On Experimental Tests of the General Theory of Relativity*

L. I. Schiff
Institute of Theoretical Physics, Department of Physics, Stanford University, Stanford, California
(Received October 6, 1959)

This paper explores the extent to which the three "crucial tests" support the full structure of the general theory of relativity, and do not merely verify the equivalence principle and the special theory of relativity, which are well established by other experimental evidence. It is shown how the first-order changes in the periods of identically constructed clocks and the lengths of identically constructed measuring rods can be found without using general relativity, and how the red shift and the deflection of light can be computed from them. Only the planetary orbit precession provides a real test of general relativity. Terrestrial or satellite experiments that would go beyond supplying corroborative evidence for the equivalence principle and special relativity would be extremely difficult to perform, and would, for example, require a frequency standard with an accuracy somewhat better than one part in 10^{18}.

I. INTRODUCTION

THE general theory of relativity is now accepted as the most satisfactory theory of gravitation. This acceptance rests partly on its conceptual and structural elegance, and partly on its agreement with experimental observation. Three "crucial tests" are usually cited as experimental verifications of the theory: the red shift of spectral lines emitted by atoms in a region of strong gravitational potential, the deflection of light rays that pass close to the sun, and the precession of the perihelion of the orbit of the planet Mercury. The main purpose of the present paper is to examine to what extent the full formalism of general relativity theory is called upon in the calculation of these three effects, and to what extent they may be correctly inferred from weaker assumptions that are well established by other experimental evidence.

It was recognized by Einstein,[1] four years before the advent of general relativity, that the red shift can be computed on the basis of the equivalence principle. In the same paper he also attempted to calculate the deflection of light, but obtained half the correct value. In spite of this, it is actually possible to obtain the light deflection, as well as the red shift, in a valid manner without using the full theory. Although such a derivation is probably well known, there does not seem to be a publication that describes it.

The availability of satellites and of extremely accurate frequency standards has in recent times stimulated interest in broadening the experimental basis of general relativity theory. When such experiments are being considered, it is important to understand the extent to which they support the full structure of general relativity theory, and do not merely verify the equivalence principle and the special theory of relativity; these latter are already established with adequate accuracy by the Eötvös experiments[2] and a great deal of work with high-energy particles. To this end, we first use the equivalence principle and special relativity to relate time measures and length measures at different places in a gravitational field, and compute the red shift from the time relation (Sec. II). Then from both the time and length relations the correct value for the deflection of light is obtained (Sec. III). In Sec. IV, several implications of these results are mentioned.

II. COMPARISON OF TIME AND LENGTH MEASURES; THE RED SHIFT

Two identically constructed clocks are placed at rest, a distance h apart along the lines of force in a uniform or nearly uniform gravitational field of acceleration g, as in Fig. 1(a). In accordance with the equivalence principle, any comparison of the periods of these clocks can be made as well in a gravitation-free region, in which they are accelerated upward with the acceleration g, as in Fig. 1(b). We accomplish this by comparing

* Supported in part by the United States Air Force through the Air Force Office of Scientific Research.

[1] A. Einstein, Ann. Physik **35**, 898 (1911).

[2] R. v. Eötvös, D. Pekár, and E. Fekete, Ann. Physik **68**, 11 (1922).

clocks A and B in turn with a third identically constructed clock C, which is permanently at rest, as they sweep by in near coincidence. We assume that the fact that A and B are undergoing acceleration as they pass C does not affect this comparison. Since C is at rest in a gravitation-free region, it is part of an inertial coordinate system, and makes a suitable standard for comparing A and B with each other.

Suppose that clock A has upward speed v_A when it passes C. Then if the period of C is T, the period of A that is seen by an observer on C is, according to special relativity,

$$T_A = T(1-v_A^2/c^2)^{-\frac{1}{2}} \approx T(1+v_A^2/2c^2), \quad (1)$$

where the approximation assumes that the speed of light c is much greater than v_A. Similarly, when clock B passes C with speed v_B a second comparison shows that the period of B observed by C is

$$T_B = T(1-v_B^2/c^2)^{-\frac{1}{2}} \approx T(1+v_B^2/2c^2). \quad (2)$$

On eliminating T between Eqs. (1) and (2), we find to the same approximation that

$$T_B \approx T_A[1+(v_B^2-v_A^2)/2c^2] = T_A(1+gh/c^2), \quad (3)$$

since $v_B^2 = v_A^2 + 2gh$. Thus the inertial observer on C can inform both A and B that the period of clock B exceeds that of clock A by the fractional amount (gh/c^2).

