If it universally appears, by experiments and astronomical observations, that all bodies about the earth gravitate towards the earth . . . in proportion to the quantity of matter that they severally contain; that the moon likewise . . . gravitates towards the earth . . . and all the planets one towards another; and the comets in like manner towards the sun; we must, in consequence of this rule, universally allow that all bodies whatsoever are endowed with a principle of mutual gravitation.

NEWTON, *Principia* (1686)

8 Universal gravitation

THE DISCOVERY OF UNIVERSAL GRAVITATION

IN CHAPTERS 6 and 7 we have built up the kind of foundation in dynamics that Newton himself was the first to establish. In a nutshell, it is the quantitative identification of force as the cause of acceleration, coupled with the purely kinematic problem of relating accelerations to velocities and displacements. We shall now consider, as a topic in its own right, the first and most splendid example of how a *law of force* was deduced from the study of motions.

It is convenient, and historically not unreasonable, to consider separately three aspects of this great discovery:

1. The analysis of the data concerning the orbits of the planets around the sun, to the approximation that these orbits are circular with the sun at the center. Several people besides Newton were closely associated with this problem.
2. The proof that gravitation is universal, in the sense that the law of force that governs the motion of objects near the earth's surface is also the law that controls the motion of celestial bodies. It seems clear that Newton was the true discoverer of this result, through his analysis of the motion of the moon.
3. The proof that the true planetary orbits, which are ellipses rather than circles, are explained by an inverse-square law of force. This achievement, certainly, was the product of Newton's genius alone.

In the present chapter we shall be able to discuss the first of these questions quite fully, using only our basic results in the kinematics and dynamics of particles. The second question requires us to learn (as Newton himself originally had to) how to analyze the gravitating properties of a body, like the earth, which is so obviously not a geometrical point when viewed from close to its surface. We shall present one approach to the problem here and complete the story in Chapter 11, where this special feature of the gravitational problem is discussed. The third question, concerning the exact mathematical description of the orbits, is something that we shall not go into at all at this stage; such orbit problems will be the exclusive concern of Chapter 13.

THE ORBITS OF THE PLANETS

We have described in Chapter 2 how the knowledge of the motions of the classical planets—Mercury, Venus, Mars, Jupiter, and Saturn—was already exceedingly well developed by the time of the astronomer Ptolemy around 150 A.D. By this we mean that the angular positions of these planets as a function of time had been catalogued with remarkable accuracy and over a long enough span for their periodic returns to the same position in the sky to be extremely well known. We have pointed out previously, however, that the interpretation of such results depends on the model of the solar system that one uses. Let us now look more carefully at the original observational data and the conclusions that can be drawn from them.

The first thing to recognize is that, whether or not one accepts the earth as the real center of the universe, it *is* the center as far as all primary observations are concerned. From this vantage point, the motion of each planet can be described, to a first approximation, as a small circle (the epicycle) whose center moves around a larger circle (the deferent). Now there are some facts about the motions of two particular planets—Mercury and Venus—that point the way to some far-reaching conclusions. These are

1. That for these two planets, the time for the center C of the epicycle [Fig. 8–1(a)] to travel once around the deferent is exactly 1 solar year—i.e., the same time that it takes the sun to complete one circuit around the ecliptic.

2. The planets Mercury and Venus never get far from the sun. They are always found within a limited angular range from the line joining the earth to the sun (about $\pm 22\frac{1}{2}°$ for Mercury,

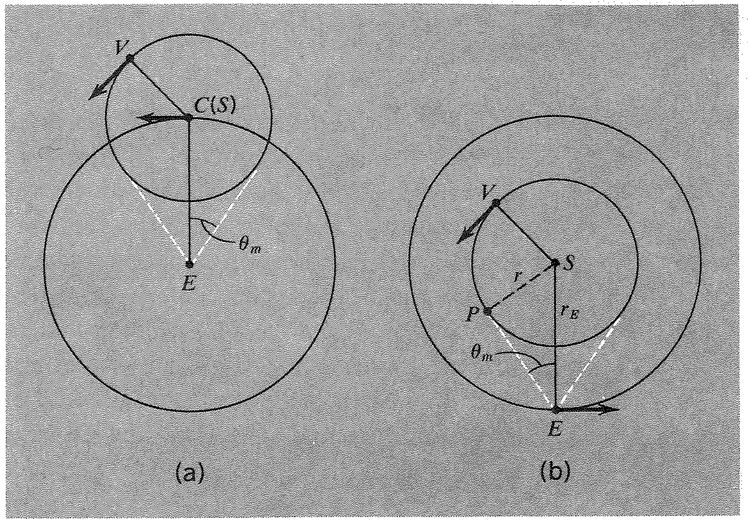


Fig. 8–1 (a) Motions of the sun and Venus as seen from the earth. Venus always lies within the angular range $\pm\theta_m$ of the sun's direction. (b) Heliocentric picture of the same situation.

$\pm 46°$ for Venus). Both of these facts are beautifully accounted for if we go over to the heliocentric, Copernican system [Fig. 8–1(b)]. We see that the larger circle of Fig. 8–1(a) corresponds in this case to the earth's own orbit around the sun, of radius r_E, and the smaller circle—the epicycle—represents the orbit of the other planet (Venus or Mercury, as the case may be). Given this interpretation, we can proceed to make quantitative inferences about the radii of the planetary orbits themselves. This is a crucial advance of the Copernican scheme over the Ptolemaic. Although Ptolemy had excellent data, they were for him just the source of purely geometric parameters, but with Copernicus we arrive at the basis of a truly physical model. Thus in Fig. 8–1(b) the maximum angular deviation, θ_m, of the planet P from the earth–sun line ES defines the planet's orbit radius r by the equation

$$\frac{r}{r_E} = \sin\theta_m \qquad (r_E > r) \tag{8–1a}$$

The radius of the earth's orbit is clearly a natural unit for measuring other astronomical distances, and has long been used for this purpose:

$$1 \text{ astronomical unit (AU)} = \text{mean distance from earth to sun } (1.496 \times 10^{11} \text{ m})$$

In terms of this unit, we then have

For Mercury: $r \approx \sin 22\frac{1}{2}° \approx 0.38$ AU
For Venus: $r \approx \sin 46° \approx 0.72$ AU

When it comes to the other planets (Mars, Jupiter, and Saturn) the tables are turned. These planets are not closely linked to the sun's position; they progress through the full 360° with respect to the earth–sun line. This can be readily explained if we interchange the roles of the two component circular motions, so that the large primary circle (the deferent) is taken to be the orbit of the planet, now larger than that of the earth, and the epicycle is seen as the expression of the earth's orbit around the sun. In the case of Jupiter, for example, the Ptolemaic picture is represented by Fig. 8–2(a) and the Copernican picture by Fig. 8–2(b). Thus the periodic angular swing, $\pm\theta_m$, of the epicycle is now related to the ratio of orbital radii through the equation

$$\frac{r_E}{r} = \sin \theta_m \qquad (r_E < r) \tag{8–1b}$$

in which the roles of r and r_E are reversed with respect to Eq. (8–1a). Ptolemy's recorded values of θ_m for Mars, Jupiter, and Saturn were about 41°, 11°, and 6°, respectively. These would then lead to the following results:

For Mars: $r \approx \csc 41° \approx 1.5$ AU
For Jupiter: $r \approx \csc 11° \approx 5.2$ AU
For Saturn: $r \approx \csc 6° \approx 9.5$ AU

Thus with the Copernican scheme (and this was its great triumph) it became possible to use the long-established data to construct

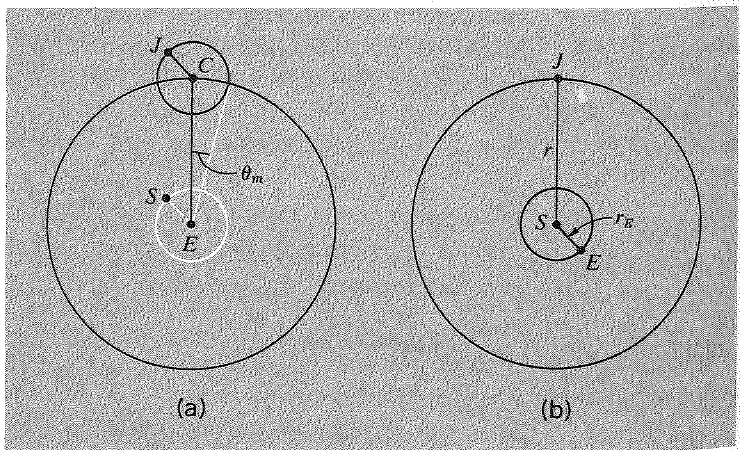


Fig. 8–2 (a) Motions of the sun and Jupiter as seen from the earth. The angle θ_m here characterizes the magnitude of the retrograde (epicyclic) motion. (b) Heliocentric picture of the same situation.

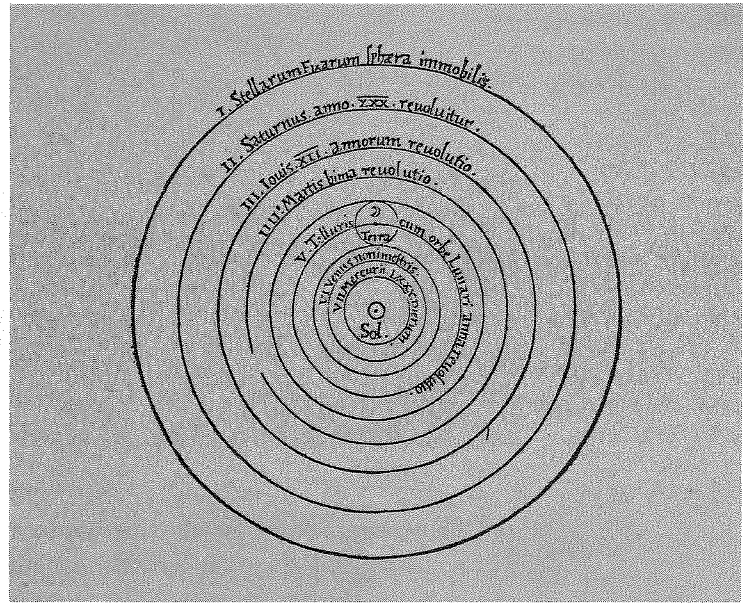


Fig. 8–3 Universe according to Copernicus. (Reproduced from his historic work, De Revolutionibus.)

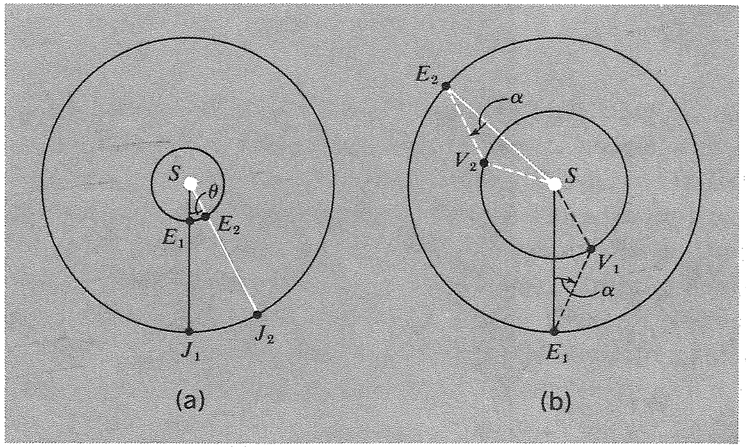
a picture of the planets in their orbits in order of their increasing distance from the sun. Figure 8–3 is a reproduction of the historic diagram by which Copernicus displayed the results in his book (*De Revolutionibus*) in 1543.

The data with which Copernicus worked (and Ptolemy, too, 1400 years before him) were actually far too good to permit a simple picture of the planets describing circular paths at constant speed around a common center. Thus Copernicus carried out a detailed analysis to find out how far the center of the orbit of each planet was offset from the sun. But even with this adjustment, the detailed change with time of the angular positions of the planets could not be fitted unless the motion around the orbit was made nonuniform. Copernicus, like Ptolemy before him, introduced auxiliary circular motions to deal with the problem, but this, as we now know, was not the answer and we shall not discuss its complexities. For the moment we shall use the basic idealization of uniform circular orbits and set aside until later the refinements that were first mastered by Kepler when he recognized the planetary paths as being ellipses.

PLANETARY PERIODS

The problem of determining the periodicities of the planets, like that of finding the shapes of their orbits, must begin with what

Fig. 8–4 *(a) Relative positions of the sun, the earth, and Jupiter at the beginning (SE_1J_1) and end (SE_2J_2) of one synodic period. (b) Comparable diagram for the sun, the earth, and Venus, allowing for the fact that Venus must be offset from the line between sun and earth if it is to be visible.*



can be observed from the moving platform that is our earth. The recurring situation that can be most easily recognized is the one in which the sun, the earth, and another planet return, after some characteristic time, to the same positions relative to one another. The length of this recurrence time is known as the *synodic period* of the planet in question. In terms of a heliocentric model of the solar system, this is easily related to the true (sidereal) period of one complete orbit of the planet around the sun.

Consider first the case of one of the outer planets, say Jupiter. Figure 8–4(a) shows a situation that can be observed from time to time. The positions of the sun, the earth, and Jupiter lie in a straight line. Observationally this could be established by finding the date on which Jupiter passes across the celestial meridian at midnight, thus placing it 180° away from the sun.[1]

Now if one such alignment is represented by the positions E_1 and J_1 of the earth and Jupiter, the next one will occur rather more than 1 year later, when the earth has gained one whole revolution on Jupiter. This is shown by the positions E_2 and J_2. Jupiter has traveled through the angular distance θ while the earth goes through $2\pi + \theta$. Both Ptolemy and Copernicus knew

[1]The celestial meridian is the projection, on the celestial sphere, of a plane containing the earth's axis and the point on the earth's surface where the observer is located. It is thus a great circle on the celestial sphere, running from north to south through the observer's zenith point, vertically above him. Noon is the instant at which the sun crosses this celestial meridian in its daily journey from east to west.

that the length of the synodic period separating these two configurations is close to 399 days. Let us denote the synodic period in general by the symbol τ. Then if the earth makes n_E complete revolutions per unit time and Jupiter makes n_J revolutions per unit time, we have

$$n_E\tau = n_J\tau + 1$$

But n_E and n_J are the reciprocals of the periods of revolution T_E and T_J of the two planets. Thus we have

$$\frac{\tau}{T_E} = \frac{\tau}{T_J} + 1$$

and solving this for T_J we have

$$T_J = \frac{T_E}{1 - T_E/\tau} \qquad (8\text{–}2a)$$

Putting $T_E/\tau \approx 365/399 \approx 0.915$, we thus find that

$$T_J \approx \frac{T_E}{0.085} \approx 11.8 \text{ yr}$$

The same type of observation and calculation can be applied to Mars and Saturn and the other outer planets that we now know. When we come to Venus and Mercury, however, the situation, as with the determination of orbital radii, is a little different. First is the practical difficulty that we cannot, at least with the naked eye, see these planets when they are in line with the sun, because it would require looking directly toward the sun to do so. We can easily get around this by considering any other situation [see Fig. 8–4(b)] in which the angle between the directions *ES* and *EV* is measured. This particular diagram shows Venus as a morning star, appearing above the horizon 1 hr or so before the sun as the earth rotates from west to east. The same value of the angle α will recur after one synodic period. This takes over $1\frac{1}{2}$ yr—about 583 days, to be more precise. In this case, however, it is Venus that has gained one revolution on the earth. Thus instead of the form of the equation that applies to the outer planets, we now have

$$n_V\tau = n_E\tau + 1$$

leading to the result

$$T_V = \frac{T_E}{1 + T_E/\tau} \qquad (8\text{–}2b)$$

Putting $T_E/\tau \approx 365/583 \approx 0.627$, we find that

$$T_V \approx \frac{T_E}{1.627} \approx 224 \text{ days}$$

It is a curious fact that Copernicus, in the introductory general account of his model of the solar system, quotes values of the planetary periods which are so rough that some of them could even be called wrong. These values are marked on his diagram (Fig. 8–3) and are repeated in his text: Saturn, 30 yr; Jupiter, 12 yr; Mars, 2 yr; Venus, 9 months; Mercury, 80 days. The worst cases are Mars (2 yr instead of about $1\frac{1}{2}$) and Venus (9 months instead of about $7\frac{1}{2}$). This seems to have led some people to think that Copernicus had only a crude knowledge of the facts, which was certainly not the case. Perhaps he was careless about quoting the periods because his real interest was in the geometrical details of the planetary orbits and distances. The truth of the matter, in any event, is that his quantitative knowledge of both periods and radii, as spelled out in detail later in his book, was so good that the best modern values do not, for the most part, differ significantly from the ones he quoted. This is shown in Table 8–1, which lists both the Copernican and the modern data on the classical planets. (Incidentally, the values to be extracted from Ptolemy's data are almost iden-

TABLE 8–1: DATA ON PLANETARY ORBITS

	Orbital radius, AU		*Synodic period, days*	*Sidereal period*	
Planet	*Copernicus*	*Modern*	*Copernicus*	*Copernicus*	*Modern*
Mercury	0.376	0.3871	115.88	87.97 days	87.97 days
Venus	0.719	0.7233	538.92	224.70 days	224.70 days
Earth	1.000	1.0000	—	365.26 days	365.26 days
Mars	1.520	1.5237	779.04	1.882 yr	1.881 yr
Jupiter	5.219	5.2028	398.96	11.87 yr	11.862 yr
Saturn	9.174	9.5389	378.09	29.44 yr	29.457 yr

tical with those of Copernicus, an astonishing tribute to those astronomers whose measurements, from about 750 B.C. up to the time of Ptolemy's own observations around 130 A.D., provided the basis of his analysis.)

