1 Newton's gravitational theory

> It was occasioned by the fall of an apple,
> As he sat in a contemplative mood …
>
> William Stukeley, *Memoirs of Sir Isaac Newton's Life*

Few theories can compare in the accuracy of their predictions with Newton's theory of universal gravitation. The predictions of celestial mechanics for the positions of the major planets agree with observation to within a few arcseconds over time intervals of many years. The discovery of Neptune and the rediscovery of Ceres are among the spectacular successes that testify to the accuracy of the theory. But Newton's theory is not perfect: The predicted motions of the perihelia for the inner planets deviate somewhat from the observed values. In the case of Mercury the excess perihelion precession amounts to 43 arcseconds per century. This small deviation was discovered through calculations by LeVerrier in 1845, and it was confirmed by Newcomb in 1882. The explanation of this perihelion precession became one of the early successes of Einstein's relativistic theory of gravitation.

Telescopic observations of planetary angular positions stretching over hundreds of years are needed to detect the excess perihelion precession. However, with the development of radar astronomy it has become possible to measure the distances to the inner planets directly and very accurately by means of the travel time of a radio signal sent from the Earth to the planet and reflected back. With such radar observations of distances, the small deviations from Newton's theory can be detected after just a few years of observation.

Although Newton's theory is not perfect, it is in excellent agreement with observation in the limiting case of motion at low velocities in a weak gravitational field. Any relativistic theory of gravitation ought to agree with Newton's theory in this limiting case. We therefore begin with a brief exposition of some aspects of Newton's theory.

1.1 The law of universal gravitation

According to Newton, the law governing gravitational interactions is "that there is a power of gravity pertaining to all bodies, proportional to the several quantities of matter which they contain . . . The force of gravity towards the several equal parts of any body is inversely as the square of the distance of places from the particles" (Newton, 1686).

If one particle (m') is at the origin and the other (m) is at a distance r, then the force is in radial direction, and it has a magnitude

$$F = \frac{Gmm'}{r^2} \tag{1.1}$$

The value of the gravitational constant in Eq. (1.1) is $G = 6.6743 \times 10^{-8}$ dyne · cm^2/g^2.

Strictly speaking, the masses that enter the force law (1.1) are the gravitational masses, which are the sources and the "receptors" of gravitation, in the same way that the electric charge is the source and the receptor of electromagnetic forces. In principle, the gravitational mass is distinct from the inertial mass, which enters on the left side of the equation of motion, $ma = F$. Experimentally, these two kinds of masses are found to be equal, and we will examine the experimental evidence for this equality in Section 1.6. In the following discussion of gravitational fields and potentials (Sections 1.1–1.4), the masses are always gravitational.

If we adopt a naive interpretation of the force law (1.1), gravitation is action-at-distance: A mass at one point acts directly and instantaneously on another mass even though the other mass is not in contact with it. Newton had serious misgivings about such a ghostly tug-of-war of distant masses and suggested that the interaction should be conveyed by some material medium. The modern view is that gravitation, like electromagnetism and all other fundamental interactions, acts locally through fields: A mass at one point produces a field, and this field acts on whatever masses with which it comes into contact. The gravitational field may be regarded as the material medium sought by Newton; the field is material because it possesses an energy density. The description of interactions by means of local fields has the further advantage of leading to a relativistic theory in which gravitational effects propagate at finite velocity. Instantaneous action-at-distance makes no sense as a relativistic theory because of the lack of an absolute time; what is instantaneous propagation in one reference frame need not be instantaneous in another. Of course, in the case of static or quasi-static mass distributions, retardation effects are insignificant, and there is then no practical distinction between local interaction and action-at-distance.

