

General Relativity and Gravitation

The General Theory of Relativity (GR), created by Albert Einstein between 1907 and 1915, is a theory both of gravitation and of spacetime structure. It is based on the assumption that matter, via its energy-momentum, interacts with the metric of spacetime, which is considered (in contrast to Newtonian physics and SPECIAL RELATIVITY) as a dynamical field having degrees of freedom of its own (GRAVITATIONAL RADIATION). In brief, ‘matter tells spacetime how to curve, and spacetime tells matter how to move’ (J A Wheeler). GR is a generalization of Newton’s theory of gravitation and of Special Relativity, which are both approximations to GR under appropriate conditions (see below). More generally, GR provides a common, coherent basis for classical (as opposed to quantum), macroscopic physics. Its relation to quantum physics is not yet well understood (QUANTUM GRAVITY). While this poses an important problem of principle, it does not present a practical difficulty to astrophysics, since the atomic and subatomic scales, on which quantum laws operate, are very small compared to macroscopic scales, where gravity dominates.

According to present fundamental physics, the only one among the four basic interactions which acts between all kinds of matter is the gravitational interaction. Since it always acts as an attraction and thus cannot be shielded, and because its range is unbounded, it dominates the behavior of matter on large scales in spite of its extreme weakness. Therefore, gravity plays a significant role in nearly all parts of astronomy and astrophysics. (In the atomic and subatomic domain gravity is unmeasurably weak and totally negligible, as is highlighted by the ‘fact’ that the gravitational force between a proton and an electron is weaker than the Coulomb force by the factor 5×10^{-38} .)

Relativistic gravity is important in the following areas of research: high-precision astrometry (REFERENCE FRAMES AND TIME SCALES IN GR, SPACE AND TIME REFERENCES: CONCEPTS), compact objects such as NEUTRON STARS and BLACK HOLES and systems thereof (binary PULSARS, low-mass X-RAY BINARY STARS), active galactic nuclei and quasars, supernovae and gravitational collapse, identification of dark stars and planets (MICROLENSING), cosmology (COSMOLOGY: STANDARD MODEL, DARK MATTER, COSMOLOGICAL CONSTANT, distribution of luminous and DARK MATTER via weak GRAVITATIONAL LENSING). In the near future a new window onto the universe is expected to open when GRAVITATIONAL WAVES become observable.

Though GR covers a larger range of phenomena than Newton’s theory, the latter provides a satisfactory description whenever the system under study is (nearly) isolated, the relevant velocities are small compared to the speed of light, $\frac{v}{c} \ll 1$, and bodies are not too compact, $\frac{2GM}{Rc^2} \ll 1$. Moreover, Newton’s theory is used to describe the basic facts of gravity; it often serves as a point of departure for approximations in GR, and its shortcomings motivate the introduction of GR and help to understand

and highlight GR’s achievements. For these reasons, the Newtonian theory of spacetime and gravity is reviewed before GR is considered.

Newton’s theory of gravity

Can be summarized as follows.

- 1(a) On spacetime, the set of events ‘here-now’, there exists an absolute time, measurable by means of arbitrarily moving, undisturbed clocks. Distances between bodies at one instant of time satisfy the laws of Euclidean geometry. Hence, motions of bodies can be referred to rigid frames of reference associated with coordinates (t, x^a) specifying times t and positions $(x^a) = \vec{x}$.
- 1(b) There exist preferred reference frames, inertial frames, with respect to which the motions of N bodies which form a nearly isolated system (like the solar system or a double star), are given by Newton’s gravitational equations of motion

$$\ddot{\vec{x}}_A = \sum_{B \neq A} \frac{M_B}{r_{AB}^2} \vec{e}_{AB} \quad (1)$$

which form the basis of classical celestial mechanics. The positive constants M_A are called the gravitational masses of the bodies, r_{AB} denotes the distance between the simultaneous positions of the bodies A and B , \vec{e}_{AB} indicates the unit vector pointing from \vec{x}_A to \vec{x}_B , and dots signify time differentiation. Equation (1) characterizes an inertial frame as a rigid frame with respect to which the equations of motion do not contain velocity dependent (Coriolis acceleration) terms and relative to which the center of mass of such a system moves uniformly in a straight line.