KEPLER'S THIRD LAW

The data of Table 8–1 point clearly to a systematic relationship between the planetary periods and distances. This is displayed

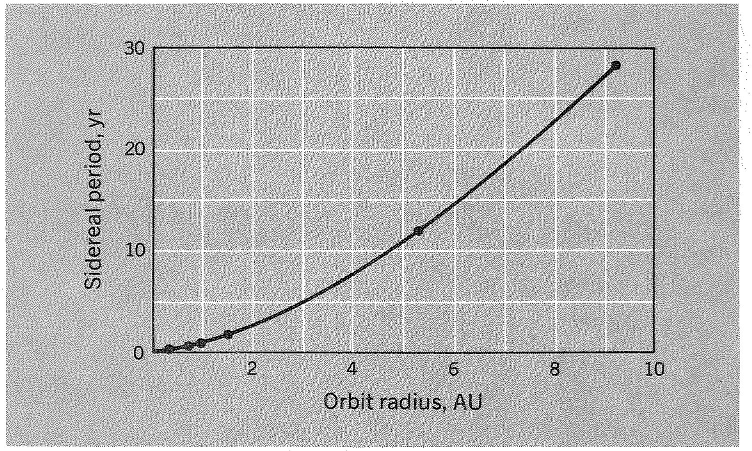


Fig. 8–5 Smooth curve relating the periods and the orbital radii of the planets.

graphically in Fig. 8–5. The precise form of the relationship was first discovered by Johann Kepler in 1618 and published by him the following year in his book *The Harmonies of the World.* In it he triumphantly wrote: "I first believed I was dreaming . . . But it is absolutely certain and exact that the ratio which exists between the periodic times of any two planets is precisely the ratio of the $\frac{3}{2}$th powers of the mean distances . . . ." Table 8–2

TABLE 8–2: KEPLER'S THIRD LAW

Planet	*Radius r of orbit of planet, AU*	*Period T, days*	r^3/T^2, $(AU)^3/(day)^2 \times 10^{-6}$
Mercury	0.389	87.77	7.64
Venus	0.724	224.70	7.52
Earth	1.000	365.25	7.50
Mars	1.524	686.98	7.50
Jupiter	5.200	4,332.62	7.49
Saturn	9.510	10,759.20	7.43

shows the data used by Kepler and a test of the near constancy of the ratio r^3/T^2. Figure 8–6 is a different presentation of the planetary data (actually in this case the data of Copernicus from Table 8–1) plotted in modern fashion on log-log graph paper so as to show this relationship:

$$T \sim r^{3/2} \tag{8–3}$$

This is known as Kepler's third law, having been preceded, 10 years earlier, by the statement of his two great discoveries (quoted

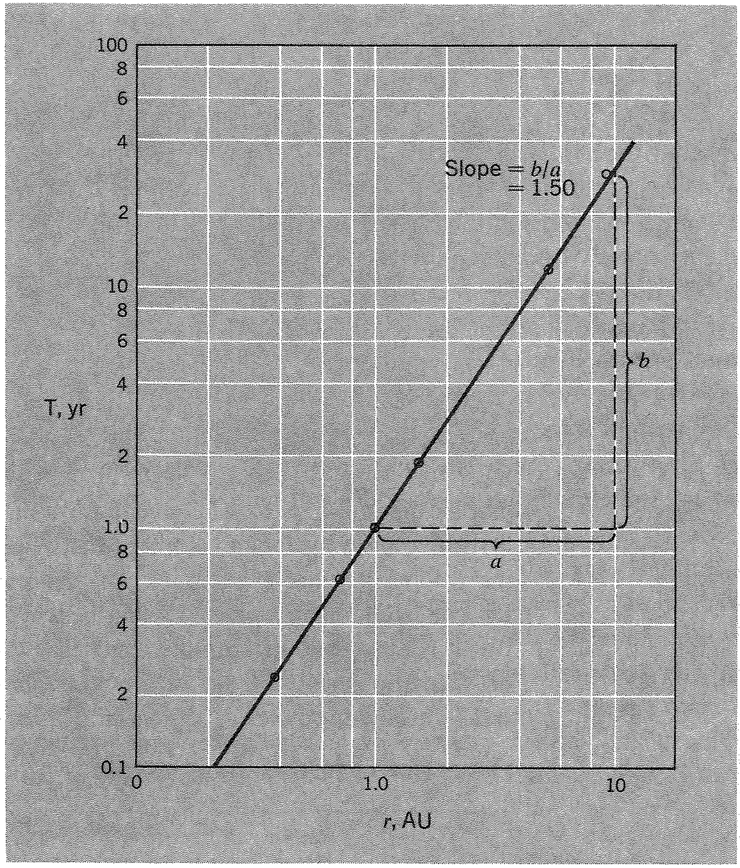


Fig. 8–6 Log-log plot of planetary period T versus orbit radius r, using data quoted by Copernicus. The graph shows that T is proportional to $r^{3/2}$ (Kepler's third law).

in the Prologue) concerning the elliptical paths of the individual planets.

The dynamical explanation of Kepler's third law had to await Newton's discussion of such problems in the *Principia.* A very simple analysis of it is possible if we again use the simplified picture of the planetary orbits as circles with the sun at the center. It then becomes apparent that Eq. (8–3) implies that an inverse-square law of force is at work. For in a circular orbit of radius r we have

$$a_r = -\frac{v^2}{r}$$

Expressing v in terms of the known quantities, r and T, we have

$$v = \frac{2\pi r}{T}$$

$$a_r = -\frac{4\pi^2 r}{T^2} \quad \text{(toward the center)} \tag{8–4}$$

From Newton's law, then, we infer that the force on a mass in a circular orbit must be given by

$$F_r = ma_r = -\frac{4\pi^2 mr}{T^2} \tag{8–5}$$

From Kepler's third law, however, we have the relation

$$\frac{r^3}{T^2} = K \tag{8–6}$$

where K might be called Kepler's constant—the same value of it applies to all the planets traveling around the sun. From Eq. (8–6) we thus have $1/T^2 = K/r^3$, and substituting this in Eq. (8–5) gives us

$$F_r = -\frac{4\pi^2 Km}{r^2} \tag{8–7}$$

The implication of Kepler's third law, therefore, when analyzed in terms of Newton's dynamics, is that the force on a planet is proportional to its inertial mass m and inversely proportional to the square of its distance from the sun. Newton's contemporaries Halley, Hooke, and Huygens all appear to have arrived at some kind of formulation of an inverse-square law in the planetary problem, although Newton's, in terms of his definite concept of forces acting on masses, seems to have been the most clear-cut. The general idea of an influence falling off as $1/r^2$ was probably not a great novelty, for it is the most natural-seeming of all conceivable effects—something spreading out and having to cover spheres of larger and larger area, in proportion to r^2, so that the intensity (as with light from a source) gets weaker according to an inverse-square relationship.

The proportionality of the force to the attracted mass, as required by Eq. (8–7), was a feature of which only Newton appreciated the full significance. With his grasp of the concept of interactions exerted mutually between pairs of objects, Newton saw that the reciprocity in the gravitational interaction must mean that the force is proportional to the mass of the attracting object just as it is to the mass of the attracted. Each object is the attracting agent as far as the other one is concerned. Hence the magnitude of the force exerted on either one of a mutually gravitating pair of particles must be expressed in the famous mathematical statement of universal gravitation:

$$F = -\frac{Gm_1m_2}{r^2} \tag{8-8}$$

where G is a constant to be found by experiment, and m_1 and m_2 are the inertial masses of the particles. We shall return to the matter of determining G in practice, but first we shall discuss the famous problem that led Newton toward some of his greatest discoveries concerning gravitation.

THE MOON AND THE APPLE

It is an old story, but still an enthralling one, of how Newton, as a young man of 23, came to think about the motion of the moon in a way that nobody had ever done before. The path of the moon through space, as referred to the "fixed" stars, is a line of varying curvature (always, however, bending toward the sun), which crosses and recrosses the earth's orbit. But of course there is a much more striking way of looking at it—the familiar earth-centered view, which shows the moon describing an approximately circular orbit around the earth. To this extent it is quite like the planetary-orbit problem that we have just been discussing. But Newton, with his extraordinary insight, constructed an intellectual bridge between this motion and the behavior of falling objects—the latter being such a commonplace phenomenon that it needed a genius to recognize its relevance. He saw the moon as being just an object falling toward the earth like any other—as, for example, an apple dropping off a tree in his garden. A very special case, to be sure, because the moon was so much farther away than any other falling object in our experience. But perhaps it was all part of the same pattern.

As Newton himself described it,[1] he began in 1665 or 1666 to think of the earth's gravity as extending out to the moon's orbit, with an inverse-square relationship already suggested by Kepler's third law. We could of course just restate the centripetal acceleration formula and apply it to the moon, but it is illuminating to trace the course of Newton's own way of discussing the problem. In effect he said this: Imagine the moon at any point A in its orbit (Fig. 8–7). If freed of all forces, it

[1]See the Prologue of this book.

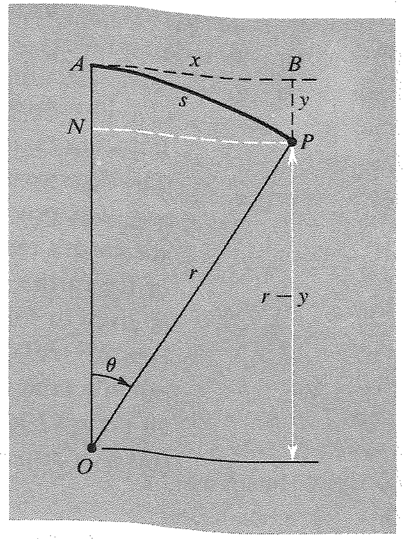


Fig. 8–7 Geometry of a small portion of a circular orbit, showing the deviation y from the tangential straight-line displacement AB (=x) that would be followed in the absence of gravity.

would travel along a straight line AB, tangent to the orbit at A. Instead, it follows the arc AP. If O is the center of the earth, the moon has in effect "fallen" the distance BP toward O, even though its radial distance r is unchanged. Let us calculate how far the moon falls, in this sense, in 1 sec, and compare it with the distance of about 16 ft that an object projected horizontally near the earth's surface would fall in that same time.

First, a bit of analytic geometry. If we denote the distance AB as x, and the distance BP as y, it will be an exceedingly good approximation to put

$$y \approx \frac{x^2}{2r} \tag{8-9}$$

One way of obtaining this result is to consider the right triangle ONP, in which we have

$$ON = r - y \qquad NP = x \qquad OP = r$$

Hence, by Pythagoras' theorem,

$$(r - y)^2 + x^2 = r^2$$
$$x^2 = 2ry - y^2$$

Since $y \ll r$ for any small value of the angle θ, Eq. (8–9) follows as a good approximation. Furthermore, since (again for small θ) the arc length AP ($= s$) is almost equal to the distance AB, we can equally well put

$$y \approx \frac{s^2}{2r} \tag{8-10}$$

In order to put numbers into this formula we need to know both the radius and the period of the moon's orbital motion. The distance to the moon, as known to Newton, depended on the two-step process devised by the ancient astronomers—finding the earth's radius and finding the moon's distance as a multiple of the earth's radius. A reminder of these classic measurements is given in Figs. 8–8 and 8–9 and the accompanying discussion (pp. 259–261). The final result, familiar to everyone, is that the moon's orbit radius r is about 240,000 miles $\approx 3.8 \times 10^8$ m. Its period T is 27.3 days $\approx 2.4 \times 10^6$ sec. Therefore, in 1 sec it travels a distance along its orbit given by

$$\text{(in 1 sec)} \qquad s = \frac{2\pi \times 3.8 \times 10^8}{2.4 \times 10^6} \approx 1000 \text{ m}$$

During this same time it falls a vertical distance, which we will denote y_1 to identify it, given [via Eq. (8–10)] by

$$\text{(in 1 sec)} \qquad y_1 \approx \frac{10^6}{7.6 \times 10^8} = 1.3 \times 10^{-3} \text{ m}$$

In other words, in 1 sec, while traveling "horizontally" through a distance of 1 km, the moon falls vertically through just over 1 mm, or about $\frac{1}{20}$ in.; its deviation from a straight-line path is indeed slight. On the other hand, for an object near the earth's surface, projected horizontally, the vertical displacement in 1 sec is given by

$$y_2 = \tfrac{1}{2}gt^2 = 4.9 \text{ m}$$

Thus

$$\frac{y_1}{y_2} = 2.7 \times 10^{-4} \approx \frac{1}{3700}$$

Newton knew that the radius of the moon's orbit was about 60 times the radius of the earth itself, as the ancient Greeks had first shown. And with an inverse-square law, if it applied equally well at all radial distances from the earth's center, we would expect y_1/y_2 to be about 1/3600. It *must* be right! And yet, what an astounding result. Even granted an inverse-square law of attraction between objects separated by many times their diameters, one still has the task of proving that an object a few feet above the earth's surface is attracted as though the whole mass of the earth were concentrated at a point 4000 miles below the ground. Newton did not prove this result until 1685, nearly 20 years after his first great insight into the problem. He published nothing, either, until it all came out, perfect and complete, in the *Principia* in 1687. One way of solving the problem follows on p. 262 (after the special section below).

FINDING THE DISTANCE TO THE MOON

The earth's radius

About 225 B.C. Eratosthenes, who lived and worked at Alexandria near the mouth of the Nile, reported on measurements made on the shadows cast by the sun at noon on midsummer day. At Alexandria (marked A in Fig. 8–8) the sun's rays made an angle of 7.2° to the local vertical, whereas corresponding measurements made 500 miles farther south at Syene (now the site of the Aswan Dam) showed the sun to be exactly overhead at noon. (In other words, Syene lay almost exactly on the Tropic of Cancer.) It follows at once from these figures that the arc AS, of length 500 miles, subtends an angle of 7.2° or $\frac{1}{8}$ rad at the center of the earth. Hence

$$\frac{500}{R_E} = \frac{1}{8}$$

or

$$R_E \approx 4000 \text{ miles}$$

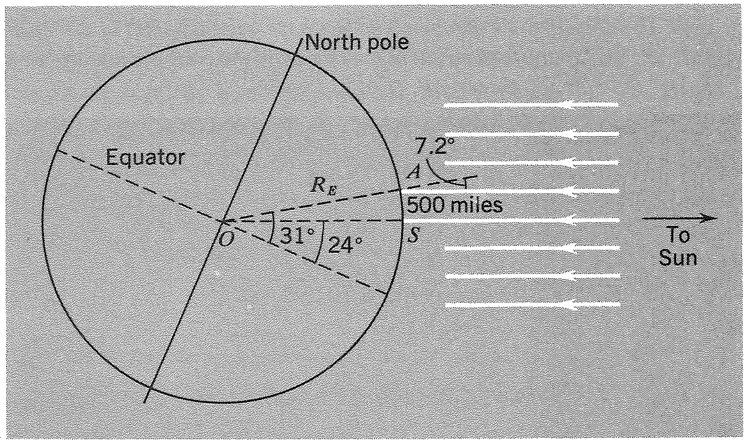


Fig. 8–8 Basis of the method used by Eratosthenes to find the earth's radius. When the midday sun was exactly overhead at Syene (S) its rays fell at 7.2° to the vertical at Alexandria (A).

The moon's distance measured in earth radii

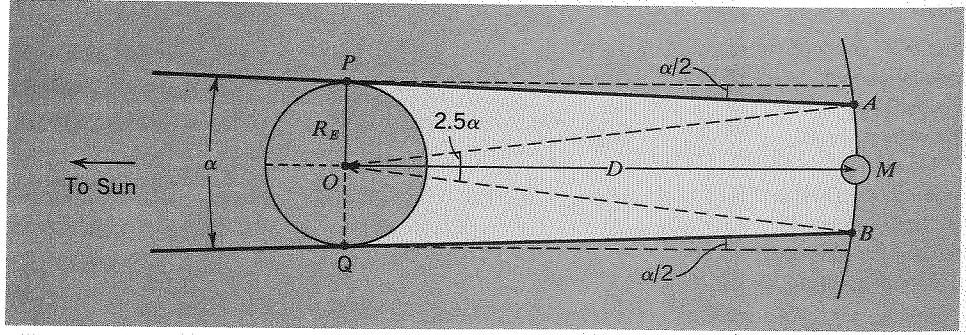
Hipparchus, a Greek astronomer who lived mostly on the island of Rhodes, made observations in about 130 B.C. from which he obtained a remarkably accurate estimate of the moon's distance. His method was one suggested by another great astronomer, Aristarchus, about 150 years earlier.