In our Solar System, Newton's theory is an excellent approximation. The condition for the validity of Newton's theory can be conveniently stated in terms of the potential energy $V(r)$, which for the inverse-square force (1.1) is

$$V(r) = -\frac{Gmm'}{r} \tag{1.2}$$

In general, we can say that relativistic effects will be small, provided that the potential energy of the moving particle is much less than the rest-mass energy and that the speed is much less than the speed of light. For a mass m orbiting with speed v around a central mass m', we can express these conditions as

$$|V(r)| \ll mc^2 \text{ and } v \ll c \tag{1.3}$$

where c is the speed of light. Note that the condition on the potential energy is equivalent to $r \gg Gm'/c^2$. Hence the deviations from Newton's theory are expected to be very small if the distance from the central mass is sufficiently large and the speed sufficiently low. For the Sun, with a mass $m' = M_\odot \cong 2.0 \times 10^{33}$ g, we have $Gm'/c^2 \cong 1.5$ km,

Table 1.1 Some laboratory measurements of the gravitational constant*

Experimenter(s)	Year	Method	G (10^{-8} dyne $\cdot$ cm^2/g^2)
Cavendish	1798	Torsion-balance deflection	6.75(±5)**
Poynting	1891	Beam balance	6.70(±4)
Boys	1895	Torsion-balance deflection	6.658(±7)
Eötvös	1896	Torsion-balance period	6.66(±1)
Luther and Towler	1982	Torsion-balance period	6.6726(±5)
Gundlach and Merkowitz	2000	Torsion-balance acceleration	6.6742(±1)
Quinn et al.	2001	Torsion balance deflection	6.6756(±3)
Armstrong and Fitzgerald	2003	Torsion balance, compensated	6.6738(±3)
Schlamminger et al.	2006	Beam balance	6.6743(±1)

* Full references for experiments before 1909 are given by Poynting (1911) and by de Boer (1984). Other references are given by Schlamminger et al. (2006).
** The number in parentheses is the experimental uncertainty in the last decimal listed.

and the condition $r \gg 1.5$ km is obviously very well satisfied, even for comets with a perihelion close to the surface of the Sun.

The gravitational constant G that appears in Eq. (1.1) is not known with the high precision of other fundamental constants. Whereas the values of e and $\hbar$ are known to eight significant figures, the value of G is known to only five significant figures. Measurements of G are difficult because of the extreme weakness of the gravitational force between masses of laboratory size. The gravitational force between masses of planetary size is not weak, but this is of no help in determining G, because only the combination Gm' (where m' is the mass of the attracting body) appears in the equations of motion of bodies with purely gravitational interactions; hence, planetary observations cannot determine the separate values of G and m'.

Table 1.1 gives selected values of laboratory measurements of G. The values are listed in chronological order; the earlier ones are included for their historical interest, and the more recent values are the best available today. Figure 1.1 shows the torsion balance used by Cavendish in his pioneering measurements of G late in the 18th century. A beam with two small masses (B, B) is suspended from a thin fiber. These small masses are gravitationally attracted by the two large lead spheres (W, W), and this results in a measurable deflection of the beam through some angle around the vertical. From the known torsional constant of the fiber and the geometry of the balance, the gravitational constant can then be calculated.

The recent measurements by Gundlach and Merkowitz (2000) and by Schlamminger et al. (2006) have given the most precise value for G. Surprisingly, these determinations agree almost exactly, although they were performed by entirely different methods.

Gundlach and Merkowitz used a small, delicate torsion balance with four "large" masses of 8 kg each mounted on a rotating turntable. During each experimental run, the turntable was accelerated at exactly the rate needed to keep the small-mass beam at a fixed angular distance from the large masses, with the torsion fiber in equilibrium (no twist in the fiber). This procedure eliminates "noise" from the gravitational background