- 1(c) If a body, or a part of a body, is subjected not only to gravity, its motion relative to an inertial frame is governed by the equation

$$m(\ddot{\vec{x}} - \vec{g}(t, \vec{x})) = \vec{F}. \quad (2)$$

Here, the positive constant m is the inertial mass of the body, \vec{g} denotes the acceleration the particle would have if subjected to gravity only, and \vec{F} is the non-gravitational force acting on the particle. The product $m\vec{g}$ is traditionally called the gravitational force, although it is not needed to describe purely gravitational motions, see equation (1). In contrast to other forces, the gravitational force depends on the particle solely through the latter’s inertial mass.

- 1(d) Measurements indicate that the ratio gravitational mass/inertial mass for laboratory-size bodies has a universal value. This ratio,

$$\frac{M}{m} = G = 6.673(10) \times 10^{-8} \text{ cm}^3 \text{ g}^{-1} \text{ s}^{-1} \quad (3)$$

though far less precisely known than most other physical constants, is assumed to have a universal

value, the constant of gravitation. Accordingly, M and m are considered as different measures of one physical quantity, ‘mass’. Note, however, that the gravitational mass of the Sun defined in accordance with equation (1) is known about two orders of magnitude more accurately than G : $M_{\odot} = 1.32704(2) \times 10^{26} \text{ cm}^3 \text{ s}^{-2}$.

- 1(e) If matter is described as continuously distributed, with (inertial) mass density ρ , velocity \vec{v} and pressure p , the effect of gravity on motion is expressed, in generalization of (2), as

$$\rho \left(\frac{\partial \vec{v}}{\partial t} + \vec{v} \cdot \vec{\nabla} \vec{v} - \vec{g}(t, \vec{x}) \right) = -\vec{\nabla} p \quad (4)$$

and the gravitational acceleration \vec{g} can be obtained from the gravitational potential

$$\phi(t, \vec{x}) = - \int \frac{G\rho(t, \vec{x}')}{|\vec{x} - \vec{x}'|} d^3x', \quad (5)$$

by taking its gradient,

$$\vec{g} = -\vec{\nabla}\phi. \quad (6)$$

Equations (5) and (6) generalize the expression for the acceleration of equation (1). For a spatially bounded mass distribution, the law (5) can be replaced by Poisson’s equation

$$\Delta\phi = 4\pi G\rho \quad (7)$$

together with the boundary condition

$$\lim_{|\vec{x}| \rightarrow \infty} \phi(t, \vec{x}) = 0. \quad (8)$$

- 1(f) The laws (1) and (2) for particles, or (4)–(8) for spatially bounded matter distributions, determine the motions of all matter elements uniquely if its positions and velocities at one time are specified. Without the boundary condition (8), the laws (6) and (7) do not determine \vec{g} , and thus do not determine the motion; Newton’s theory applies only to systems which can be treated as isolated, but not to the apparently spatially unbounded cosmic matter distribution.

All laws stated above take the same forms in all inertial frames, they are invariant under Galilean transformations which express symmetries of Newtonian spacetime.

Shortcomings of the Newtonian theory

Which led, first, to the special relativity theory (SR) and, later, to GR, include the following facts and considerations.