The method involves a clear understanding of the positional relationships of sun, earth, and moon. We know that sun and moon subtend almost exactly the same angle α at the earth. Hipparchus measured this angle to be 0.553° ($\approx 1/103.5$ rad); he also knew what Aristarchus before him had found—that the sun is far more distant than the moon. Hipparchus used this knowledge in an analysis of an eclipse of the moon by the earth (Fig. 8–9). The shaded region indicates the area that is in complete shadow; its boundary lines PA and QB make an angle α with one another, because this is the angle between rays coming from the extreme edges of the sun. The moon passes through the shadowed region, and from the measured time that this passage took, Hipparchus deduced that the angle subtended at the earth by the arc BA was 2.5 times that subtended by the moon itself. Thus $\angle AOB \approx 2.5\alpha$.

Let us now do some geometry. If the distance from the earth's center to the moon is D, the length of the arc BA is very nearly equal to the earth's diameter PQ diminished by the amount αD:

$$AB \approx 2R_E - \alpha D$$

Fig. 8–9 Basis of the method used by Hipparchus to find the moon's distance. The method depended on observing the duration (and hence the angular width) of the moon's total eclipse in the shadow of the earth, as represented by the arc AB.



But we also have

$$\frac{AB}{D} \approx 2.5\alpha$$

or

$$AB \approx 2.5\alpha D$$

Substituting this in the first equation we have

$$3.5\alpha D \approx 2R_E$$

or

$$\frac{D}{R_E} \approx \frac{2}{3.5\alpha}$$

Since $\alpha \approx 1/103.5$ rad, this gives

$$\frac{D}{R_E} \approx \frac{207}{3.5} \approx 59$$

Combining this with the value of R_E itself, we have

$$D \approx 236{,}000 \text{ miles}$$

Modern methods

Refined triangulation techniques give a mean value of 3,422.6″, or 0.951°, for the angle subtended at the moon by the earth's radius. Using the modern value of the earth's radius

$$(R_E = 6378 \text{ km} = 3986 \text{ miles})$$

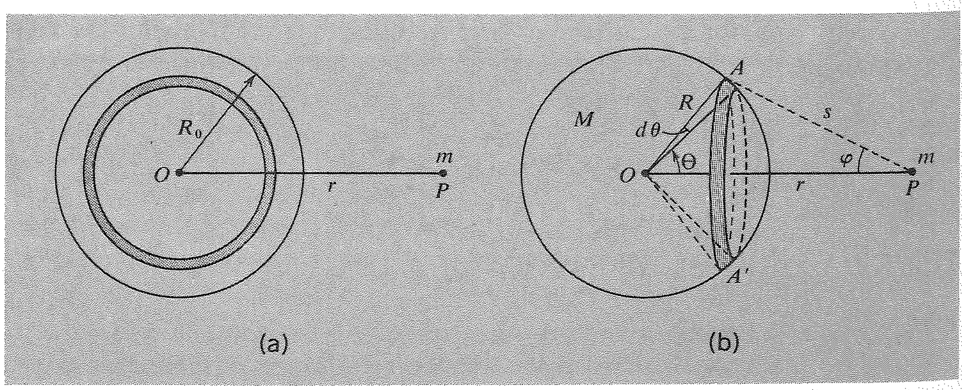
one obtains almost exactly 240,000 miles for the moon's mean distance. Such traditional methods, however, are far surpassed by the technique of making a precision measurement of the time for a radar echo or laser reflection to return to earth. The flight time of such signals (only about 2.5 sec for the roundtrip) can be measured to a fraction of a microsecond, giving range determinations that are not only of unprecedented accuracy but are also effectively instantaneous.

THE GRAVITATIONAL ATTRACTION OF A LARGE SPHERE

It has long been suggested that Newton's failure to publicize his discovery about an inverse-square law of the earth's gravity extending to the moon was due primarily to an actual numerical discrepancy, resulting from his use of an erroneous value for the earth's radius. This would, via Eq. (8–10), falsify the value of the moon's distance of fall, since r (the radius of the moon's orbit) was calculated, according to the method discovered by Hipparchus, in terms of the earth's radius. When Newton first did the calculation he was home in the countryside, out of reach of reference books, and it is reliably recorded that he calculated the earth's radius by assuming that 1° of latitude is 60 miles, instead of the correct figure of nearly 70 miles. Be this as it may, it remains almost certain that Newton, with his outstandingly thorough and critical approach to problems, would never have regarded the theory as complete until he had solved the problem of gravitation by large objects. Let us now consider a way of analyzing this problem. (In Chapter 11 we shall tackle it again in a more sophisticated way.)

Suppose we have a large solid sphere, of radius R_0, as shown in Fig. 8–10(a), and wish to calculate the force with which it attracts a small object of mass m at an arbitrary point P. We shall assume that the density of the material of the sphere may vary with distance from the center (as is the case for the earth, to a very marked degree) but that the density is the same at all points equidistant from the center. We can then consider the solid sphere to be built up of a very large number of thin uniform

Fig. 8–10 (a) A solid sphere can be regarded as built up of a set of thin concentric spherical shells. (b) The gravitational effect of an individual shell can be found by treating it as an assemblage of circular zones.



spherical shells, like the successive layers of an onion. The total gravitational effect of the sphere can be calculated as the superposition of the effects of all these individual shells. Thus the basic problem becomes that of calculating the force exerted by a thin spherical shell of arbitrary radius, assuming that the fundamental law of force is that of the inverse square between point masses.

In Fig. 8–10(b) we show a shell of mass M, radius R, of negligible thickness, with a particle of mass m at a distance r from the center of the shell. If we consider a small piece of the shell, near point A, the force that it exerts on m is along the line AP. It is clear from the symmetry of the system, however, that the resultant force due to the whole shell must be along the line OP; any component of force transverse to OP due to material near A will be canceled by an equal and opposite contribution from material near A'. Thus if we have an element of mass dM near A, we need only consider its contribution to the force along OP, i.e., the radial direction from the center of the shell to m. Hence we have

$$dF_r = -\frac{Gm\,dM}{s^2}\cos\varphi \qquad (8\text{–}11)$$

Let us now consider a complete belt or zone of the shell, shown shaded in the diagram. It represents the portion of the shell that is contained between the directions θ and $\theta + d\theta$ to the axis OP, and the same mean values of s and φ apply to every part of it. Thus, if we calculate its mass, we can substitute this value as dM in Eq. (8–11) to obtain the contribution of the whole belt to the resultant gravitational force along OP. Now the width of the belt is $R\,d\theta$ and its circumference is $2\pi R \sin\theta$; thus its area is $2\pi R^2 \sin\theta\,d\theta$. The area of the whole shell is $4\pi R^2$; hence the mass of the belt is given by

$$dM = \frac{2\pi R^2 \sin\theta\,d\theta}{4\pi R^2}\,M = \frac{M}{2}\sin\theta\,d\theta$$

Thus Eq. (8–11) gives us

$$dF_r = -\frac{GMm}{2}\,\frac{\cos\varphi\sin\theta\,d\theta}{s^2} \qquad (8\text{–}12)$$

Our task now is to sum the contributions such as dF_r over the whole of the shell, i.e., over the whole range of values of s, φ, and θ. This looks like a formidable task, but with the help of a little calculus (another of Newton's inventions!) the solution turns out to be surprisingly straightforward.

From the geometry of the situation [Fig. 8–10(b)], it is possible to express both of the angles θ and φ in terms of two fixed distances, r and R, and the variable distance s. By two separate applications of the cosine rule we have

$$\cos\theta = \frac{r^2 + R^2 - s^2}{2rR} \qquad \cos\varphi = \frac{r^2 + s^2 - R^2}{2rs}$$

From the first of these, by differentiation, we have

$$\sin\theta\, d\theta = \frac{s\, ds}{rR}$$

Hence, substituting the values of $\cos\varphi$ and $\sin\theta\, d\theta$ in Eq. (8–12), we obtain

$$dF_r = -\frac{GMm}{4r^2R}\frac{(r^2 + s^2 - R^2)\, ds}{s^2}$$

The total force is found by integrating this expression from the minimum value of s $(= r - R)$ to its maximum value $(r + R)$. Thus we have

$$F_r = -\frac{GMm}{4r^2R}\int_{r-R}^{r+R}\frac{r^2 + s^2 - R^2}{s^2}\, ds \tag{8–13}$$

The integral is just the sum of two elementary forms; we have

$$\int\frac{r^2 + s^2 - R^2}{s^2}\, ds = \int ds + (r^2 - R^2)\int\frac{ds}{s^2}$$
$$= s - \frac{r^2 - R^2}{s}$$

Inserting the limits, we then find that

$$\int_{r-R}^{r+R}\frac{r^2 + s^2 - R^2}{s^2}\, ds = [(r + R) - (r - R)] - \left(\frac{r^2 - R^2}{r + R} - \frac{r^2 - R^2}{r - R}\right)$$
$$= 2R - (r - R) + (r + R)$$
$$= 4R$$

Substituting this value of the definite integral in Eq. (8–13) we have

$$F_r = -\frac{GMm}{r^2} \tag{8–14}$$

What a wonderful result! It is of extraordinary simplicity, and the radius R of the shell does not appear at all. It is uniquely a consequence of an inverse-square law of force between particles; no other force law would yield such a simple result for the net effect of an extended spherical object.

Once we have Eq. (8–14), the total effect of a solid sphere follows at once. Regardless of the particular way in which the density varies between the center and the surface (provided that it depends only on R) the complete sphere does indeed act as though its total mass were concentrated at its center. It does not matter how close the attracted particle P is to the surface of the sphere, as long as it is in fact outside. Take a moment to consider what a truly remarkable result this is. Ask yourself: Is it obvious that an object a few feet above the apparently flat ground should be attracted as though the whole mass of the earth (all 6,000,000,000,000,000,000,000 tons of it!) were concentrated at a point (the earth's center) 4000 miles down? It is about as far from obvious as could be, and there can be little doubt that Newton had to convince himself of this result before he could establish, to his own satisfaction, the grand connection between terrestrial gravity and the motion of the moon and other celestial objects.

OTHER SATELLITES OF THE EARTH

Newton's thinking quite explicitly embraced the possibility—at least theoretically—of having other satellites of the earth. Figure 8–11 is an illustration from Newton's book, *The System of the World* (which is incorporated in the *Principia*); it shows the transition from the effectively parabolic trajectories of short-range projectiles (although the apparent parabolas are really small parts of ellipses) to a perfectly circular orbit and then to other elliptic orbits of arbitrary dimensions.

Let us derive the formulas for the required velocity v and the period T of a satellite launched horizontally in a circular orbit at a distance r from the center of the earth. The necessary

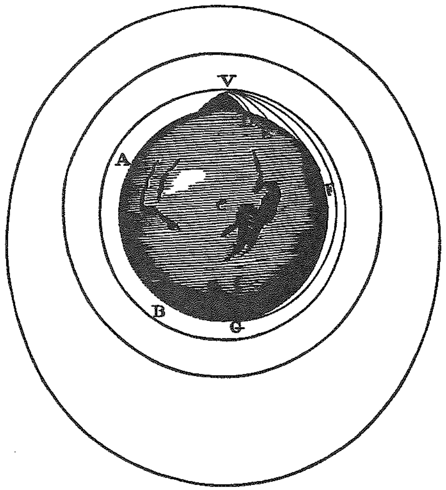


Fig. 8–11 Newton's diagram showing the transition from normal parabolic trajectories to complete orbits encircling the earth. (From The System of the World.)

force to maintain circular motion is provided by gravitational attraction:

$$\frac{mv^2}{r} = G\frac{M_E m}{r^2}$$

where M_E is the mass of the earth, m the mass of the satellite, and G the universal gravitational constant. Solving for v,

$$v = \left(\frac{GM_E}{r}\right)^{1/2} \tag{8–15}$$

It is often convenient to express this result in terms of more familiar quantities. We can do this by noticing that, for an object of mass m at the earth's surface, the gravitational force on it, by Eq. (8–8), is

$$F_g = \frac{GM_E m}{R_E^2}$$

But this is the force that can be set equal to mg for the mass in question. Hence we have

$$mg = \frac{GM_E m}{R_E^2}$$

or

$$GM_E = gR_E^2$$

Substituting this in Eq. (8–15), we get

$$v = \left(\frac{gR_E^2}{r}\right)^{1/2}$$

The period, T, of the satellite is then given by

$$T = \frac{2\pi r}{v} = \frac{2\pi r^{3/2}}{g^{1/2} R_E} \tag{8–16}$$

Putting $g = 9.8\ \text{m/sec}^2$, $R_E = 6.4 \times 10^6$ m, we have a numerical formula for the period of any satellite in a circular orbit of radius r around the earth:

$$\text{(Earth satellites)} \qquad T \approx 3.14 \times 10^{-7} r^{3/2} \tag{8–17}$$

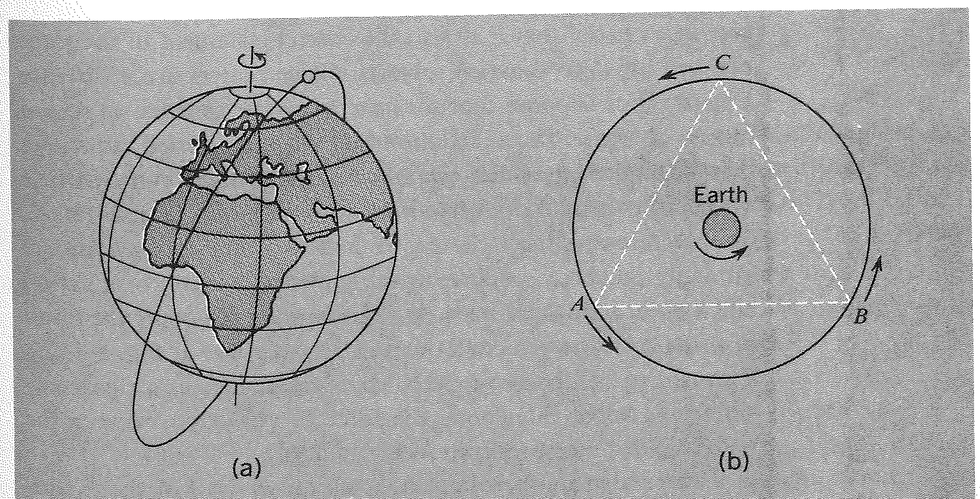
where T is in seconds and r in meters.

For example, a satellite at minimum practicable altitude (about 200 km, say), has $r \approx 6.6 \times 10^6$ m, and hence

$$T \approx 5.3 \times 10^3\ \text{sec} \approx 90\ \text{min}$$

The first man-made satellite, Sputnik I (October 1957) had an orbit as shown in Fig. 8–12(a). Its maximum and minimum distances from the earth's surface were initially 228 and 947 km, respectively, giving a mean value of r equal to about 6950 km.

Fig. 8–12 (a) Orbit of Sputnik I, the first man-made satellite (October 1957). (b) Synchronous satellite communication system. Orbital diameter in relation to earth's diameter is approximately to scale.



With this value of r, Eq. (8–17) gives an orbital period of about 96 min, which agrees closely with the observed figure.

Particular interest attaches to synchronous satellites that have an orbital period equal to the period of the earth's rotation on its axis. If placed in orbit in the earth's equatorial plane, such satellites will remain above the same spot on the earth's surface, and a set of three of them, ideally in a regular triangular array as shown in Fig. 8–12(b), can provide the basis of a worldwide communications system with no blind spots. Putting $T = 1$ day in Eq. (8–17), one finds $r \approx 42{,}000$ km or 26,000 miles. Thus such satellites must be about 22,000 miles above the earth's surface, i.e., about $5\frac{1}{2}$ earth radii overhead. The first such satellite to be successfully launched was Syncom II in July 1963.

Equation (8–16), on which the above calculations are based, has a very noteworthy feature. A satellite traveling in a circular orbit of a given radius has a period independent of the mass of the satellite. Thus a massive spaceship of many tons will, for the same value of r, have precisely the same orbital period as a flimsy object such as one of the Echo balloons, with a mass of only about 100 kg—or, for that matter, a small piece of interplanetary debris with a mass of only a few kilograms. This result is a direct consequence of the fact that the gravitational force on any object is strictly proportional to its own mass.

THE VALUE OF G, AND THE MASS OF THE EARTH

Although the result expressed by Eq. (8–14) was obtained by considering a large sphere and a small particle, one can quickly convince oneself that it is also the correct statement of the force between any two spherical objects whose centers are a distance r apart. For suppose that we have two such spheres, as shown in Fig. 8–13(a). The calculation that we have carried out shows that one sphere (say the one on the left) attracts every particle of the other as if the left-hand mass were a point [Fig. 8–13(b)]. This therefore reduces the problem to the mutual gravitational attraction between a sphere (the right-hand sphere, of mass m) and a point particle of mass M. But now we can apply the result of the last section a second time. Thus we arrive at Fig. 8–13(c), with two point masses separated by a distance r, as a rigorously correct basis for calculating the force of attraction between the two extended masses shown in Fig. 8–13(a).