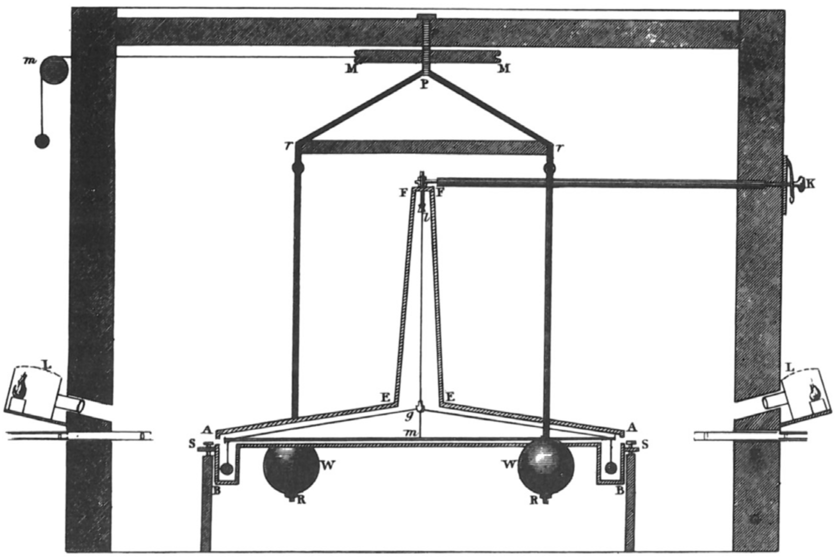

Fig. 1.1 The apparatus used by Cavendish. The large lead spheres (W, W) attract the small spheres (B, B) which are attached to the beam of the torsion balance. (From Cavendish, 1798)

and errors arising from irregularities in the torsion contributed by the twisted fiber of the Cavendish arrangement.

In contrast, the apparatus of Schlamminger et al. was colossal, with two large masses of 7.5 metric tons, consisting of pure mercury in two large cylindrical tanks, placed alternatively below or above two test masses of copper of about 1 kg each. The test masses were hung from the beam of an accurate beam balance, which registered the change in force on the test masses when the large masses were shifted from below the test masses to above them (see Fig. 1.2). Mercury was selected as the material for the large masses because its uniform density permits accurate calculation of the gravitational force exerted by each large mass on each test mass. The observed magnitude of the change in force between the two alternative configurations shown in Figure 1.2 then permits the evaluation of the gravitational constant G.

1.2 Tests of the inverse-square law

Is it possible that there are deviations from the inverse-square law at large distances or at small distances? By "large distances" we mean distances of up to 10^4 or 10^5 light-years; such distances are large compared with the dimensions of the Solar System, but small compared with the typical dimensions of the universe.[1] To investigate deviations from the inverse-square law, it is expedient to begin with the general mathematical constraints that relativistic field theory imposes on possible alternatives to the inverse-square law.

[1] At very large distances (more than 10^7 light-years), there may be cosmological deviations from the $1/r^2$ force (see Section 7.3). These deviations are not our concern in the present context.

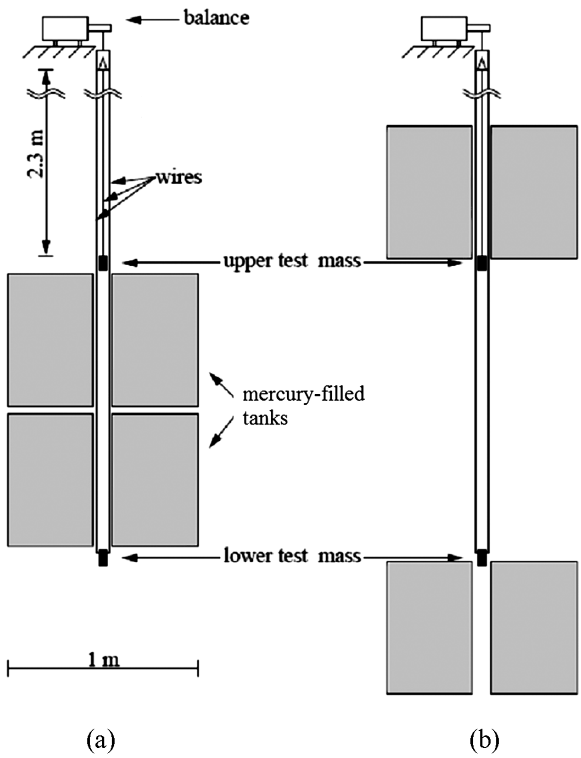

Fig. 1.2 The apparatus used by Schlamminger et al. The two test masses are suspended by wires from the balance arm at the top, and these wires pass through the axial holes of the two large mercury-filled tanks (gray). (a) In the first configuration, the two tanks are adjacent, and the gravitational pull is downward on the upper test mass, upward on the lower test mass. (b) In the second configuration, the two tanks are widely separated, and the pulls on the test masses are reversed. The beam balance detects this change of pull on the test masses. (From Schlamminger et al., 2006)