- 2(a) The assumptions about time and space listed above under 1(a) are incompatible with electro-dynamical and optical phenomena involving moving bodies, with well established facts about moving clocks and

with many other experimental results. Those facts for which gravity can be neglected or, more precisely, where the temporal and spatial scales on which the gravitational acceleration changes, are negligible compared to the relevant scales, are well accounted for by Special Relativity. According to SR, there is no universal, absolute time. Rather, each inertial frame has its own measure of time and also of distance, and the coordinates of inertial frames are related not by Galilean, but by Poincaré transformations. The metric of spacetime determines, for each pair of events (t_1, \vec{x}_1) and (t_2, \vec{x}_2) , whether there exists an inertial frame with respect to which the events: (a) are simultaneous; or (b) happen at the same point of space; or (c) whether the events can be connected by a light signal. Formally, the metric is given, in any inertial frame, by the interval

$$\Delta s^2 = -c^2(t_2 - t_1)^2 + (\vec{x}_2 - \vec{x}_1)^2 \quad (9)$$

and the alternatives just listed correspond to

- (a) $\Delta s^2 > 0$
- (b) $\Delta s^2 < 0$
- (c) $\Delta s^2 = 0$.

According to SR, no interaction or signal can be transmitted faster than with the speed of light in vacuum, c .

- 2(b) In Newtonian theory, inertial frames can only be identified globally for isolated systems. ‘Real’ inertial frames associated with spatially separated, nearly isolated systems such as galaxies, are relatively accelerated and cannot be fitted into a single, ‘truly inertial’ frame of reference. Also, freely falling test particles which can be observed locally, exhibit relative accelerations in any real, inhomogeneous gravitational field; ‘free’ particles exactly obeying the law of inertia do not exist since all particles ‘feel’ gravity in the same way, see equation (2) with $\vec{F} = 0$.
- 2(c) As stated in 1(f), Newtonian theory does not provide predictive dynamical laws for spatially unbounded matter distributions.
- 2(d) Newtonian celestial mechanics is unable to account for the ‘anomalous’ advance of the perihelion of Mercury’s orbit, noticed first by Leverrier in 1859.

General Relativity Theory overcame these shortcomings and predicted additional phenomena.

Basic concepts and laws of General Relativity

- 3(a) In order to preserve locally the approximate validity of SR, it is assumed that, for any event E of spacetime, there exists a coordinate system (t, x^a) such that the statements made in connection with equation (9) are valid for spacetime points P which are ‘close to E ’, in the sense that

$$\Delta s^2 = -c^2 \Delta t^2 + \Delta \vec{x}^2 + 0(4). \quad (10)$$

(The symbol $O(4)$ indicates that additional terms are at least as small as fourth powers of the leading ones, e.g., $c^2 \Delta t^2 \cdot \Delta \bar{x}^2$, and thus are negligible for small displacements $\Delta t, \Delta \bar{x}$ connecting E to P.) Such coordinate systems are called locally inertial at E. They may be visualized/realized as connected with a small, non-rotating, freely falling cabin, like a space shuttle. Two such coordinate systems, referring to different ‘central’ events E_1 and E_2 , represent approximately inertial systems near E_1 and E_2 , respectively, but in contrast to Newtonian or SR inertial systems they are nonlinearly related, and thus exhibit relative accelerations between those reference frames. The basic assumption, symbolized in equation (10), at the same time preserves SR as a local approximation and weakens the concept of an inertial frame as suggested by the considerations in 2(b). Mathematically this assumption means that spacetime carries a Lorentz metric whose components in a general coordinate system, $g_{\alpha\beta}(x^\gamma)$, are functions of those coordinates ($\alpha, \beta = 0, 1, 2, 3$). The fact that these functions in general cannot be reduced to constants in a finite coordinate domain means geometrically that spacetime is curved.

- 3(b) The most direct physical interpretation of the metric is that it determines the times shown by undisturbed clocks or natural time carriers like atoms. The motion, or ‘history’, of a clock is represented by a time-like curve (figure 1), and the proper time τ it measures is given by the ‘arc length’

$$\tau = \int |g_{\alpha\beta} dx^\alpha dx^\beta|^{1/2} \quad (11)$$

along that world line. This statement is the GR-substitute for Newton’s axiom about absolute time.