The above result is important in the analysis of the experiment, already described in Chapter 5, for finding the universal

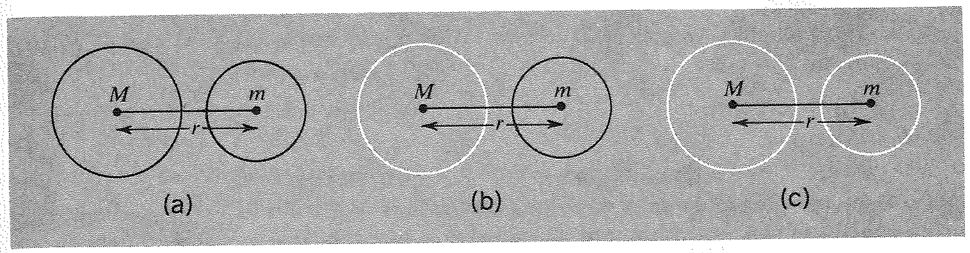


Fig. 8–13 (a) Two gravitating spheres at small separation. (b) Effect of one sphere (M) can be calculated by treating it as a point mass. (c) The argument can be repeated, so that the attraction between the spheres can be calculated as though both were point masses.

gravitation constant, G, from the measured force between two spheres of known masses. In order to get the biggest possible effect with an interaction that is so extremely weak, it is usual to arrange things so that the centers of the spheres are separated by only a little more than the sum of the radii. It is then a great convenience to be able to calculate the force, even under these conditions, on the basis of Eq. (8–14). Notice, however, that the result holds only for spheres. Some of the measurements to determine G have made use of cylindrical masses, because of the greater ease of machining them to high precision. In such cases it becomes necessary to calculate the net force by an explicit integration over the spatial distribution of material.

The presently accepted value of G, as obtained from laboratory measurements of the force exerted between two known masses, is (as already quoted in Chapter 5):

$$G = 6.670 \times 10^{-11}\ \mathrm{m^3/kg\text{-}sec^2} \qquad (8\text{–}18)$$

Newton himself did not know the value of G, although he made a celebrated guess at the mean density of the earth, from which he could have obtained a conjectural figure. In Book III of the *Principia*, he remarks at one point as follows: "Since . . . the common matter of our earth on the surface thereof is about twice as heavy as water, and a little lower, in mines, is found about three, or four, or even five times heavier, it is probable that the quantity of the whole matter of the earth may be five or six times greater than if it consisted all of water . . . ."

If we denote the mean density of the earth as ρ and its radius as R, the gravitational force exerted on a particle of mass m just at the earth's surface is given by

$$F = \frac{GMm}{R^2} \tag{8-19}$$

where

$$M = \frac{4\pi}{3}\rho R^3$$

Hence

$$F = \frac{4\pi}{3}(G\rho R)m$$

Since, however, this is just the force that gives the particle an acceleration g in free fall, we also have

$$F = mg$$

It follows, then, that

$$g = \frac{4\pi}{3}G\rho R \tag{8-20}$$

If in this equation we put $g \approx 9.8\ \text{m/sec}^2$, $R \approx 6.37 \times 10^6$ m, and (using Newton's estimate) $\rho \approx 5000$ to $6000\ \text{kg/m}^3$, we find that

$$G \approx (6.7 \pm 0.6) \times 10^{-11}\ \text{m}^3/\text{kg-sec}^2$$

Thus Newton's estimate was almost exactly on target. In practice, of course, the calculation is done the other way around. Given the directly determined value of G [Eq. (8–18)] we substitute in Eqs. (8–19) and (8–20) to find the mass and the mean density of the earth. The result of these substitutions (with $R = 6.37 \times 10^6$ m) is

$$M = 5.97 \times 10^{24}\ \text{kg}$$
$$\rho = 5.52 \times 10^3\ \text{kg/m}^3$$

LOCAL VARIATIONS OF g

If we take the idealization of a perfectly spherical, symmetrical earth, then the gravitational force on an object of mass m at a distance h above the surface is given by

$$F = \frac{GMm}{(R+h)^2}$$

If we identify F with m times the value of g at the point in question, we have

$$g(h) = \frac{GM}{(R+h)^2} \tag{8-21}$$

For $h \ll R$, this would imply an almost exactly linear decrease of g with height. Using the binomial theorem, we can rewrite Eq. (8–21) as follows:

$$g(h) = \frac{GM}{R^2}\left(1 + \frac{h}{r}\right)^{-2}$$
$$= \frac{GM}{R^2}\left(1 - \frac{2h}{R} + \frac{3h^2}{R^2}\cdots\right)$$

Hence, for small h, we have

$$g(h) \approx g_0\left(1 - \frac{2h}{R}\right) \tag{8-22}$$

where $g_0 = GM/R^2$, the value of g at points extremely close to the earth's surface. [Alternatively, we can use a calculus method that can be extremely useful whenever we want to consider the *fractional* variation of a quantity. It is based on the fact that the differentiation of the natural logarithm of a quantity leads at once to the fractional variation. In the present case we have

$$g(r) = \frac{GM}{r^2}$$

Therefore,

$$\ln g = \text{const.} - 2 \ln r$$

Differentiating,

$$\frac{\Delta g}{g} \approx -2\frac{\Delta r}{r}$$

Hence, putting $r = R$, $g = g_0$, and $\Delta r = h$, we have

$$\Delta g \approx -2g_0\frac{h}{r}$$

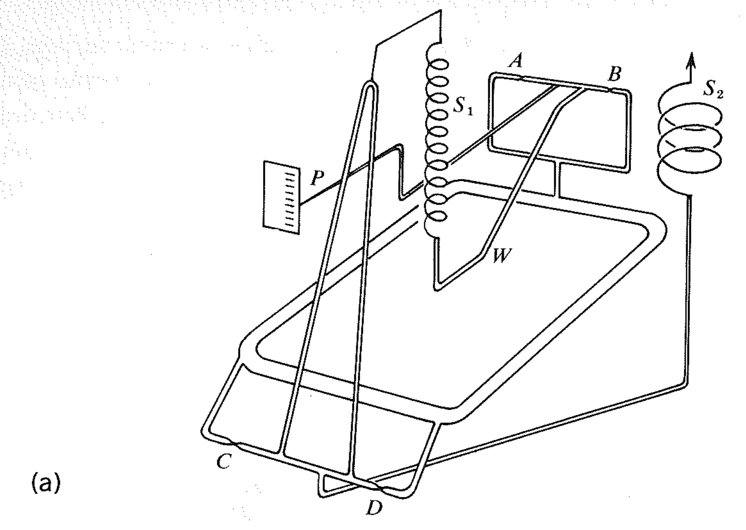
which leads us back to Eq. (8–22). Notice how this method frees us of the need to concern ourselves with the values of any multiplicative constants that appear in the original equation—e.g., the value of GM in the equation for g. A recognition of this fact can enable one to avoid a lot of unnecessary arithmetic in the computation of small changes of one quantity that depends upon another according to some well-defined functional relationship.]

Newton's contemporary, Robert Hooke, made several efforts to detect a variation of the gravitational attraction with height. He did this by looking for any changes in the measured weights of objects at the tops of church towers and the bottoms of deep wells. Not surprisingly, he was unable to find any difference. By Eq. (8–22) one would have to ascend to a point about 1000 ft above ground (e.g., the top of the Empire State Building) before the decrease of g was even as great as 1 part in 10,000. As we shall see in a moment, however, such variations are detected with the greatest of ease by modern techniques.

Superimposed on the systematic variations of the gravitational force with height are the variations produced by inhomogeneities in the material of the earth's crust. For example, if one is standing above a subterranean deposit of salt or sand, much lower in density than ordinary rocks, one would expect the value of g to be reduced. Such changes, although extremely small, can be measured with remarkable accuracy by modern instruments and have become a very valuable tool in geophysical prospecting.

Almost all modern gravity meters are static instruments, in which a mass is in equilibrium under the combined action of gravity and an elastic restoring force supplied by a spring. In other words, it is just a very sensitive spring balance. A change in g as the instrument is moved from one point to another leads to a minute change in the equilibrium position, and this is detected by sensitive optical methods or electrically by, for example, making the suspended mass part of a tuned circuit whose capacitance, and hence frequency, is changed by the slight displacement. To be useful, such instruments must be capable of detecting fractional changes of g of 10^{-7} or less. Figure 8–14(a) shows the basic construction of one such device. With it one can trace out contours of constant g over a region of interest. Figure 8–14(b) shows the results of such a survey, after allowance has been made for effects due to varying altitude, surface features, and so on. Such contours can give good indications of ore concentrations. The primary unit of measurement in these gravity surveys is known as the *gal* (after Galileo):

$$1 \text{ gal} = 1 \text{ cm/sec}^2 \approx 10^{-3}\, g$$



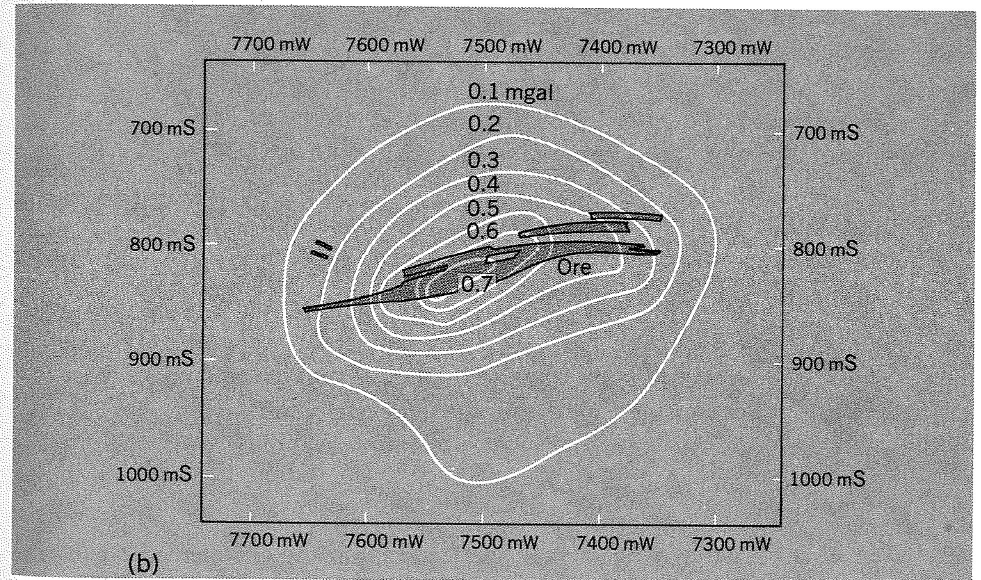


Fig. 8–14 (a) Sketch of basic features of a sensitive gravimeter, made of fused quartz. The arm marked W acts as the main weight. It is pivoted at A and B and carries a pointer P. The restoring force is provided by a control spring S_1 and a null reading can be obtained with the help of the calibrated spring S_2. (b) Example of a gravity survey over an area about 400 by 500 m, with contours of constant g indicating an ore deposit. (After a survey made by the Boliden Mining Co., Sweden, and reproduced in D. S. Parasnis, Mining Geophysics, *Elsevier, Amsterdam, 1960.)*

This is far too large for convenience, so most surveys, like that of Fig. 8–14(b), show contours labeled in terms of milligals (1 mgal $= 10^{-3}$ cm sec^2 $\approx 10^{-6}\,g$). Under the best conditions, relative measurements accurate to 0.01 mgal may be achievable. One can appreciate how impressive this is by noting that a change of g by 0.01 mgal (1 part in 10^8) corresponds to a change in elevation of only about 3 cm at the earth's surface!

THE MASS OF THE SUN

Let us return to the simple picture of the solar system in which each planet describes a circular orbit about a fixed central sun (Fig. 8–15). We have seen, in the discussion of Kepler's third law, how the use of Newton's law of motion implies that the force on the planet is given, in terms of its mass, orbital radius, and period, by the following equation [Eq. (8–5)]:

$$F_r = -\frac{4\pi^2 mr}{T^2}$$

According to the basic law of the force, however, as expressed by Newton's law of universal gravitation [Eq. (8–8)], the value of F_r is given by the equation

$$F_r = -\frac{GMm}{r^2}$$

where M is the mass of the sun. From the equality of these two expressions, we obtain the following result:

$$T^2 = \frac{4\pi^2 r^3}{GM} \qquad (8\text{–}23)$$

We may again note the feature, already commented on in connection with earth satellites, that the period is independent of the mass of the orbiting object, in this case the earth itself or some other planet. What *does* matter is the mass of the sun, and if we turn Eq. (8–23) around, we have an equation that tells us the value of this mass, M, in terms of observable quantities:

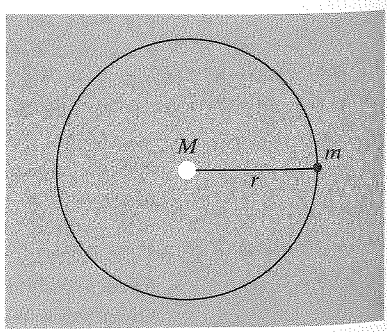


Fig. 8–15 Planetary orbit approximated by a circle with the sun at the center.

$$M = \frac{4\pi^2}{G}\frac{r^3}{T^2} \qquad (8\text{–}24)$$

Kepler's third law expresses the fact that the value of r^3/T^2 is indeed the same for all the planets. The statement of this result does not, however, require the use of absolute values of r—or, for that matter, of T either. It is sufficient to know the values of r and T for the various planets as multiples or fractions of the earth's orbital radius and period. In order to deduce the mass of the sun from Eq. (8–24), however, the use of absolute values is essential. We have seen, earlier in this chapter, how the length of the earth's year has been known with great accuracy since the days of antiquity. A knowledge of the distance from the earth to the sun is, however, rather recent. The development of this knowledge makes an interesting story, which is summarized in the special section following. The final result, expressed as a mean distance in meters, can be substituted as the value of r in Eq. (8–24), along with the other necessary quantities as follows:

$$r_E \approx 1.50 \times 10^{11}\ \text{m}$$
$$T_E \approx 3.17 \times 10^{7}\ \text{sec}$$
$$G = 6.67 \times 10^{-11}\ \text{m}^3/\text{kg-sec}^2$$

We then find that

$$M_{\text{sun}} \approx 2.0 \times 10^{30}\ \text{kg}$$

FINDING THE DISTANCE TO THE SUN

The first attempt to estimate the distance of the sun was made by the great Greek astronomer, Aristarchus, in the third century B.C., and he arrived at a result which, although quite erroneous, held the field for many centuries. His method, sound in principle but made ineffectual by the great remoteness of the sun, is indicated in Fig. 8–16(a). He knew that one half of the moon was illuminated by the sun and that the phases of the moon were the result of viewing this illuminated hemisphere from the earth.

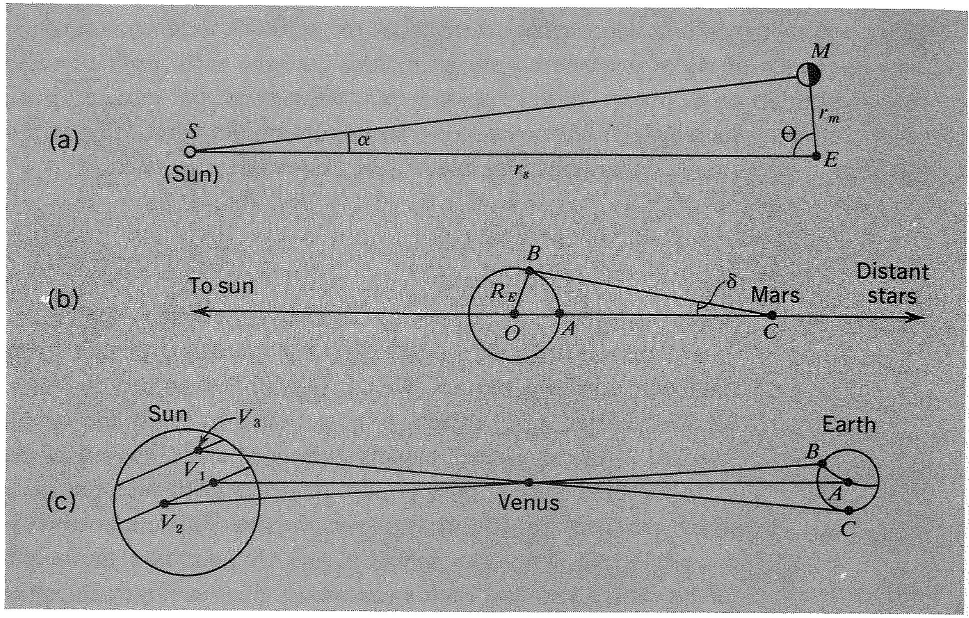


Fig. 8–16 (a) Method attempted by Aristarchus to find the sun's distance by measuring the angle SEM at half-moon. (b) Triangulation method of establishing the scale of the solar system by finding the distance of Mars, using the earth's radius as a base line. (c) Direct determination of the sun's distance by observing the transit of Venus from different points on the earth.