We now argue that in a nonuniform gravitational field we can replace the quantity gh in Eq. (3) by the difference in gravitational potential between the positions of clocks A and B. This implies that we make a series of intercomparisons between a number of clocks so

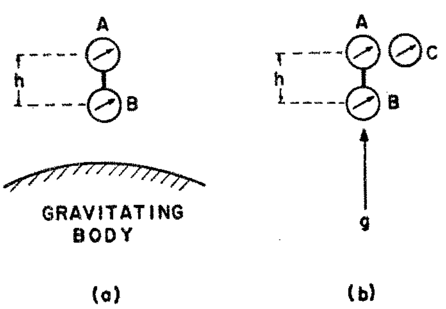

FIG. 1. (a) Two identically-constructed clocks, A and B, are at rest in a gravitational field. (b) The gravitating body is replaced by an upward acceleration g of clocks A and B, and a stationary clock C is introduced to compare their periods.

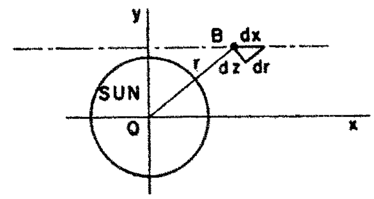

FIG. 2. A light pulse at B passes close to the limb of the sun, moving in the positive x direction.

arranged that the field is nearly uniform from one to the next. Then if the field arises from a spherically symmetric mass M, Eq. (3) becomes

$$T_B \approx T_A[1+(GM/c^2r_B)-(GM/c^2r_A)], \quad (4)$$

where G is the universal constant of gravitation, and r_A and r_B are the distances of clocks A and B from the center of the gravitating mass.

The gravitational red shift is easily obtained from Eq. (4). Let B be a clock, or more realistically an atom, at the surface of a star, and A an identical atom at the surface of the earth $(r_A \approx \infty)$ where the velocity and gravitational potential can be neglected. Then Eq. (4) tells us that the stellar atom vibrates more slowly than the terrestrial atom. Furthermore, the entire system is stationary in time, so that no vibrations are gained or lost in transit from star to earth. Thus the spectral lines that reach the earth are shifted toward the red by the fractional amount (GM/c^2R), where R is the radius of the star.

The same kind of comparison can be made for measuring rods. If we compare two identically constructed rods that are laid along the field lines (radial direction), we find in place of Eq. (4) that the lengths of rods A and B are related by

$$L_B \approx L_A[1-(GM/c^2r_B)+(GM/c^2r_A)]. \quad (5)$$

The sign differences between Eqs. (4) and (5) occur because the Lorentz time dilation used in the first derivation is replaced by the Lorentz length contraction in the second derivation.

Finally, it is easily seen that a comparison of two identically constructed rods that are laid perpendicular to the field lines (tangential direction) leads to the relation

$$L_B = L_A. \quad (6)$$

III. DEFLECTION OF LIGHT

A light pulse, instantaneously at B (see Fig. 2), passes close to the limb of the sun while moving in the positive x direction. An observer A on the earth, whose velocity and gravitational potential can be neglected, wishes to plot the path of this light pulse. He can do this by using Huygens' principle if he knows how the plotted speed of the pulse depends on its position and its direction of motion. This in turn can be expressed in terms of the local speed measured by a stationary observer at the instantaneous position of the light pulse, by using Eqs. (4)–(6) to convert the units of time and length that define the local speed into those that define the plotted speed.

We call the stationary observer B, and set up two other observers B' and B'' in the following way. At the instant under consideration B' and B'' are at the same point in space, away from gravitational fields, and all three have zero relative velocity. B' is accelerated with the value of the acceleration of gravity at B, and B'' is permanently at rest. Because of the equivalence principle B and B' agree on all measurements made at this instant. Also, in accordance with the assumption made at the beginning of Sec. II, the acceleration of B' with respect to B'' does not affect their comparison of time and length intervals at this same instant. Thus the three observers agree on their measurements of first-order infinitesimal time and length intervals; in higher order the first-order changes in relative position and velocity will in some cases destroy the agreement. Since speed can be expressed as the ratio of first-order infinitesimal length and time intervals, all three observers will measure the same speed for the light pulse. But B'' is part of an inertial coordinate system and so, according to special relativity, finds that the speed of light is c in all directions. We conclude that B also observes the speed c for the light pulse.