- 3(c) According to Einstein’s equivalence principle all local, non-gravitational laws of physics take the same mathematical form at the origin of a local inertial coordinate system. This principle restricts the relativity principle on which SR is based since it implies that the laws of SR hold only approximately in small domains; at the same time it generalizes its approximate validity to arbitrary gravitational fields. The principle is used to generalize laws from Special Relativity to General Relativity. Applied to the law of inertia of SR, it leads to the statement that the spacetime curve, or world line, of a freely falling particle is a straightest curve, or (timelike) geodesic, in spacetime. Applied to Maxwell’s laws of electromagnetism it provides the corresponding laws in curved spacetime, i.e., in an arbitrary gravitational field. In the geometrical optics approximation this implies that light rays (\equiv photon world lines) are null geodesics, i.e., geodesics everywhere tangent to the local light cones (figure 1). Analogously, one obtains the equations of motion for fluids or gases in gravitational fields. While in these

and similar cases of classical physics Einstein’s equivalence principle works perfectly, it proved to be complicated to generalize quantum field theory to curved spacetimes. For reasons given in the introduction this is, however, so far of no practical importance to astrophysics.

The mathematical forms of the general-relativistic laws for particles and fields show that in GR the metric $g_{\alpha\beta}$ of spacetime also plays the role of the gravitational potential which has ten components that replace the one Newtonian potential ϕ .

- 3(d) The third, and most important, ingredient of General Relativity is Einstein’s gravitational field equation. It relates the metric field of spacetime to the distribution and motion of matter and energy. More precisely, the metric tensor $g_{\alpha\beta}$ and its first and second derivatives determine the so-called curvature tensor (which measures the tidal field) and a tensor derived from it, called the Einstein tensor $G^{\alpha\beta}$. Moreover, the energy density, energy current density and stresses of matter and fields are combined into a ten-component quantity $T^{\alpha\beta}$, the stress-energy-momentum tensor. These fields are required to satisfy the (tensor) equation

$$G^{\alpha\beta} = \frac{8\pi G}{c^4} T^{\alpha\beta}. \quad (12)$$

This system of nonlinear partial differential equations replaces Poisson’s equation (7) and does for gravity what Maxwell’s equations do for electromagnetism.

- 3(e) In contrast to SR, the topology of spacetime is not given once and for all. Rather, the problem of finding a spacetime model involves to determine a four-dimensional manifold and, on it, physical fields satisfying equation (12). This aspect of GR is particularly relevant for the study of black holes and cosmology.
- 3(f) The unification of special relativity and Newtonian gravitation theory brought about by Einstein’s GR has solved an old conceptual problem emphasized especially by Ernst Mach. In Newtonian theory (and in special relativity), the structure of spacetime as exhibited in the set of inertial frames, acts on matter, but it cannot be observed independently of the motion of matter, and it is not influenced by matter. It is, in other words, an absolute, unobservable cause. To assume such a cause contradicts the scientific mode of thinking (‘widerspricht dem wissenschaftlichen Verstande’) according to Mach, and Einstein consented. According to GR, the metric of spacetime is not given once and for all, but is a dynamical field interacting with matter, acting and being acted upon; in this respect GR overcomes Mach’s objection. The dynamical nature of the metric is the most important insight of principle of GR.

Since the metric depends on the circumstances, one cannot introduce coordinates with a definite physical

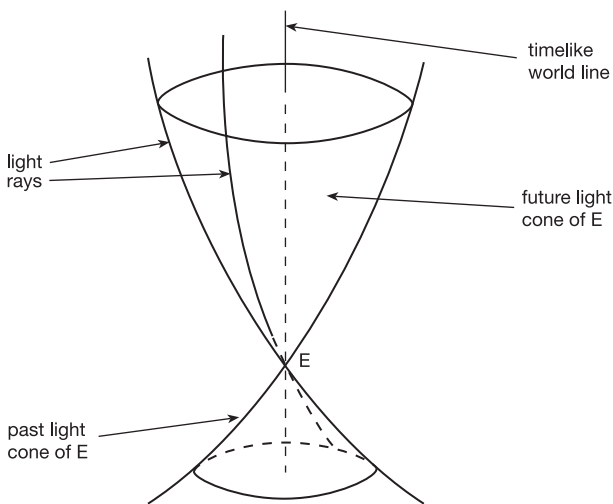


Figure 1. The future and past light cones of event E, generated by light rays, and a timelike world line which could represent a clock's motion.

meaning independently of the physical situation. Therefore, coordinates only serve to 'name' and distinguish spacetime points and to express their 'nearness' in a qualitative (topological) sense. For the same reason, the laws of the theory have to be formulated so as to hold in arbitrary coordinate systems; they have to be generally covariant. This requirement, together with the principle of equivalence, guided Einstein on his way towards GR.