When the moon is exactly half full, the angle SME is 90°. If, in this situation, an exact measurement can be made of the angle θ, the difference in directions of the sun and moon as seen from a point on the earth, we can deduce the angle α $(= 90° - \theta)$ subtended at the sun by the earth–moon distance r_M. Aristarchus judged θ to be about 87°, which gives $\alpha \approx 3° \approx \frac{1}{20}$ rad and hence $r_S \approx 20 r_M$. Since, however, the measured angle is θ and not α, the error in the final result may be (and is) very great. Our present knowledge tells us that the value of θ in the situation represented by Fig. 8–16(a) is actually about 89.8° instead of 87°; this relatively small change in θ raises the ratio r_S/r_M to several hundred instead of 20.

A completely different attack on the problem was initiated by Kepler, although its full exploitation was not possible until much later. Even so it at once became clear that the sun is more distant than Aristachus had concluded. The basis of the method is indicated in Fig. 8–16(b). It involves observations on the planet Mars. When Mars is closest to the earth, it lies on a line joining both planets to the sun. Under these conditions the distance between them is the difference between their orbital radii. Now if Mars is viewed from two different points on the earth, it should appear in slightly different directions with respect to the vastly more distant background of "fixed" stars. The particular angular difference, δ, for observers placed at A and B is called the *parallax;* it is the angle subtended by the earth's radius at the position of Mars. To measure this angle one does not need to have observers at two different points on the earth; the rotation of the earth itself would carry an observer from A to B in about 6 hr during a given night. Now Kepler was able to deduce from the very careful observations of his master, Tycho Brahe, that the value of δ must be less than 3 minutes of arc, which is about 1/1200 rad; he could conclude that the distance to Mars in this configuration must be greater than 1200 earth radii or about 5 million miles. Then, using the known *relative* values of orbital radii from the Copernican scheme (Table 8–1), it follows that the distance of the sun from the earth is more than 2400 earth radii, i.e., more than 10 million miles.

John Flamsteed, a contemporary of Newton to whose observations Newton owed a great deal (he was Astronomer Royal from 1675 to 1720), reduced the upper limit on the parallax of Mars to about 25 seconds of arc, and concluded that the sun's distance was at least 80 million miles. An Italian astronomer, Cassini, arrived at a specific value of about 87 million miles at about the same time, using observations made by himself in Europe and by a French astronomer, Richer, at Cayenne in South America. Another contemporary of Newton's—Edmund Halley[1]—proposed a method that finally led, 100 years later, to the first precision measurements of the sun's distance. The method involved what is known as a *transit of Venus*, i.e., a passage of Venus across the sun's disk, as seen from the earth. Figure 8–16(c) illustrates the basis of the method. As it passes across the sun, Venus looks like a small black dot. Its apparent path, and also the times at which the transit begins or ends, depend on the position of the observer on earth. Since the motion of Venus

[1]Edmund Halley, best known for the comet named after him, succeeded Flamsteed in 1720 as Astronomer Royal. Long before this, however, he had been very active in physics and astronomy. He became a devoted friend and admirer of Newton, and it was largely through his help and persuasion that the *Principia* was published.

is accurately known, the timing of the transit can be used to yield accurate measures of the differences in angular positions of Venus as seen by observers at different positions. From such observations the parallax of Venus can be inferred, after which one can use an analysis just like that for Mars. These transits are fairly rare, because the orbits of the earth and Venus are not in the same plane, but Halley pointed out that a pair of them would occur in 1761 and 1769, and again in 1874 and 1882, and then in 2004 and 2012. From the first two of these (both occurring long after Halley's own death) the solar parallax was found to be definitely between about 8.5 and 9.0 seconds of arc, corresponding to a distance of between about 92 and 97 million miles. Thus the currently accepted result was approached. (The best measurements of this type have been made on the asteroid Eros at its closest approach to the earth.)

Further refinements came with the observations made in the late nineteenth century. One of the most notable of these was the use of an accurately known value of the speed of light to deduce the diameter of the earth's orbit from the accumulated time lag, over a period of 6 months, in the observed eclipses of the moons of Jupiter. The situation is indicated in Fig. 8–17. While the earth moves from E_1 to E_2, Jupiter moves only from J_1 to J_2. This introduces an extra transit time of about 16 min for the light that tells us that one of Jupiter's moons has, for example, just appeared from behind the planet. Knowing that the speed of light is 186,000 miles/sec in empty space, we can deduce that the earth's orbital radius is equal to this speed times about 480 sec, or about 90 million miles. (The calculation was originally done just the other way around, by the Danish astronomer Roemer in 1675. Using an approximate value of the distance from earth to sun, he made the first quantitative estimate of the speed of light.)

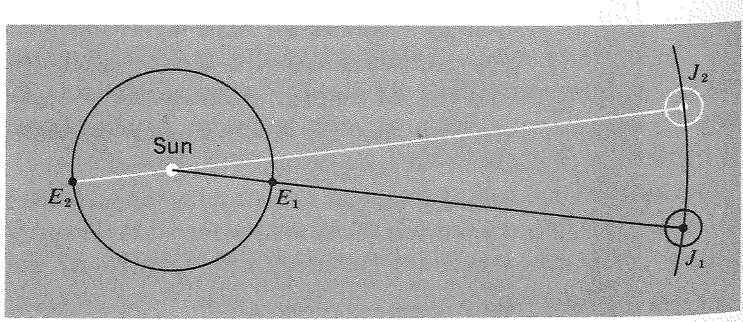


Fig. 8–17 Measurement of the diameter of the earth's orbit by observing the eclipses of Jupiter's moons and the apparent delays due to the travel time of light through space.

Although the modern measurements of the sun's distance are of great accuracy, we must still reckon with the fact that this distance varies during the course of a year. If we ignore this relatively small variation, however, we can make use of the average value, already quoted near the beginning of this chapter:

$$r_E = 1 \text{ AU} = 1.496 \times 10^{11} \text{ m}$$

MASS AND WEIGHT

Perhaps the most profound contribution that Newton made to science was the fundamental connection that he recognized between the inertial mass of an object and the earth's gravitational force on it—a force roughly equal to the measured weight of the object. (Remember, we have defined *weight* as the magnitude of the force, as measured for example on a spring balance, that holds the object at rest relative to the earth's surface.)

It had been known since Galileo's time that all objects near the earth fall with about the same acceleration, **g**. Until Newton it was just a kinematic fact. But in terms of Newton's law it took on a much deeper significance. If an object is observed to have this acceleration, there must be a force $\mathbf{F}_g$ acting on it given by $\mathbf{F} = m\mathbf{a}$, i.e.,

$$\mathbf{F}_g = m\mathbf{g} \qquad (8\text{–}25)$$

It then becomes a vitally significant *dynamical* fact that, since the acceleration **g** is the same for all objects, the force $\mathbf{F}_g$ causing it is strictly proportional to the inertial mass. To appreciate how remarkable this result is, imagine starting from scratch to investigate the force of attraction between two objects in a purely static experiment. One measures the force by balancing it with a springlike device—a torsion fiber. One finds a quantity, which might be called (by analogy with electrical interactions) a gravitational charge, q_g. This "charge" is characteristic of any object and has, as far as these experiments are concerned, nothing at all to do with the inertial mass, which is defined solely in terms of acceleration (under the action of forces produced, for example, by stretched springs). One experiments with objects of all sorts of materials, in different states of aggregation, and so on. It then turns out that, in each and every instance, the gravitational charge is strictly proportional to an independently established quantity, the inertial mass. Is this just a remarkable coincidence, or does it point to something very fundamental? For a long time this apparent coincidence was regarded as one of the unexplained mysteries of nature. It took the sagacity of an Einstein to suspect that gravitation may, in a sense, be *equivalent* to acceleration. Einstein's "postulate of equivalence," that the gravitational charge q_g and the inertial mass m are measures of the same quantity, provided the basis of his own theory of gravitation as embodied in the general theory of relativity. We shall come back to this in Chapter 12, when we discuss noninertial frames of reference.

We are quite accustomed to exploiting the proportionality of F_g to m in our use of the equal-arm balance [Fig. 8–18(a)]. What we are doing is balancing the torques of two forces, but what we are actually interested in is the equality of the amounts of material. We make use of the fact that, to very high accuracy, the value of g is the same at the positions of both masses, and we do not need to bother about what its particular value is. Thanks to the proportionality of gravitational force to mass, we could, with an equal-arm balance and a set of standard weights, measure out a required quantity of any substance equally well on the earth, the moon, or Mars. The spring balance [Fig. 8–18(b)], on the other hand, has a calibration that depends directly on a particular value of g. Its readings are in effect readings of force, even though we use them as a basis for measuring out required amounts of mass. We might find it convenient to use a spring balance for this purpose on the moon, but if its scale were marked in kilograms, we should have to mask this out and attach a fresh calibration with the help of standard masses.

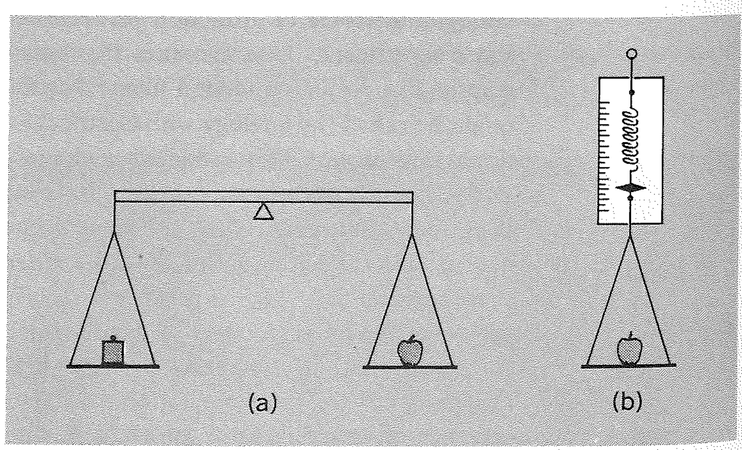


Fig. 8–18 (a) Weighing with an equal-arm balance—in effect a direct comparison of masses, valid whatever the value of g.
(b) Weighing with a spring balance—a measurement of the gravitational force, directly dependent on the value of g.

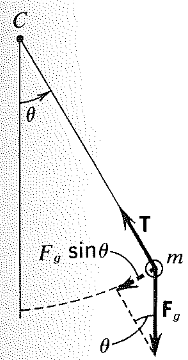


Fig. 8–19 Forces acting on the bob of a simple pendulum.

Newton himself recognized that the strict proportionality of the gravitational force to the inertial mass, as evidenced by the identical local acceleration of falling objects, was a key feature in his own statement of universal gravitation as expressed in Eq. (8–8). He therefore made a series of very careful pendulum experiments to test whether a pendulum of a given length, but with a variety of objects used as the bob, always had the same period. To see how this works, consider an object of inertial mass m hung on a string (Fig. 8–19). The two forces on it (ignoring air resistance) are the tension **T** and the gravitational force $\mathbf{F}_g$. The tension **T** is at every instant perpendicular to the path of the pendulum bob and so has no effect on the tangential acceleration a_θ. The tangential acceleration is due to the tangential component of $\mathbf{F}_g$. From Fig. 8–19,

$$F_\theta = ma_\theta = -F_g \sin\theta$$

from which

$$a_\theta = -(F_g/m)\sin\theta \tag{8–26}$$

At every angle θ the *acceleration* a_θ depends on the ratio F_g/m. Therefore—for given initial conditions—the *velocity* of the bob at every angle θ will be determined by this ratio. So also the *period* for one complete round trip will depend upon the ratio (F_g/m). Newton observed the periods of pendulums with different bobs but with equal lengths. From his observation that the periods of all such pendulums were equal within his experimental error, Newton concluded that F_g was proportional to m to better than 1 part in 1000.

More recent experiments (beginning with Baron Eötvös in Budapest in the nineteenth century) have made use of a very clever idea that permits a *static* measurement. It depends on recognizing that an object hanging *at rest relative to the earth* in fact has an acceleration toward the earth's axis because, by virtue of the earth's rotation, it is traveling in a circular path of radius $r = R\cos\lambda$, where λ is the latitude (see Fig. 8–20). This means that a net force of magnitude $m\omega^2 r$ must be acting on it, where m is the inertial mass. How is this force provided? The answer is that, when a body hangs on a string near the earth's surface, the string, exerting a force **T**, is not in quite the same

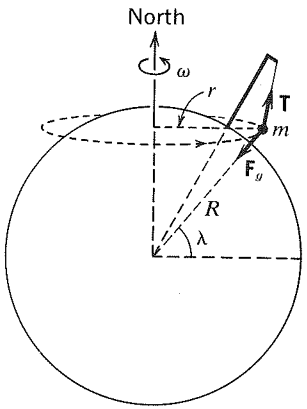


Fig. 8–20 Basis of the Eötvös method for comparing the inertial mass and the gravitational mass of an object that is at rest relative to the earth and hence is being accelerated toward the earth's axis.

direction as the gravitational force $\mathbf{F}_g$. And if $\mathbf{F}_g$ is not strictly proportional to m, the angle between $\mathbf{T}$ and $\mathbf{F}_g$ will be different for different objects. To search for any such variations, a very sensitive torsion balance is used, carrying dissimilar objects at the two ends of a horizontal bar [Fig. 8–21(a)]. If the directions of the tensions $\mathbf{T}_1$ and $\mathbf{T}_2$ are different [Fig. 8–21(b)], there will be small horizontal components [Fig. 8–21(c)] that act in opposite directions with respect to the horizontal bar but that give torques in the same sense. On the other hand, if the directions of $\mathbf{T}_1$ and $\mathbf{T}_2$ are identical, even if their magnitudes are not quite the same, there is no net torque tending to twist the torsion fiber. To test for the existence of any such torque, Eötvös placed the whole apparatus in a case that could be rotated. The horizontal beam carrying the two masses was aligned in an east–west direction [Fig. 8–22(a)] and a reading was taken of its position with respect to the case. The whole system was then rotated through 180°, as in Fig. 8–22(b). If you analyze both situations on the basis of Fig. 8–21, you will find that, with respect to the center line of the case, the angle of twist would be reversed by this operation; hence, if any net torque existed, its existence would be revealed.

Fig. 8–21 Principle of the Eötvös torsion-balance measurement: (a) Two approximately equal masses hang from a torsion bar. (b) If the objects do not have identical ratios of inertial to gravitational mass, the tensions in the suspending strings must be in slightly different directions. In equilibrium, the direction of the main supporting fiber must be intermediate between the directions of $\mathbf{T}_1$ *and* $\mathbf{T}_2$. *(c) This implies the possibility of a net torque that twists the torsion bar about a vertical axis.*

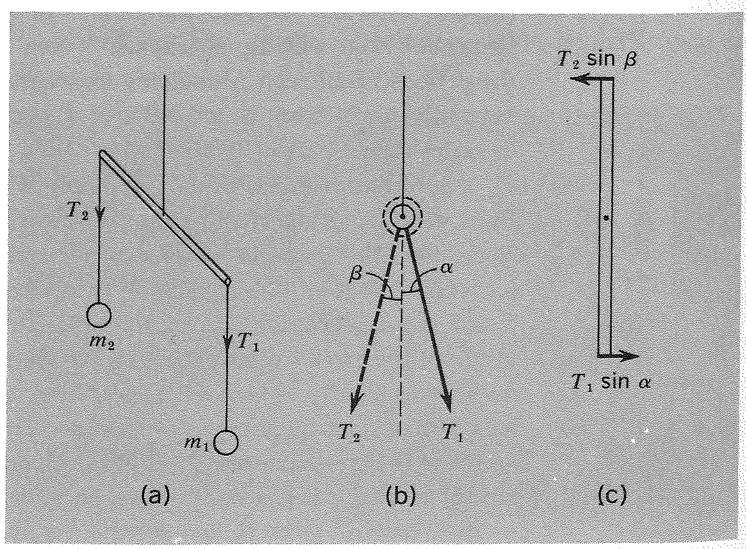
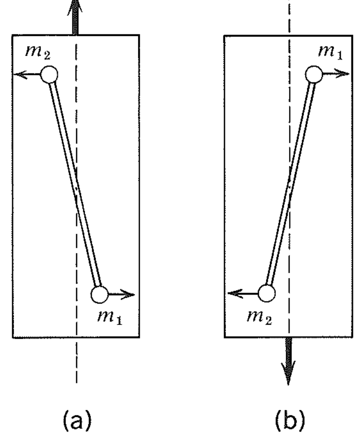


Fig. 8–22 To see whether a net torque exists in the Eötvös experiment, the whole apparatus is turned through 180°. This would reverse the sense of the torque.



More recently, some elegant modernized experiments of this type have been performed by R. H. Dicke and his collaborators.[1] By such experiments it has been shown that the strict proportionality of $\mathbf{F}_g$ to m holds to 1 part in 10^{10} or better.