It is easiest to express these constraints in terms of the potential. The inverse-square law has the special potential given by Eq. (1.2), whereas the general potential consistent with field theory turns out to be

$$V(r) = -\alpha \frac{Gmm'}{r} e^{-r/\lambda} \tag{1.4}$$

where α and λ are constants. This is called a *Yukawa potential*; obviously, the $1/r$ potential (1.2) is a special Yukawa potential with $\alpha = 1$ and $\lambda = \infty$. The constant λ is called the *range* of the potential – if the distance r appreciably exceeds λ, the potential and the force it produces become negligible. Besides (1.4), the only other possibility is some superposition of several Yukawa potentials, which would mean that we are dealing with several gravitational fields. In this case, the long-distance behavior of the net potential is dominated by the Yukawa potential of the largest λ, because this potential will linger to the largest distance.

If we focus on the Yukawa potential of the longest λ, what do the available observational data tell us about the value of this λ? We know that the range of the gravitational force is very long – we know that our Galaxy as well as clusters of galaxies are held together by gravitation, which implies that the gravitational potential does not deviate much from $1/r$ out to distances of $r \cong$ size of cluster $\cong 10^{24}$ cm. From this we can conclude that $\lambda > 10^{24}$ cm.

Incidentally, the value of λ is related to the mass of the graviton, a (hypothetical) particle of spin two, which is to gravitation what the photon is to electromagnetism.

According to relativistic quantum theory, the mass of the graviton is inversely proportional to the range of the Yukawa potential,

$$m_\Gamma = \hbar/\lambda c \tag{1.5}$$

If the gravitational force is inverse-square, the range of the force is infinite, and the mass of the graviton is zero. If we rely on the observational limit $\lambda > 10^{24}$ cm, we obtain $m_\Gamma < 10^{-62}$ g for the graviton mass (Goldhaber and Nieto, 2010).

However, this mass limit rests on the assumption that the observed discrepancies between the observed orbital velocities of stars in the outer reaches of galaxies and the visible mass of these galaxies are accounted for by dark, invisible, mass (sometimes called "missing mass"). This extra dark mass supposedly makes a large contribution to the total mass M of the galaxy, and thereby endows the orbiting stars with a larger orbital velocity, according to the usual relation between centripetal acceleration and gravitational force, $v^2/r = GM/r^2$. The existence of such dark mass has been challenged, and several schemes have been proposed for modifications of the behavior of gravity at large distances. For instance, the MOND scheme (*MO*dified *N*ewtonian *D*ynamics) proposed by Milgrom (1983) conjectures that at large distances the strength of the gravitational force is modified from $1/r^2$ to $1/r$, so in the outer reaches of galaxies the force of gravity remains much stronger than expected from Newton's law. As a purely ad hoc scheme, MOND has not found much favor among astronomers, who reckon that invisible, dark mass is the lesser of two evils. In any case, for galaxies, the proposed modification of gravity would come into play only at distances of about 10^{22} cm, so for shorter distances we can still rely on the $1/r^2$ law.

It is of some interest to compare the 10^{24} or 10^{22}-cm limit with the analogous observational limit on the mass of the photon that can be set by examination of galactic magnetic fields. Such magnetic fields are known to extend over distances of 10^{21} cm, and with this limit on λ we obtain $m_\gamma < 10^{-59}$ g for the photon mass. The mass of the graviton is constrained to a smaller value than the mass of the photon because gravitational fields are observed over larger distances than electromagnetic fields.

Because the value of λ for the long-distance part of the gravitational potential is certainly very large, and because a value $\lambda = \infty$ is consistent with our observational data, we will hereafter assume throughout this chapter that at large distances the gravitational potential reduces to the Newtonian $1/r$ potential, so there are no long-distance deviations from the inverse-square law.