Some consequences of GR

Causality

The mathematical analysis of the field equation (12) has shown that the metric (\equiv gravitational field) propagates like a wave. Its value in the vicinity of an event A can influence the field at a later event B only if B is contained in the interior of or on the future light cone of A, i.e. the field propagates not faster than light. To determine the field at B, it suffices to know the field on a 'spacelike', three-dimensional subspace S of spacetime earlier than B, on that part C of S which is contained in the past light cone of B. This is illustrated, with one space dimension suppressed, in figure 2. These statements also hold for the combined gravitational and matter fields. They generalize corresponding laws of special relativity (e.g. for electromagnetic fields), and contrast sharply with the instantaneous action-at-a-distance of Newtonian forces.

Violation of local energy-momentum conservation

The field equation (12) implies that the energy tensor satisfies an equation which generalizes the laws of local conservation of energy and momentum from SR to GR, in accordance with the equivalence principle,

$$\nabla_{\beta} T^{\alpha\beta} = 0. \quad (13)$$

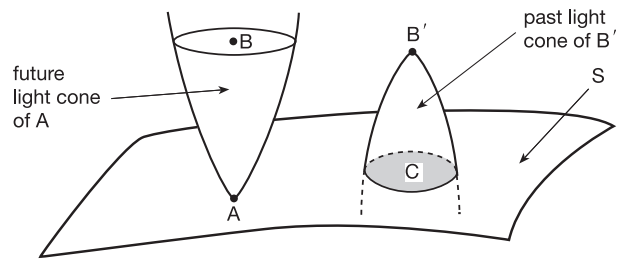


Figure 2. An event B in the interior of the future light cone of A and the region C of S which causally determines what happens at B'.

Special cases of this equation provide GR-generalizations of the equations of motion (1), (2) and (4) of Newtonian theory. The derivative operator ∇ depends on the metric and expresses that the localizable energy and momentum of matter and of non-gravitational fields are not strictly conserved, due to the action of gravity. For an isolated system, however, a global conservation law holds which includes a contribution from the gravitational field, see section Isolated systems, global energy-momentum below.

Equations of motion and gravitational waves

A fundamental task for any theory of gravitation is the prediction of the motions of bodies such as planets and stars. Approximation methods based on the gravitational field equation (12) and its consequence (13) have been developed to solve that problem. Solutions which represent the gravitational fields of well-separated, nearly spherical, slowly moving ($v \ll c$) and weakly stressed ($p \ll \rho c^2$) bodies forming an isolated system reproduce, in lowest order, the results of Newtonian dynamics and provide higher 'Post-Newtonian' order, general-relativistic corrections, see section Tests of the field equation below. These methods have been successful in describing effects in the solar system like the perihelion advance, and even the effect of the emission of gravitational waves on the motions of the binary pulsar are in agreement with the theory. We do not yet understand in detail what happens if, under the damping action of gravitational radiation, two neutron stars and/or black holes forming a binary system approach each other and then merge to form a single compact object. Such a process is expected to produce a strong gravitational wave signal, gravitational waves.

At large distances from the source, gravitational waves have properties similar to electromagnetic waves; they propagate with the fundamental speed c , are transverse, and have two independent states of polarization. However, while the patterns of two basic states of linear polarization of an electromagnetic wave form an angle of 90° as indicated symbolically by $|$ and $-$, in the gravitational case they form an angle of 45° only as indicated by $+$ and \times . This corresponds, in quantum language, to the fact that photons have spin 1 (in units of \hbar), while gravitons are expected to have spin 2.