The description of the above experiments points to a closely related phenomenon—a systematic variation with latitude of the measured weight of an object. If we take the idealization of a perfectly spherical earth (Fig. 8–23), the equilibrium of the object is maintained by applying a force of magnitude W at an angle α to the radius such that the following conditions are satisfied:

$$W \sin \alpha = m\omega^2 r \sin \lambda$$

$$F_g - W \cos \alpha = m\omega^2 r \cos \lambda$$

where $r = R \cos \lambda$. Since α is certainly a very small angle, it is justifiable to put $\cos \alpha \approx 1$ in the second equation, thus giving the result

$$W \approx F_g - m\omega^2 R \cos^2 \lambda$$

[1]See R. H. Dicke, *Sci. Am.*, **205**, 84 (Dec. 1961).

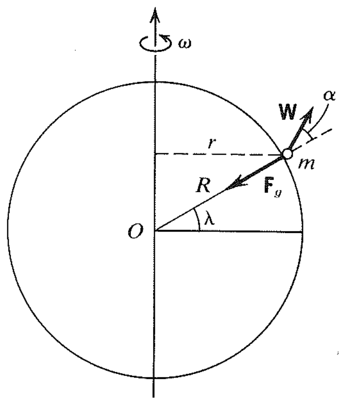


Fig. 8–23 The force needed to balance the weight of an object is different in both direction and magnitude from the force of gravity.

It follows that

$$W(\lambda) \approx W_0 + m\omega^2 R \sin^2 \lambda$$

where W_0 is the measured weight on the equator. Putting $W_0 = mg_0$, we can also obtain a corresponding expression for the latitude dependence of g:

$$g(\lambda) = g_0 + \omega^2 R \sin^2 \lambda$$

If in this expression we substitute $\omega = 2\pi/86{,}400 \text{ sec}^{-1}$ and $R = 6.4 \times 10^6$ m, we find $\omega^2 R \approx 3.4 \times 10^{-2}$ m/sec^2, which with $g_0 \approx 9.8$ m/sec^2 gives us

$$g(\lambda) \approx 9.8(1 + 0.0035 \sin^2 \lambda) \qquad \text{m/sec}^2$$

This formula is more successful than it deserves to be, for we have no right to ignore the significant flattening of the earth, due again to the rotation, which makes the equatorial radius of the earth about 1 part in 300 greater than the polar radius. This ellipticity has two consequences: It puts a point on the equator farther away from the earth's center than it otherwise would be, but it also in effect adds an extra belt of gravitating material around the equatorial region. The resultant value of g at sea level, taking these effects into account, is quite well described by the following formula:

$$g(\lambda) = 9.7805(1 + 0.00529 \sin^2 \lambda) \qquad (8\text{–}27)$$

Thus our simple calculation has the correct form, but its value for the numerical coefficient of the latitude-dependent correction is only about two thirds of the true figure.

WEIGHTLESSNESS

It is appropriate, after the detailed discussion of the relations among mass, gravitational force, and weight, to devote a few words to the property that is called weightlessness. The very explicit distinction that we have drawn between the gravitational force on an object and its measured weight makes use of what is called an *operational* definition of the latter quantity. The weight, as we have defined it, is the magnitude of the force that will hold an object at rest relative to the earth. Our definition of weightlessness derives very naturally from this: *An object is weightless whenever it is in a state of completely free fall.* In this state each part of the object undergoes the same acceleration, of whatever value corresponds to the strength of gravity at its location. (In saying this we assume that g does not change appreciably over the linear extent of the object.) An object that is prevented from falling, by being restrained or supported, inevitably has internal stresses and deformations in its equilibrium state. This may become very obvious, as when a drop of mercury flattens somewhat when it rests on a horizontal surface. All such stresses and deformations are removed in the weightless state of free fall. The mercury drop, for example, is free to take on a perfectly spherical shape.

The above definition of weightlessness can be applied in any gravitational environment, and this is the way it should be. The bizarre dynamical phenomena of life in a space capsule do not depend on getting into regions far from the earth, where the gravitational forces are much reduced, but simply on the fact that the capsule, and everything in it, is falling freely with the same acceleration, which in consequence goes undetected. For example, if a spacecraft is in orbit around the earth, 200 km above the earth's surface, the gravitational force on the spacecraft, and on everything in it, is still about 95% of what one would measure at sea level, but the phenomena associated with what we call weightlessness are just as pronounced under these conditions as they are in another spacecraft 200,000 miles from earth, where the earth's gravitational attraction is down to $\frac{1}{2500}$ of that at the earth's surface. In both situations an object released inside the spacecraft will remain poised in midair. The same would be true in a spacecraft that was simply falling radially toward the earth's center rather than pursuing a circular or elliptical orbit around the earth. When viewed in these terms the phenomena of weightlessness are not in the least mysterious, although they are still startling because they conflict so strongly with our normal experience.

LEARNING ABOUT OTHER PLANETS

The recognition of the universality of gravitation gives us a powerful tool for obtaining information about planets other than the earth, and indeed about celestial objects in general. In particular, if a planet has satellites of its own, we can find its mass by an analysis exactly similar to the one we used for deducing the mass of the sun from the motions of the planets themselves. This provides the simplest way of finding the mass of any planet that has satellites. Such satellites, if a planet has more than one of them, also provide a further test of Kepler's third law, taking the planet itself as the central gravitating body.

Newton himself applied an analysis of this kind to Jupiter, using data for its four most prominent satellites. These were the satellites (or "moons") that made history when Galileo discovered them with his new astronomical telescope in 1610 (see pp. 287–290). Figure 8–24(a) shows their changing positions as seen through a modern telescope, and Fig. 8–24(b) reproduces a few of the sketches that Galileo himself made, night after night, over a period of many months. Figure 8–24(c) is a graph constructed from Galileo's quantitative records, using the readings that can be unambiguously associated with the outermost of the four satellites. A period of about 16 days can be inferred. Galileo had no hesitation in interpreting his observations in terms of the four satellites following circular orbits that were seen edgewise—giving, as we would describe it, the appearance of simple harmonic motion at right angles to our line of sight. On the basis of further measurements Galileo arrived at rather accurate values of the orbital periods of all four satellites and at moderately good values of the orbital radii expressed as multiples of the radius of Jupiter itself.

Newton, in the *Principia*, used similar data of greater precision, obtained by his contemporary John Flamsteed. Table 8–3 (p. 290) lists these data, and in Fig. 8–25 they are plotted logarithmically (cf. Fig. 8–6) in such a way as to show how they give another demonstration of the correctness of Kepler's third law. The slope is accurately $\frac{3}{2}$.

The use of Jupiter's radius as a unit for measuring the orbital radii was not merely a convenience. As we have already noted in our discussion of the mass of the sun, the absolute scale of the solar system was not known with any great certainty in Newton's day. It is interesting, however, that using the data as presented in Table 8–3, without absolute values of the radii, one can deduce the mean density, ρ_J, of Jupiter. By analogy with our analysis of earth satellites [p. 266, and in particular Eq. (8–15)], we have

$$v = \frac{2\pi r}{T} = \left(\frac{GM_J}{r}\right)^{1/2}$$

whence

$$M_J = \frac{4\pi^2 r^3}{GT^2}$$

Putting

$$M_J = \frac{4\pi}{3}\rho_J R_J^3$$

we get

$$\rho_J = \frac{3\pi}{G}\frac{n^3}{T^2}$$

Substituting $n^3/T^2 \approx 7.5 \times 10^{-9}\ \text{sec}^{-2}$, and $G = 6.67 \times 10^{-11}\ \text{m}^3/\text{kg-sec}^2$, we find

$$\rho_J \approx 1050\ \text{kg/m}^3$$

i.e., about the same density as water.

If a planet does not have satellites of its own, the magnitude of its mass may be inferable from a detailed analysis of the tiny disturbing effects—called perturbations—that planets exert on one another's orbits. This technique has been used for Mercury and Venus. The unraveling of these mutual interactions is a complicated and difficult matter, however, and in at least one case it posed a problem that was not adequately answered for a long time. This was the interaction between the two most massive planets, Jupiter and Saturn, which caused irregularities of a very puzzling kind in the orbits of both. It was even considered possible that the basic law of gravitation would need to be modified slightly away from a precise inverse-square relationship. The solution to the mystery was finally achieved, almost a century after the publication of the *Principia*, by the French mathematician Laplace, building on work by his great fellow

The moons of Jupiter

In 1608 Hans Lippershey, in Holland, patented what may have been the first successful telescope. Galileo learned of this, and soon made telescopes of his own design. His third instrument, of more than 30 diameters' magnification, led him to a dramatic discovery, as recounted by him in his book, The Starry Messenger*:*

> *"On the seventh day of January in this present year 1610, at the first hour of the night, when I was viewing the heavenly bodies with a telescope, Jupiter presented itself to me, and . . . I perceived that beside the planet there were three small starlets, small indeed, but very bright. Though I believed them to be among the host of fixed stars, they aroused my curiosity somewhat by appearing to lie in an exact straight line parallel to the ecliptic . . . I paid no attention to the distances between them and Jupiter, for at the outset I thought them to be fixed stars, as I have said. But returning to the same investigation on January eighth — led by what, I do not know — I found a very different arrangement. . . ."*

A few more nights of observation were enough to convince Galileo what he was seeing: "I had now [by January 11] decided beyond all question that there existed in the heavens three stars wandering about Jupiter as do Venus and Mercury about the sun. . . . Nor were there just three such stars; four wanderers complete their revolutions about Jupiter. . . . Also I measured the distances between them by means of the telescope. . . . Moreover I recorded the times of the observations . . . for the revolutions of these planets are so speedily completed that it is usually possible to take even their hourly variations."

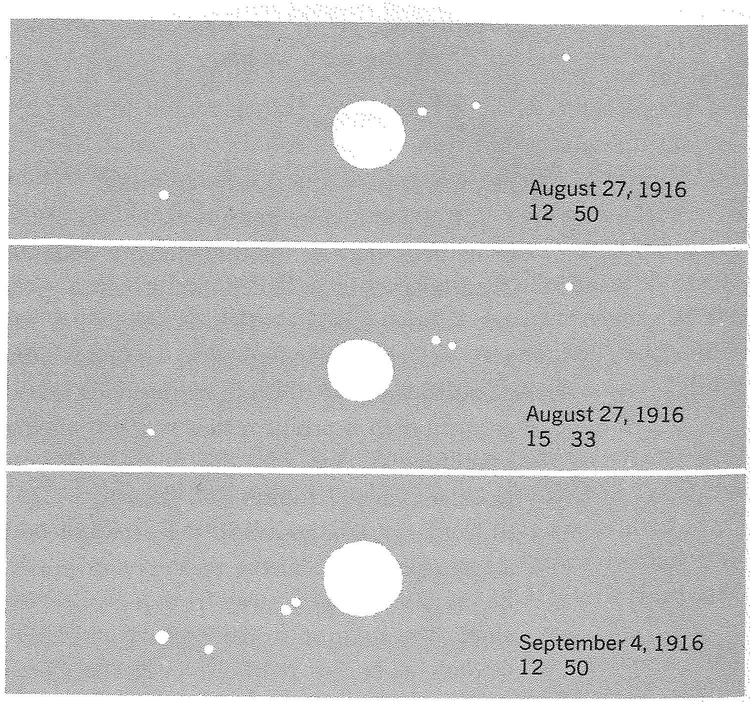


Fig. 8–24(a) Jupiter and its four most prominent satellites as seen through a modern telescope. The first and second photographs illustrate Galileo's observation that noticeable changes occur within a single night. (Yerkes Observatory photographs)

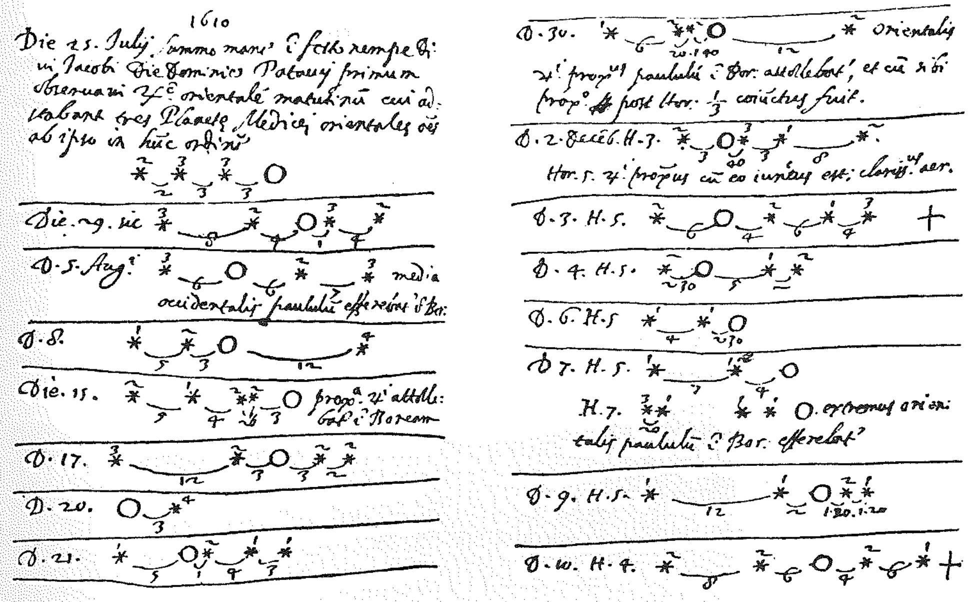


Fig. 8–24(b) Facsimile of a page of Galileo's own handwritten records of his observations during the later months (July–October) of 1610.

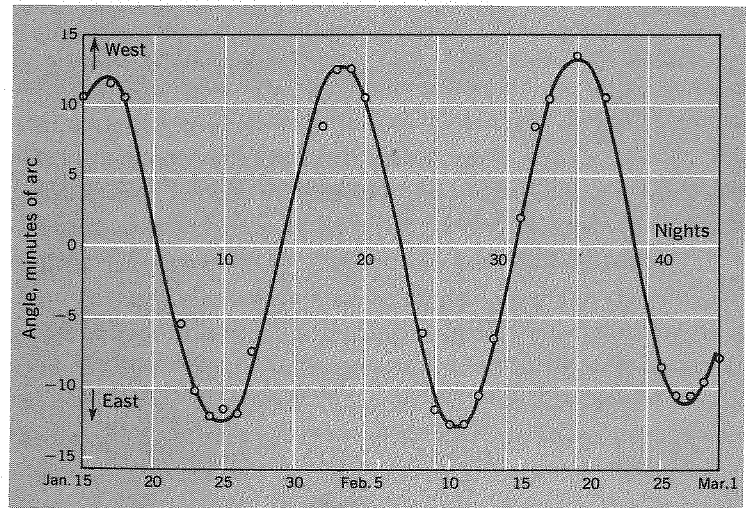


Fig. 8–24(c) A graph constructed from Galileo's own records, showing the periodic motion of Callisto, the outermost of the four satellites visible to him. The period of about 16 3/4 days is clearly exhibited.

countryman Lagrange. It turned out that a curious kind of resonance effect was at work, resulting from the fact that the periods of Jupiter and Saturn are almost in a simple arithmetic relationship ($5T_J \approx 2T_S$). This made large an otherwise negligible term in the perturbation, with a repetition period so long (~900 yr) that it seemed to be increasing without limit. When the mystery was finally resolved the belief in Newton's theory was, of course, strengthened still further.

TABLE 8-3: DATA ON SATELLITES OF JUPITER[a]

Satellite	$n = r/R_J$	*Period* (T)	n^3/T^2, *sec*$^{-2}$
Io	5.578	1.7699 days ≈ 1.53×10^5 sec	7.4×10^{-9}
Europa	8.876	3.5541 days ≈ 3.07×10^5 sec	7.5×10^{-9}
Ganymede	14.159	7.1650 days ≈ 6.19×10^5 sec	7.5×10^{-9}
Callisto	24.903	16.7536 days ≈ 1.45×10^6 sec	7.4×10^{-9}

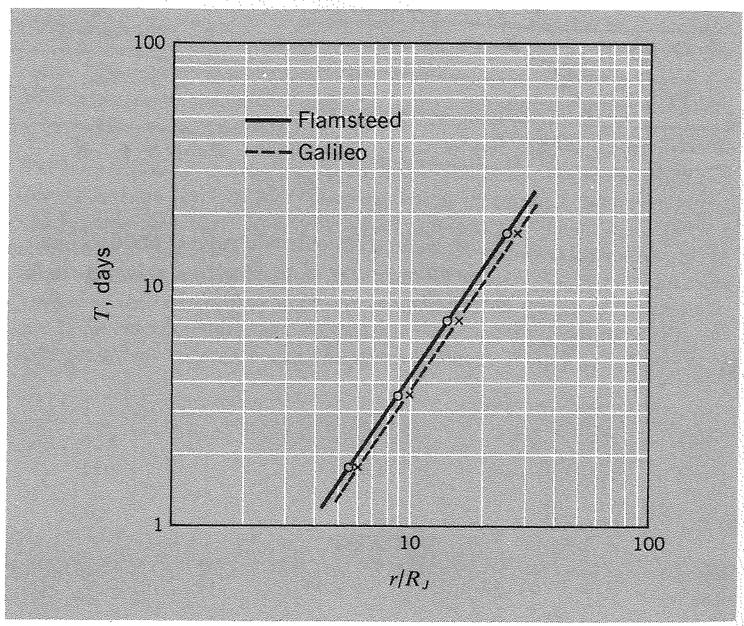


Fig. 8–25 A log-log graph displaying the applicability of Kepler's third law to the Galilean satellites of Jupiter. It may be seen that Galileo's results are little different from those obtained by John Flamsteed nearly 100 years later.