There remains the question of possible deviations from the inverse-square law at short distances, generated by an additional Yukawa potential with a short range λ. Such deviations arise in speculative theories involving extra dimensions, such as string theories. These extra dimensions are supposed to be tightly curled up on a short scale of distance, so they remain unobservable. The extra dimensions probably come into play only on a distance scale of 10^{-33} cm, the Planck distance that characterizes the scale of gravitational quantum fluctuations. However, according to some radical conjectures, one or more of the extra dimensions might have a distance scale much larger than 10^{-33} cm, maybe as large as a fraction of a millimeter. An extra dimension with such a length scale would escape observation if all the familiar particles are somehow confined to three dimensions (or, more precisely, four dimensions, if we count the time dimension), and only gravity spreads into the extra millimetric dimension. In our 3-D space, the only observable effect would then be a modification of the behavior of gravity at millimeter distances. For distances larger than the size of the extra dimension, the modification of gravity can be approximately represented by a superposition of the Newtonian $1/r$ potential and an extra Yukawa potential with a finite value of λ,

$$V(r) = -\frac{Gmm'}{r} - \alpha\frac{Gmm'}{r}e^{-r/\lambda} \tag{1.6}$$

In theories with an extra dimension, the constant λ is expected to be of the order of magnitude of the size of the extra dimension, and α is expected to be of the order of magnitude of 1.

Independently of the motivation underlying Eq. (1.6), experimenters often use this equation to parametrize deviations from the Newtonian potential, not only on millimetric scales but also on Solar-System scales. Note that Eq. (1.6) gives an inverse-square force at large distances, but a complicated behavior for distances smaller than λ. However, for $r \ll \lambda$ the potential reduces to

$$V(r) \cong -\frac{Gmm'}{r}(1+\alpha) \tag{1.7}$$

so the force reverts to an inverse-square force, with a modified value $(1+\alpha)G$ for the gravitational constant. If the range λ of the Yukawa potential in Eq. (1.6) is of the order of, say, a few hundred meters, then the gravitational constant measured in laboratory experiments is $(1+\alpha)G$, whereas the gravitational constant for interplanetary forces is G.

Limits on λ and on α can be extracted from a variety of orbital, geophysical, and laboratory observations and experiments.

Orbital Observations. High-precision measurements of the distances to Mercury, Venus, Mars, and Jupiter have been obtained by radar ranging, either with radar signals directly reflected by the surface of the planet or with signals returned by a transponder on a spacecraft during a flyby or while in orbit around the planet. In combination with determinations of planetary orbital periods, obtained by traditional astronomical observations, the distance data permit a rigorous test of Kepler's third law and therefore a test of the inverse-square law. A recent analysis of all the available data imposes a tight limit on the strength of the extra Yukawa potential, $|\alpha| < 10^{-8}$ for λ between 10^{10} and 10^{14} cm (Fischbach and Talmadge, 1999).

An analogous test can be performed for the orbits of the Moon or of artificial satellites around the Earth. The distance to the Moon has been measured with high precision by laser ranging, by means of laser pulses reflected by corner reflectors placed on the Moon during the Apollo 11 mission. Such precise measurements have also been performed on the LAGEOS artificial satellite. The lunar laser-ranging data show no detectable deviations from the inverse-square law and set a tight limit of $|\alpha| < 10^{-10}$ for $\lambda \approx 10^{10}$ cm (Fischbach and Talmadge, 1999; also reviews by Adelberger, Heckel, and Nelson, 2003; Adelberger et al., 2009; and Newman, Berg, and Boynton, 2009).

Geophysical measurements. Geophysical investigations of the inverse-square law hinge on a method for the determination of G first proposed by Airy in 1856. This method exploits the variation of the acceleration of gravity with depth below the surface of the Earth (or height above the surface). If we descend into a deep mine shaft, we find that g varies with depth. For a uniform-density sphere, g would decrease linearly with depth. However, the Earth is not of uniform density, and g at first increases with depth and then decreases. For illustrative purposes, assume that the mass distribution of the Earth is spherical, with a density $\rho(r)$ and a mass $M(r)$ enclosed within the radius r. According to the familiar Gauss law (which applies to gravitation as it applies to electrostatics), the acceleration $g(r)$ depends only on the mass enclosed within the radius r,

$$g(r) = G\frac{M(r)}{r^2} \tag{1.8}$$

and

$$\begin{aligned} \frac{dg}{dr} &= -2G\frac{M(r)}{r^3} + \frac{G}{r^2}\frac{dM(r)}{dr} \\ &= -\frac{2g(r)}{r} + \frac{G}{r^2}4\pi r^2 \rho(r) \end{aligned} \tag{1.9}$$

With this equation, the value of G can be calculated from the measured values of $g(r)$ and dg/dr, provided we know the density ρ. Equation (1.9) is only a rough approximation; for an accurate determination of G via this method, we must also take into account the rotation of the Earth and its ellipsoidal shape.