Gravitomagnetism

It follows from equations (12) and (13) that, in addition to Newtonian-type gravitational forces depending on the mass distribution only, there exist 'gravitomagnetic' forces depending on the flow of mass, analogous to magnetic forces, which depend on the motions of the charges. Such forces are taken into account in post-Newtonian approximations, but have not yet been measured in isolation, see Tests of the field equation below.

Isolated systems, global energy-momentum

The spacetime of an isolated system in GR is required to be such that its light cones behave, at large distances from matter sources, asymptotically like those of the spacetime of special relativity. It is then possible to define a conserved total energy momentum of such a system and to show that the total energy is positive, except if there is no true gravitational field, i.e., no curvature (Yau, Witten). It is also possible to give an exact expression for parts of the total energy-momentum which are radiated out of the system between two future light cones (Bondi, Sachs).

Horizons and black holes

The causal structure of a curved spacetime, indicated in the section Causality above, sometimes has the consequence that part of spacetime can never be seen by some observers. If a spherical star collapses, that particular future light cone of the center of the star which passes the star's surface at the instant when the ratio mass/radius of the star reaches the 'Schwarzschild limit' $c^2/2G$, will never reach an outside observer (figure 3). This so-called event horizon separates those events, which can be seen from outside, from those which are hidden from sight. That part of spacetime which is thus hidden is called a black hole; the horizon is its surface. Although according to (classical) GR no radiation and no particles can be emitted from a black hole, it does have a mass and attracts other bodies accordingly. For additional information on black holes see also QUANTUM GRAVITY, where singularities of gravitational fields are also briefly discussed.

Cosmology

General Relativity allows to treat spatially unbounded matter distributions and thus provides a basis for cosmology. The basic reason for this is that the solutions to equations (12) and (13) are determined, like those of Maxwell's equations, by initial data whose influence on the evolving metric and matter propagates at most with the fundamental speed c , as explained in the section Causality. Therefore, the field at some spacetime point never depends on a spatially infinite amount of matter, as it would according to the old equations (5) or (7). (Boundary conditions can be imposed in General Relativity only on initial data specifying a solution of the field equation (12), not on the field which evolves from these data.)

To allow for a static and spatially closed world model, Einstein in 1917 supplemented his field equation (12) of 1915 by a cosmological term $-\Lambda g^{ab}$ on the right-hand side. This contribution has later been interpreted as

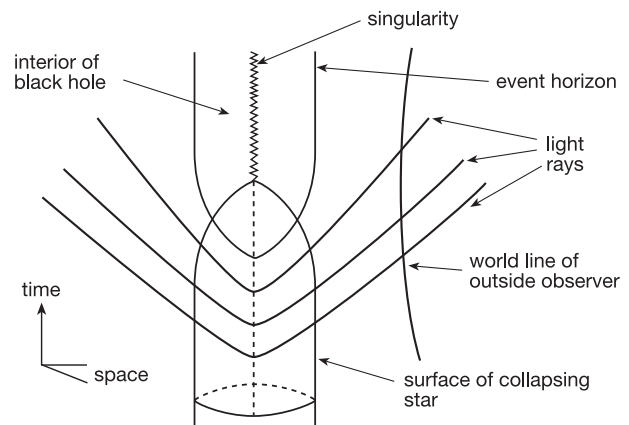


Figure 3. Collapsing star with event horizon, outside observer and light rays from the center of the star. The singularity is hidden behind the event horizon.

the expectation value of the energy-momentum of the field fluctuations of the quantum vacuum. The appropriateness of this interpretation depends on a better understanding of the quantum vacuum and its relation to gravity. So far, this interpretation has not led to a computation of the value of the COSMOLOGICAL CONSTANT Λ .

Experimental and observational tests of general relativity

Experiments and measurements supporting GR can be classified into those which test the Einstein equivalence principle and those which probe the gravitational field equation.