[a]These same data have been presented in a striking way in Eric Rogers, *Physics for the Inquiring Mind*, Princeton University Press, Princeton, N.J., 1960:

Satellite	r^3, (*miles*)^3	T^2, (*hours*)^2
Io	1.803×10^{16}	1.803×10^3
Europa	7.261×10^{16}	7.264×10^3
Ganymede	29.473×10^{16}	29.484×10^3
Callisto	160.440×10^{16}	160.430×10^3

How would you convince your friends that this close *numerical* coincidence is not evidence of a new fundamental law?

THE DISCOVERY OF NEPTUNE

Probably the most vivid illustration of the power of the gravitational theory has been the prediction and discovery of planets whose very existence had not previously been suspected. It is noteworthy that the number of known planets remained unchanged from the days of antiquity until long after Newton. Then, in 1781, William Herschel noticed the object that we now know as Uranus. He was engaged in a systematic survey of the stars, and his only clue to start with was that the object seemed slightly less pointlike than the neighboring stars. Then, having a telescope with various degrees of magnification, he confirmed that the size of the image increased with magnification, which is not true for the stars—they remain below the limit of resolution of even the biggest telescopes, and always produce images indistinguishable from those due to ideal point sources.

Once his attention had been drawn to the object, Herschel returned to it night after night and confirmed that it was moving with respect to the other stars. Also, as has happened in various other cases, it was found that the existence of the object had in fact been recorded in earlier star maps (first by John Flamsteed in 1690). These old data suddenly became extremely valuable, because they were a ready-made record of the object's movements dating back through nearly a century. When combined with new measurements carried out over many months, they showed that the object (finally to be called Uranus) was indeed a member of our solar system, following an almost circular orbit with a mean radius of 19.2 AU and a period of 84 years.

This is where our main story begins. Once it was discovered, Uranus and its motions became the subject of a continuing study, and evidence began to accumulate that there were some extremely small irregularities in its motion that could not be ascribed to perturbing effects from any known source. Figure 8–26(a), a tribute to the wonderful precision of astronomical observation, shows the anomaly as a function of time since 1690. The suspicion began to grow that perhaps there was yet another

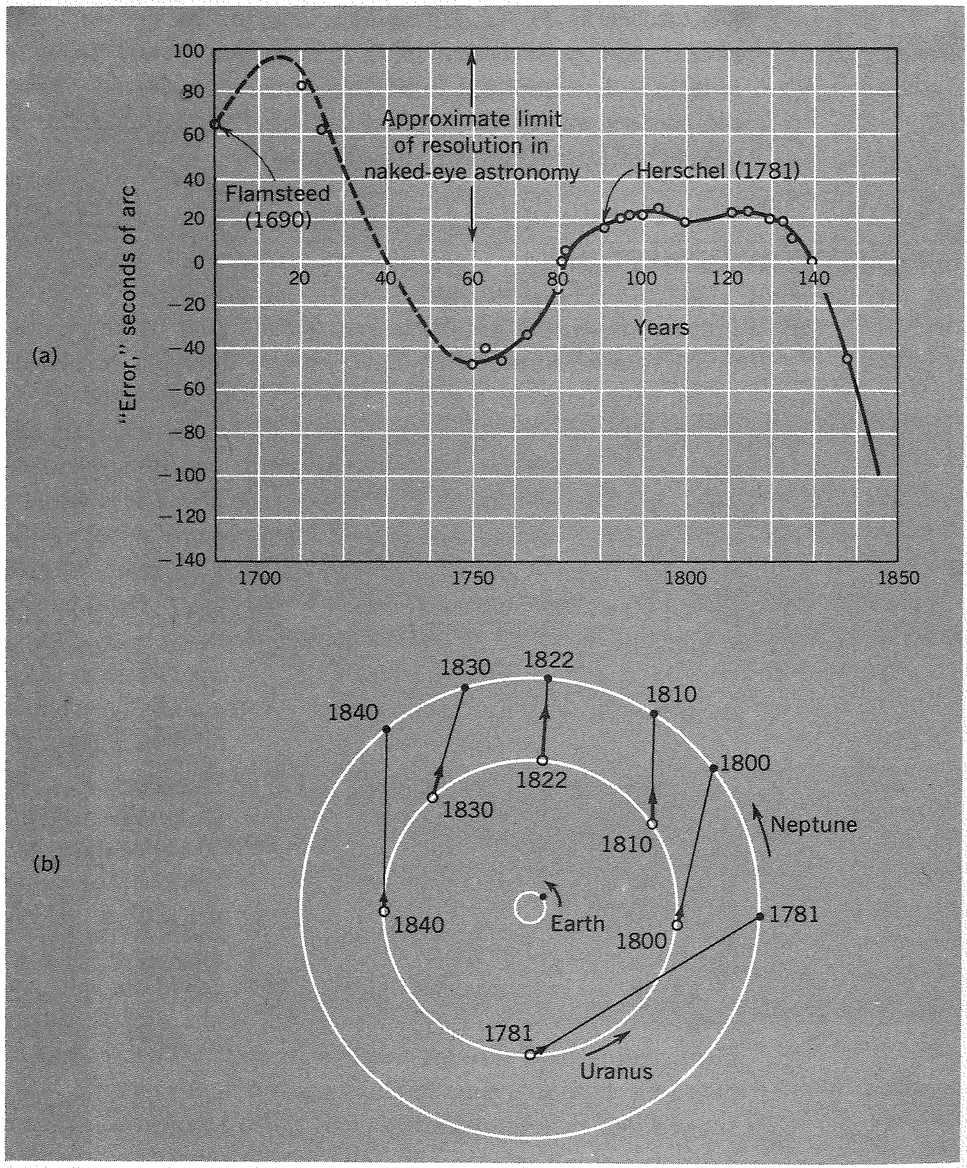


Fig. 8–26 (a) Unexplained residual deviations in the observed positions of Uranus between 1690 and 1840. (b) Basis of ascribing the deviations to the influence of an extra planet. The arrows indicate the relative magnitude of the perturbing force at different times.

planet beyond Uranus, unknown in mass, period, or distance. Two men—J. C. Adams in England and U. J. LeVerrier in France—independently worked on the problem. Both men used as a starting point the assumption that the orbital radius of the unknown planet was almost exactly twice that of Uranus. The basis of this was a curious empirical relation, known as Bode's law (actually discovered by J. D. Titius in 1772, but publicized by J. Bode), which expresses the fact that the orbital radii of the known planets can be roughly fitted by the following formula:

$$R_n \text{ (AU)} = 0.4 + (0.3)(2^n)$$

where n is an appropriate integer for each planet. Putting $n = 0$, 1, 2, we get the approximate radii for Venus, earth, and Mars (Mercury requires $n = -\infty$, which is hard to defend). Using $n = 4$, 5, and 6 one gets quite good values for Jupiter, Saturn, and Uranus. (The missing integer, $n = 3$, corresponds to the asteroid belt.) Figure 8–27 shows this relation of orbital radii with the help of a semilog plot; it is clear that a simple exponential relation (linear on this graph) works almost as well, but if one accepts Bode's law, then $n = 7$ gives $r = 38.8$ AU, and this is what Adams and LeVerrier used.

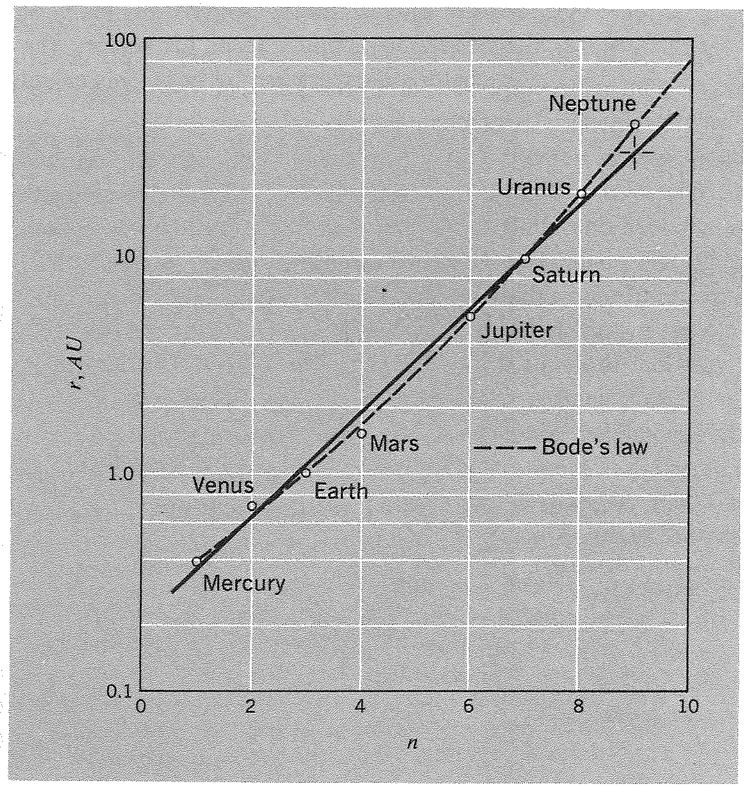


Fig. 8–27 Graph for predicting the orbital radius of the new planet with the help of Bode's law.

Given the radius, the period is automatically defined by Kepler's third law, and then it becomes possible to construct a definite picture, as shown in Fig. 8–26(b), of the way in which the new planet could alternately accelerate and retard Uranus in its orbital motion, depending on their relative positions. With the help of laborious analysis, one can then deduce where in its orbit the new planet should be on a particular date. Adams supplied such information in October 1845 to the British Astronomer Royal, G. B. Airy, who acknowledged Adams' letter, raised a question of detail, but otherwise did nothing. LeVerrier did not complete his own calculations until August 1846, but the astronomer to whom he wrote (J. G. Galle, in Germany) made an immediate search and identified the new planet (Neptune) on his very first night of observation. It was only about a degree from the predicted position (see Fig. 8–28). The next night it had visibly shifted, thereby confirming its planetary status.

Although the discovery of Neptune is in some respects a great success story, it is also a story of luck, both good and bad, and of human frailty. Adams was really first in the field, but he received no support from his seniors (he was fresh from his bachelor's degree when he began his calculations). Airy missed

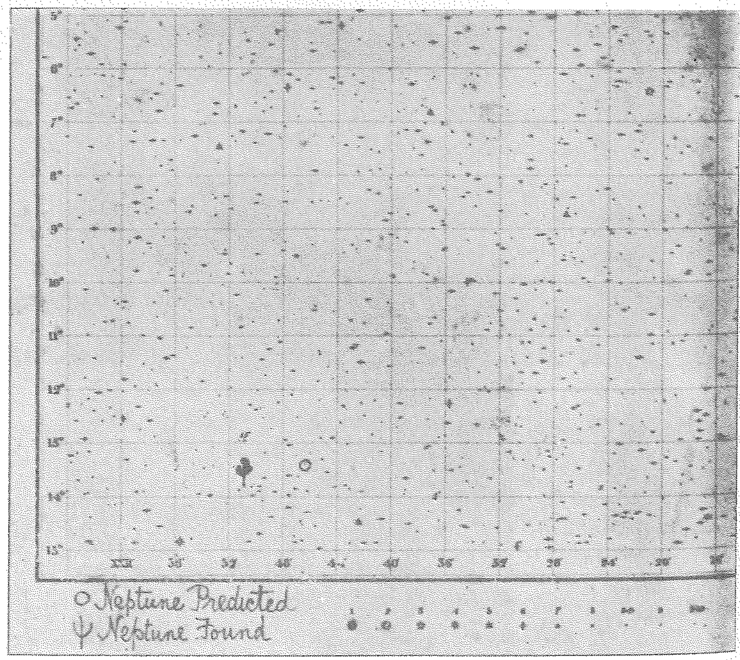


Fig. 8–28 Star map showing the discovery of Neptune, September 23, 1846. (From Herbert Hall Turner, Astronomical Discovery, *Edward Arnold, London, 1904.)*

the credit, which he might readily have won, of being the man who first identified Neptune. But the locations that both Adams and LeVerrier predicted might well have been hopelessly misleading, for in their reliance on Bode's law they used an orbital radius (and hence a period) that was far from correct. The true value is about 30 AU instead of nearly 40 as they had assumed, which means that they overestimated the period by nearly 50%.[1] It was therefore largely a lucky accident that the planet was so near to its predicted position on the particular date that Galle sought and found it. But let this not be taken as disparagement. A great discovery was made, with the help of the laws of motion and the gravitational force law, and it remains as the most triumphant confirmation of the dynamical model of the universe that Newton invented.[2] The discovery of Pluto by C. Tombaugh in 1930, on the basis of a detailed record of the irregularities of Neptune's own motion, provided an echo of the original achievement.

GRAVITATION OUTSIDE THE SOLAR SYSTEM

When Newton wrote his *System of the World*, nothing was known about the distances or possible motions of the stars. They simply provided a seemingly fixed background against which the dynamics of the solar system proceeded. There were exceptions. A few prominent stars—e.g., Sirius—known since antiquity by naked-eye astronomy, were found to have shifted position within historic time. But the serious and systematic investigation of stellar motions was begun by William Herschel. His observations, continued and refined by his son John Herschel and by other astronomers, revealed two classes of results. The first was the continuing apparent displacement of individual stars in a way that suggested that the solar system is itself involved in a general movement of the stars in our neighborhood, at a speed of the order of 10 miles/sec (comparable to the earth's own orbital speed around the sun). This, as it stood, was just an empirical fact. But the second type of result pointed directly

[1]This also means that they overestimated the mass necessary to produce the observed perturbations of Uranus. LeVerrier gave a figure of about 35 times the mass of the earth; the currently accepted value is about half this.

[2]For a detailed account of the whole matter, see H. H. Turner, *Astronomical Discovery*, Edward Arnold, London, 1904. A shorter but more readily accessible account may be found in an essay entitled "John Couch Adams and the Discovery of Neptune," by Sir H. Spencer Jones, in *The World of Mathematics* (J. R. Newman, ed.), Simon and Schuster, New York, 1956.

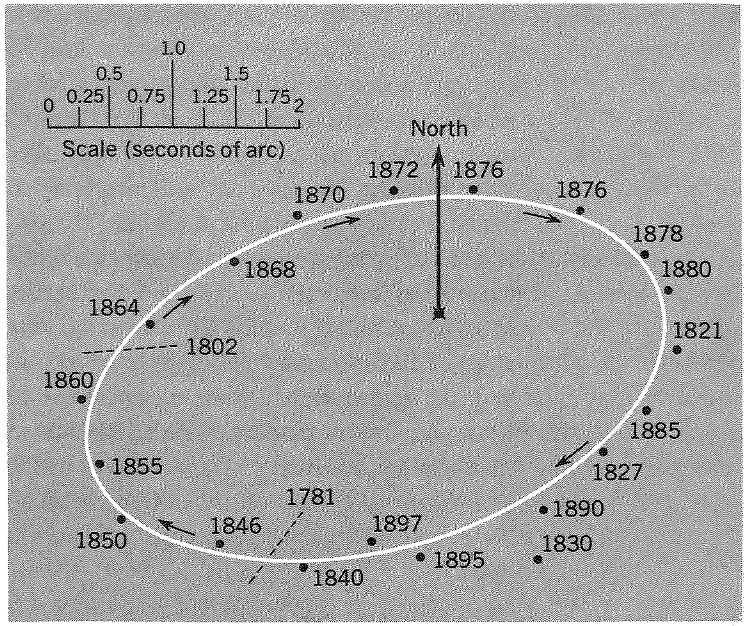


Fig. 8–29 Variation with time of the relative position vector of the members of a double-star system. (After Arthur Berry, A Short History of Astronomy, *1898; reprinted by Dover Publications, New York, 1961.)*

to the operation of Newton's dynamics. For the Herschels discovered numerous pairs of stars that were evidently orbiting around one another as *binary systems.* Figure 8–29 shows one of the best documented early examples, and the first to be subjected to a detailed analysis in terms of Kepler's laws. (It is ξ-Ursae, in one of the hind paws of the constellation known as the Great Bear.)

The period of a binary star depends on the *total* mass of the system, not on the individual masses. This is easily proved in the case in which the orbits are assumed to be circles around the common center of mass [see Fig. 8–30(a)].[1] The individual stars are always at opposite ends of a straight line passing through the center, *C*. If we write the statement of $F = ma$ for one of the stars, say m_1, we have

$$G\frac{m_2 m_1}{(r_1 + r_2)^2} = \frac{m_1 v_1^2}{r_1} = m_1 \omega^2 r_1$$

where ω ($= 2\pi/T$) is the angular velocity common to both stars. Hence

$$\omega^2 = \frac{G m_2}{r_1 (r_1 + r_2)^2}$$

[1]For a full discussion of the concept of center of mass, see Ch. 9, p. 337.