The Airy method cannot achieve the precision of laboratory determinations of G. However, it can be exploited to test the inverse-square law, as follows. Find some region where the density ρ is known, and measure g as a function of depth in the ground; then calculate G from g and dg/dr, by means of Eq. (1.9) or, rather, by means of the accurate version of this equation. If the result of this determination of G agrees with the laboratory value $G = 6.6743 \times 10^{-8}$ dyne $\cdot$ cm^2/g^2, then the result verifies the inverse-square law; if not, then it disproves the inverse-square law. Attempting to apply the Airy method, experimenters have measured gravity as a function of depth in mine shafts (Stacey et al., 1987), in boreholes in the ground (Thomas and Vogel, 1990), in the Greenland icecap (Ander et al., 1989; Zumberge et al., 1990), and underwater in the ocean (Stacey and Tuck, 1981; Zumberge et al., 1991). In a variant of the Airy method, experimenters have also measured gravity as a function of height on TV transmitter towers several hundred meters high (Eckhardt et al., 1988; Thomas et al., 1989).

Unfortunately, such geophysical tests of the inverse-square law are bedeviled by the presence of underground density variations. In all these experiments, the investigators seek to detect a deviation from the inverse-square law by comparing their measured values of g with the values calculated from the inverse-square law. However, the calculations hinge on explicit or implicit assumptions about the homogeneity of the underground material, and it is almost always easy to construct models of slightly inhomogeneous mass distributions that account for the measured data *without* invoking any deviation from the inverse-square law (Parker and Zumberge, 1989).

The measurements of the 1980s were mostly motivated by a proposal by Fischbach et al. (1986), who resuscitated an earlier discarded proposal by Lee and Yang for a "fifth force" proportional to baryon number. In contrast to Lee and Yang – who had assumed that their baryon force was a $1/r^2$ force, with a $1/r$ potential – Fischbach et al. assumed that their baryon force was based on a Yukawa potential. Such a baryon force would produce two observable effects: It would alter the behavior of the force with distance [as in Eq. (1.6)], and it would cause inequalities in the free-fall accelerations of bodies toward the ground, because samples of equal masses but different baryon numbers would experience different net forces. The fifth-force proposal stirred up considerable interest, especially when some measurements of weight as a function of height on towers suggested a deviation from inverse-square. This deviation was later found to be an illusion arising from problems with the data analysis, and the fifth force was finally laid to rest by comparisons of different mass samples by means of torsion balances, which showed that there was no effect attributable to baryon number (Adelberger et al., 2009; Gundlach, Schlamminger, and Wagner, 2009).

Laboratory Measurements. A simple way to test the inverse-square law is to compare the results of determinations of G by different experimenters. Most of these determinations were made with torsion balances. If the force between the masses deviates from the inverse-square law, then the result of a determination of G will depend on the size of the torsion balance. Cavendish used a large balance, with a beam of about 2 m and a distance of more than 10 cm between the attracting masses; modern versions of the experiment used beams as small as 2 cm and a correspondingly smaller distance between the attracting masses. The agreement between such determinations of G suggests that there are no substantial deviations from inverse-square. However, in view of the rather large experimental uncertainties in the determinations of G, the comparison does not yield any stringent limits (de Boer, 1984).

Better limits on deviations from the inverse-square law have been obtained by experiments specifically designed for this purpose. An elegant experiment by Spero et al. (1980) used a torsion balance to explore the force field inside a long cylindrical shell (see Fig. 1.3). If, and only if, the inverse-square law is valid, the force that such a cylindrical shell exerts on a small spherical mass in its interior is exactly zero. In the experiment, the small mass in Fig. 1.3 was moved back and forth relative to the cylindrical shell, to see whether it experiences any force when near the wall of the cylinder. The absence of any detectable force set a limit of about $|\alpha| < 10^{-4}$ for $\lambda \approx$ a few centimeters.