Tests of the Einstein equivalence principle

Tests of the first kind probe the approximate local validity of Special Relativity, i.e. the existence of local inertial frames or, equivalently, the possibility to assign a unique Lorentz metric (see 3(a) and (b)) to spacetime. From the viewpoint of GR these experiments serve to identify, relative to some arbitrary reference frame, a local inertial frame. Free fall experiments and Foucault pendulum experiments carried out in a terrestrial laboratory, for example, measure the acceleration and the angular velocity of the laboratory relative to a local inertial frame. Optical experiments and quantum interference experiments with neutron waves accomplish the same, and the concordance of the results establishes, up to experimental uncertainties, the uniqueness of the spacetime metric. The experimental fact that local inertial frames are accelerated relative to each other shows that spacetime is curved.

Tests of the field equation

Approximation methods to solve the field equation (12) show that, for a nearly isolated system of slowly, not too dense bodies, the metric differs from that of the flat spacetime of special relativity only in the time component;

the latter is determined by the Newtonian potential ϕ according to

$$-g_{00} = 1 + 2\frac{\phi}{c^2}. \quad (14)$$

This formula specifies how the gravitational potential affects clock rates and determines frequency shifts of electromagnetic waves; it has been tested for the Earth's gravitational field to within a measurement uncertainty of 2×10^{-4} (Vessot–Levine, gravity probe A). This formula is not specific to GR; it has to hold in any 'metric' theory of gravity to ensure the approximate validity of Newtonian theory. Tests which can distinguish between Einstein's theory, equation (12), and alternative theories of gravity require more accurate approximations to the metric. To obtain the light deflection, the metric

$$ds^2 = -\left(1 + 2\frac{\phi}{c^2}\right)c^2 dt^2 + \left(1 - 2\frac{\phi}{c^2}\right)d\vec{x}^2 \quad (15)$$

which exhibits space curvature, is sufficient. The validity of this approximate metric has also been confirmed by the retardation of electromagnetic waves in gravitational fields (fourth test, Shapiro effect) and by the geodetic precession (de Sitter effect) of the angular momentum of the Earth–Moon system. Additional terms are needed to measure the Schiff precession (mixed $dt dx^a$ -terms, to be tested in the gravity probe B experiment in 2000), as well as the Lense–Thirring precession of the line of nodes of satellites orbiting the Earth, due to the Earth's angular momentum. This effect appears to be observable through the orbits of the satellites Lageos I and III (Ciufolini). The famous perihelion motion of Mercury's orbit requires the additional term $+2\phi^2/c^4$ in the g_{00} term of (15). The tests here summarized have verified Einstein's field equation up to relative measurement uncertainties of a few parts in 10^3 . For more details, see GR: SOLAR SYSTEM EXPERIMENTS. Of still higher order of approximation is the effect of gravitational radiation reaction on the orbits of bodies; it has been verified in the case of binary pulsars (see GR TESTS: BINARY PULSARS). The major challenge of experimental gravity is the direct measurement of gravitational radiation which is expected to open a new window onto the universe.

Bibliography

The state of knowledge in each main area of general relativity theory is covered in at least one of the following reviews:

Hawking S W and Israel W (ed) 1997 *General Relativity: An Einstein Centenary Survey* (Cambridge: Cambridge University Press)

Held A 1980 *General Relativity and Gravitation* vols I and II (New York: Plenum)

Hawking S W and Israel W (ed) 1987 *Three Hundred Years of Gravitation* (Cambridge: Cambridge University Press)

A text book emphasizing relativistic astrophysics:

Straumann N 1984 *General Relativity and Relativistic Astrophysics* (Berlin: Springer)

Books concentrating on experimental and observational aspects:

Will C M 1993 *Theory and Experiment in Gravitational Physics* (Cambridge: Cambridge University Press)

Ciufolini I and Wheeler J A 1995 *Gravitation and Inertia* (Princeton, NJ: Princeton University Press)

J Ehlers