(a)

(b)

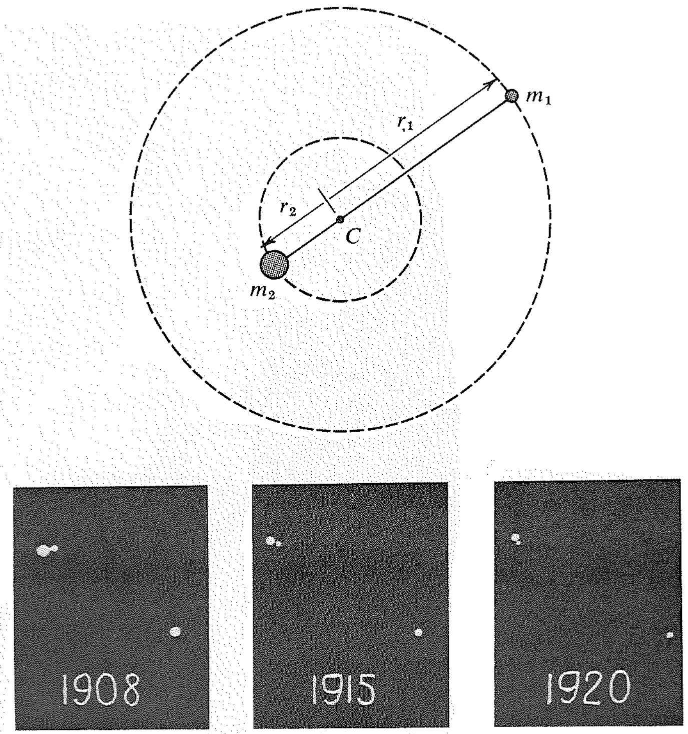


Fig. 8–30 (a) Motion of the members of a binary star system with respect to the center of mass, C, for the case of circular orbits. (b) Direct visual evidence of the motion of a binary system—Krueger 60, photographed by E. E. Barnard. (Yerkes Observatory photograph.)

However, by the definition of the center of mass, we have

$$r_1 = \frac{m_2}{m_1 + m_2} r$$

where

$$r = r_1 + r_2$$

It follows at once that

$$\omega^2 = \frac{G(m_1 + m_2)}{r^3}$$

Thus if the distance r between the stars can be obtained by direct astronomical observation (e.g., starting from a knowledge of their angular separation), the sum of the masses is at once determined. Finding the individual masses entails the somewhat harder job of measuring the motion of each star in absolute terms against the background of the "fixed" stars. Figure 8–30(b) shows convincing direct evidence of the orbital motion of an actual binary system.

Fig. 8–31 Rotating galaxy (spiral galaxy NGC 5194 in the constellation Canes Venatici). (Photograph from the Hale Observatories.)

With the development of modern astronomy, the systematic motions of our sun and its neighbors came to be seen as part of a greater scheme of movements controlled by gravity. All around us throughout the universe were the immense stellar systems—the galaxies—most of them vividly suggesting a state of general rotation, as in Fig. 8–31, for example. The most difficult structure to elucidate was the one in which we ourselves are embedded, i.e., the Milky Way galaxy. It finally became clear, however, that its basic structure is very much like that of Fig. 8–31, and that in it our sun is describing some kind of orbit around the center, with a radius of about 3×10^{20} m ($\approx$ 30,000 light-years) and an estimated period of about 250 million years ($\approx 8 \times 10^{15}$ sec). Using these figures we can infer the approximate gravitating mass, inside the orbit, that would define this motion. From Eq. (8–24) we have

$$M = \frac{4\pi^2}{G}\frac{r^3}{T^2}$$

With $G \approx 7 \times 10^{-11}$ m^3/kg-sec^2, we find that

$$M \approx \frac{40}{7 \times 10^{-11}} \times \frac{3 \times 10^{61}}{6 \times 10^{31}} \approx 3 \times 10^{41}\text{ kg}$$

Since the mass of the sun (a typical star) is about 2×10^{30} kg, we see that a core of about 10^{11} stars is implied. This is not really a figure that can be independently checked. It is a kind of ultimate tribute to our belief in the universality of the gravitational law that it is confidently used to draw conclusions like those above concerning masses of galactic systems.

EINSTEIN'S THEORY OF GRAVITATION

We have described earlier how Newton recognized that the proportionality of weight to inertial mass is a fact of fundamental significance; it played a central role in leading him to the conclusion that his law of gravitation must be a general law of nature. For Newton this was a strictly dynamical result, expressing the basic properties of the force law. But Albert Einstein, in 1915, looked at the situation through new eyes. For him the fact that all objects fall toward the earth with the same acceleration g, whatever their size or physical state or composition, implied that this must be in some truly profound way a kinematic or geometrical result, not a dynamical one. He regarded it as being on a par with Galileo's law of inertia, which expressed the tendency of objects to persist in straight-line motion.

Building on these ideas, Einstein developed the theory that a planet (for example) follows its characteristic path around the sun because in so doing it is traveling along what is called a *geodesic* line—that is to say, the most economical way of getting from one point to another. His proposition was that although in the absence of massive objects the geodesic path is a straight line in the Euclidean sense, the presence of an extremely massive object such as the sun modifies the geometry locally so that the geodesics become curved lines. The state of affairs in the vicinity of a massive object is, in this view, to be interpreted not in terms of a gravitational field of force but in terms of a "curvature of space"—a facile phrase that covers an abstract and mathematically complex description of non-Euclidean geometries.

For the most part the Einstein theory of gravitation gives results indistinguishable from Newton's; the grounds for preferring it might seem to be conceptual rather than practical. But in one celebrated instance of planetary motions there is a real discrepancy that favors Einstein's theory. This is in the so-called "precession of the perihelion" of Mercury. The phenomenon is that the orbit of Mercury, which is distinctly elliptical in shape, very gradually rotates or precesses in its own plane, so that the major axis is along a slightly different direction at the end of each complete revolution. Most of this precession (amounting to about 10 minutes of arc per century) can be understood in terms of the disturbing effects of the other planets according to Newton's law of gravitation.[1] But there remains a tiny, obstinate residual rotation equal to 43 seconds of arc per century. The attempts to explain it on Newtonian theory—for example by postulating an unobserved planet inside Mercury's own orbit—all came to grief by conflicting with other facts of observation concerning the solar system. Einstein's theory, on the other hand, without the use of any adjustable parameters, led to a calculated precession rate that agreed exactly with observation. It corresponded, in effect, to the existence of a very small force with a different dependence on distance than the dominant $1/r^2$ force of Newton's theory. The way in which a disturbing effect of this kind causes the orbit to precess is discussed in Chapter 13. Other empirical modifications of the basic law of gravitation—small departures from the inverse-square law—had been tried before Einstein developed his theory, but apart from their arbitrary character they also led to false predictions for the other planets. In Einstein's theory, however, it emerged automatically that the size of the disturbing term was proportional to the square of the angular velocity of the planet and hence was much more important for Mercury, with its short period, than for any of the other planets.

[1]The apparent amount of precession as viewed from the earth is actually about 1.5° per century, but most of this is due to the continuous change in the direction of the earth's own axis (the precession of the equinoxes — see Chapter 14).

PROBLEMS

8–1 Given a knowledge of Kepler's third law as it applies to the solar system, together with the knowledge that the disk of the sun subtends an angle of about $\frac{1}{2}°$ at the earth, deduce the period of a hypothetical planet in a circular orbit that skims the surface of the sun.

8–2 It is well known that the gap between the four inner planets and the five outer planets is occupied by the asteroid belt instead of by a tenth planet. This asteroid belt extends over a range of orbital radii from about 2.5 to 3.0 AU. Calculate the corresponding range of periods, expressed as multiples of the earth's year.

8–3 It is proposed to put up an earth satellite in a circular orbit with a period of 2 hr.

(a) How high above the earth's surface would it have to be?

(b) If its orbit were in the plane of the earth's equator and in the same direction as the earth's rotation, for how long would it be continuously visible from a given place on the equator at sea level?

8–4 A satellite is to be placed in synchronous circular orbit around the planet Jupiter to study the famous "red spot" in Jupiter's lower atmosphere. How high above the surface of Jupiter will the satellite be? The rotation period of Jupiter is 9.9 hr, its mass M_J is about 320 times the earth's mass, and its radius R_J is about 11 times that of the earth. You may find it convenient to calculate first the gravitational acceleration g_J at Jupiter's surface as a multiple of g, using the above values of M_J and R_J, and then use a relationship analogous to that developed in the text for earth satellites [Eq. (8–16) or (8–17)].

8–5 A satellite is to be placed in a circular orbit 10 km above the surface of the moon. What must be its orbital speed and what is the period of revolution?

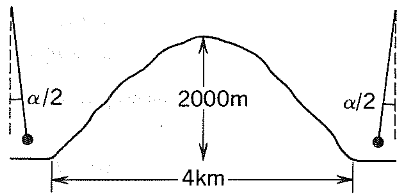
8–6 A satellite is to be placed in synchronous circular earth orbit. The satellite's power supply is expected to last 10 years. If the maximum acceptable eastward or westward drift in the longitude of the satellite during its lifetime is 10°, what is the margin of error in the radius of its orbit?

8–7 The springs found in retractable ballpoint pens have a relaxed length of about 3 cm and a spring constant of perhaps 0.05 N/mm. Imagine that two lead spheres, each of 10,000 kg, are placed on a frictionless surface so that one of these springs just fits, in its uncompressed state, between their nearest points.

(a) How much would the spring be compressed by the mutual gravitational attraction of the two spheres? The density of lead is about 11,000 kg/m³.

(b) Let the system be rotated in the horizontal plane. At what frequency of rotation would the presence of the spring become irrelevant to the separation of the masses?

8–8 During the eighteenth century, an ingenious attempt to find the mass of the earth was made by the British Astronomer Royal, Nevil Maskelyne. He observed the extent to which a plumbline was pulled out of true by the gravitational attraction of a mountain. The figure illustrates the principle of the method. The change of direction of the plumbline was measured between the two sides of the mountain.



(This was done by sighting on stars.) After allowance had been made for the change in direction of the local vertical because of the curvature of the earth, the residual angular difference α was given by $2F_M/F_E$, where $\pm F_M$ is the horizontal force on the plumb-bob due to the mountain, and $F_E = GM_Em/R_E^2$. (m is the mass of the plumb-bob.)

The value of α is about 10 seconds of arc for measurements on opposite sides of the base of a mountain about 2000 m high. Suppose that the mountain can be approximated by a cone of rock (of density 2.5 times that of water) whose radius at the base is equal to the height and whose mass can be considered to be concentrated at the center of the base. Deduce an approximate value of the earth's mass from these figures. (The true answer is about 6×10^{24} kg.) Compare the gravitational deflection α to the change of direction associated with the earth's curvature in this experiment.

8–9 Imagine that in a certain region of the ocean floor there is a roughly cone-shaped mound of granite 250 m high and 1 km in diameter. The surrounding floor is relatively flat for tens of kilometers in all directions. The ocean depth in the region is 5 km and the density of the granite is 3000 kg/m^3. Could the mound's presence be detected by a surface vessel equipped with a gravity meter that can detect a change in g of 0.1 mgal?

(*Hint:* Assume that the field produced by the mound at the surface can be approximated by the field of a mass point of the same total mass located at the level of the surrounding floor. Note that in calculating the change in g you must keep in mind that the mound has displaced its own volume of water. The density of water, even at such depths, can be taken as about equal to its surface value of about 1000 kg/m^3.

8–10 Show that the period of a particle that moves in a circular orbit close to the surface of a sphere depends only on G and the mean density of the sphere. Deduce what this period would be for any sphere having a mean density equal to the density of water. (Jupiter almost corresponds to this case.)

8–11 Calculate the mean density of the sun, given a knowledge of G, the length of the earth's year, and the fact that the sun's diameter subtends an angle of about 0.55° at the earth.

8–12 An astronaut who can lift 50 kg on earth is exploring a planetoid (roughly spherical) of 10 km diameter and density 3500 kg/m^3.

(a) How large a rock can he pick up from the planetoid's surface, assuming that he finds a well-placed handle?

(b) The astronaut observes a rock falling from a cliff. The rock's radius is only 1 m and as it approaches the surface its velocity is 1 m/sec. Should he try to catch it? (This is obviously a fanciful problem. One would not expect a planetoid to have cliffs or loose rocks, even if an astronaut were to get there in the first place.)

8–13 It is pointed out in the text that a person can properly be termed "weightless" when he is in a satellite circling the earth. The moon is a satellite, yet it is noted in many discussions that we would weigh $\frac{1}{6}$ of our normal weight there. Is there a contradiction here?

8–14 A dedicated scientist performs the following experiment. After installing a huge spring at the bottom of a 20-story-high elevator shaft, he takes the elevator to the top, positions himself on a bathroom scale inside the airtight car with a stopwatch and with pad and pencil to record the scale reading, and directs an assistant to cut the car's support cable at $t = 0$. Presuming that the scientist survives the first encounter with the spring, sketch a graph of his measured weight versus time from $t = 0$ up to the beginning of the second bounce. (*Note:* Twenty stories is ample distance for the elevator to acquire terminal velocity.)

8–15 A planet of mass M and a single satellite of mass $M/10$ revolve in circular orbits about their stationary center of mass, being held together by their gravitational attraction. The distance between their centers is D.

(a) What is the period of this orbital motion?

(b) What fraction of the total kinetic energy resides in the satellites?

(Ignore any spin of planet and satellite about their own axes.)

8–16 We have considered the problem of the moon's orbit around the earth as if the earth's center represented a fixed point about which the motion takes place. In fact, however, the earth and the moon revolve about their common center of mass.

(a) Calculate the position of the center of mass, given that the earth's mass is 81 times that of the moon and that the distance between their centers is 60 earth radii.

(b) How much longer would the month be if the moon were of negligible mass compared to the earth?

8–17 The sun appears to be moving at a speed of about 250 km/sec in a circular orbit of radius about 25,000 light-years around the center of our galaxy. (One light-year $\simeq 10^{16}$ m.) The earth takes 1 year to describe an almost circular orbit of radius about 1.5×10^{11} m around the sun. What do these facts imply about the total mass responsible for keeping the sun in its orbit? Obtain this mass as a multiple of the sun's mass M. (Note that you do not need to introduce the numerical value of G to obtain the answer.)

8–18 (A good problem for discussion.) In 1747 Georges Louis Lesage explained the inverse-square law of gravitation by postulating that vast numbers of invisible particles were flying through space in all directions at high speeds. Objects like the sun and planets block these particles, leading to a shadowing effect that has the same quantitative result as a gravitational attraction. Consider the arguments for and against this theory.

(*Suggestion:* First consider a theory in which opaque objects block the particles completely. This proposal is fairly easy to refute. Next consider a theory in which the attenuation of the particles by objects is incomplete or even very small. This theory is much harder to dismiss.)

8–19 The continuous output of energy by the sun corresponds (through Einstein's relation $E = Mc^2$) to a steady decrease in its mass M, at the rate of about 4×10^6 tons/sec. This implies a progressive increase in the orbital periods of the planets, because for an orbit of a given radius we have $T \sim M^{-1/2}$ [cf Eq. (8–23)].

A precise analysis of the effect must take into account the fact that as M decreases the orbital radius itself increases—the planets gradually spiral away from the sun. However, one can get an order-of-magnitude estimate of the size of the effect, albeit a little bit on the low side, by assuming that r remains constant. (See Problem 13–21 for a more rigorous treatment.)

Using the simplifying assumption of constant r, estimate the approximate increase in the length of the year resulting from the sun's decrease in mass over the time span of accurate astronomical observations—about 2500 years.

8–20 It is mentioned at the end of the chapter how Einstein's theory of gravitation leads to a small correction term on top of the basic Newtonian force of gravitation. For a planet of mass m, traveling at speed v in a circular orbit of radius r, the gravitational force becomes in effect the following:

$$F = \frac{GMm}{r^2}\left(1 + 6\frac{v^2}{c^2}\right)$$

where c is the speed of light. (Correction terms of the order of v^2/c^2 are typical of relativistic effects.)

(a) Show that, if the period under a pure Newtonian force GMm/r^2 is denoted by T_0, the modified period T is given approximately by

$$T \simeq T_0\left(1 - \frac{12\pi^2 r^2}{c^2 T_0^2}\right)$$

(Treat the relativistic correction as representing, in effect, a small fractional increase in the value of G, and use the value of v corresponding to the Newtonian orbit.)

(b) Hence show that, in each revolution, a planet in a circular orbit would travel through an angle greater by about $24\pi^3 r^2/c^2 T_0^2$ than under the pure Newtonian force, and that this is also expressible as $6\pi GM/c^2 r$, where M is the mass of the sun.

(c) Apply these results to the planet Mercury and verify that the accumulated advance in angle amounts to about 43 seconds of arc per century. This corresponds to what is called the precession of its orbit.