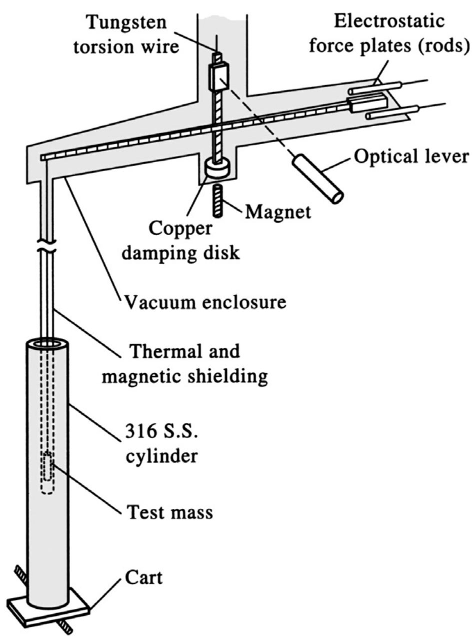

Fig. 1.3 Torsion balance with one mass suspended in the interior of a long cylindrical shell. (From Spero et al., 1980)

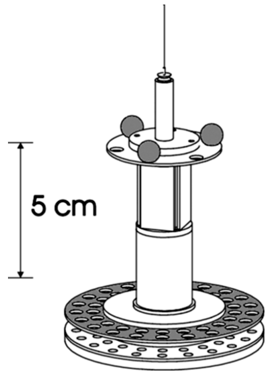

Fig. 1.4 In this torsion balance the upper perforated plate (gray) is attached to a cylinder, which is suspended from the torsional fiber (a short segment of fiber can be seen at the top). The three spherical balls are used for fine adjustments of the instrument. (From Kapner et al., 2007)

Similar results were obtained by torsion-balance experiments that compared the force exerted by a small mass placed near a torsion balance with the force exerted by a larger mass placed farther from the torsion balance (Chen, Cook, and Metherell, 1984). Somewhat larger ranges of λ, reaching somewhere above 10 m, were explored with a "gravity gradiometer" that directly tested that the gradients in the gravitational field of a mass are those appropriate to a $1/r^2$ force (Hoskins et al., 1985).

Less stringent limits on $|\alpha|$, but for larger hypothetical values of λ, were obtained by experiments performed with hydroelectric pumped-storage reservoirs. The water level of such reservoirs often rises or falls by tens of meters in just a few hours, and the change of gravity that this produces in the region above the water depends on α and λ. The change of gravity can be measured with a beam balance that has one of its pans above the water level and the other pan below water level, all in a long waterproof tube (Stacey et al., 1987). Alternatively, the change of gravity can be measured with a high-precision gravimeter, that is, a delicate spring balance (Müller et al., 1990).

Several recent experiments were designed to search for Yukawa potentials with values of λ of a millimeter or less, which are the values of greatest interest for theories with extra dimensions. An experiment by Kapner et al. (2007) used a torsion balance of a special design (see Fig. 1.4) in which the beams holding the small and the large masses of the Cavendish balance are replaced by plates with circular holes around their circumferences. The upper plate is suspended from a torsional fiber, but is placed very close to the lower plate (0.05 mm, in some experimental runs). Whenever the holes in the upper, suspended, plate are not aligned with those in the lower plate, the masses in the interstices between the holes in the upper and lower plates attract each other and exert a detectable torque on each other (this torque can be conveniently described as due to an effective repulsion of the holes). However, below the bottom plate, there is a second, hidden bottom plate (its edge is barely visible in Fig. 1.4), also with circular holes. This second bottom plate has its holes aligned with the interstices of the first bottom plate; furthermore, this second bottom plate is more massive, in proportion to the square of the distance from the upper, suspended, plate. The net result is that the bottom plates in combination exert (almost) no torque on the suspended plate. But this cancellation of the effects of the two bottom plates fails if the force is not inverse square, and thus the torsion balance is able to detect deviations from the inverse-square law.

This experiment established that $|\alpha| < 10^{-2}$ for $\lambda > 0.2$ mm, from which it can be concluded that the size of the extra dimension, if any, is smaller than that. The experiment