We now convert the local speed c, observed by B, to the plotted speed c' from which A will compute the path of the light pulse. B finds that the light pulse travels dx units of length during $dt=dx/c$ units of time. But A knows, having been informed by C, that B's unit of time is longer than A's in the ratio T_B/T_A given by Eq. (4). Thus the time interval that A must use in plotting the path of the light pulse is

$$dt'=dt(T_B/T_A)=dt[1+(GM/c^2r)]. \qquad (7)$$

We can express dx as the resultant of radial and tangential components, dr and dz, as shown in Fig. 2. Then A knows that B's unit of radial distance is shorter than A's, so that, in accordance with Eq. (5), A must use the radial interval

$$dr'=dr[1-(GM/c^2r)].$$

In similar fashion it follows from Eq. (6) that $dz'=dz$. Thus the interval dx' that A must use in plotting the path of the light pulse is

$$\begin{aligned}dx'&=(dr'^2+dz'^2)^{\frac{1}{2}}\approx[dr^2(1-2GM/c^2r)+dz^2]^{\frac{1}{2}}\\&=dx[1-(2GM/c^2r)(dr/dx)^2]^{\frac{1}{2}}\\&\approx dx[1-(GMx^2/c^2r^3)]; \qquad (8)\end{aligned}$$

we have made use here of the relations $dx^2=(dr^2+dz^2)$ and $(dr/dx)=(x/r)$. Combination of Eqs. (7) and (8) then gives

$$\begin{aligned}c'&=(dx'/dt')\\&=(dx/dt)[1-(GMx^2/c^2r^3)]/[1+(GM/c^2r)]\\&\approx c[1-(GMx^2/c^2r^3)-(GM/c^2r)]. \qquad (9)\end{aligned}$$

Note that as determined by A the speed of light is anisotropic since it depends on the angle between the radial line and the direction of propagation.

Equation (9) may be used to the order given, since deviations of the position or direction of the light pulse from the straight-line path parallel to the x axis produce higher-order changes in c'. Then from Huygens' principle, if we express c' as a function of x and y instead of x and r, the curvature of the light path is given by

$$\begin{aligned}&(1/c')(\partial c'/\partial y)\\&\approx(GM/c^2)[(3x^2y/r^5)+(y/r^3)]. \qquad (10)\end{aligned}$$

Since c' increases with increasing y, the curvature is such that the ray is concave toward the sun. Again to lowest order, y can be replaced by R, the distance of closest approach to the sun, and the total angular deflection is

$$\theta=\int_{-\infty}^{\infty}[(1/c')(\partial c'/\partial y)]_{y=R}dx. \qquad (11)$$

Substitution of (10) into (11) gives

$$\theta=(4GM/c^2R),$$

which is approximately equal to 1.7″ when R is the radius of the sun.

IV. CONCLUDING REMARKS

1. The first term on the right side of Eq. (10) arises from the difference in the lengths of standard measuring rods at A and B, and the second term from the difference in the periods of standard clocks. Since these two terms contribute equally to the deflection of light, Einstein[1] obtained half the correct value when he considered only the difference in clock periods.

2. The precession of the perihelion of the orbit of Mercury cannot be calculated by an extension of the methods of this paper for two reasons. First, we require in addition an equation of motion for a particle of finite rest mass, to replace the argument used above that the speed of light measured by B is c. Second, it turns out that when a suitable equation of motion (the geodesic equation) is introduced, the change in the clock period of order $(GM/c^2r)^2$ cannot be neglected, and this cannot be found by the methods of this paper.

3. The first correct calculation of the deflection of light and the orbit precession was made by Einstein,[3] using a method of successive approximations. He realized that the red shift is determined only by the time change of order (GM/c^2r), that the deflection of light requires in addition the distance change of the same order, and that the orbit precession requires these together with the equation of motion and the time change of order $(GM/c^2r)^2$. These points have also been made with particular force by Eddington.[4]

4. Since the first two of the three "crucial tests" can be derived from the equivalence principle and special relativity without reference to the geodesic equation or the field equations of general relativity, it follows that only the orbit precession really provides a test of general relativity. Moreover, any discrepancy between theory and experiment with regard to the red shift and the deflection of light would not only cast doubt on general relativity, but would have to be reconciled with the experimentally well-established validity of the equivalence principle and special relativity.

5. By the same token, it will be extremely difficult to design a terrestrial or satellite experiment that really tests general relativity, and does not merely supply corroborative evidence for the equivalence principle and special relativity. To accomplish this it will be necessary either to use particles of finite rest mass so that the geodesic equation may be confirmed beyond the Newtonian approximation, or to verify the exceedingly small time or distance changes of order $(GM/c^2r)^2$. For the latter the required accuracy of a clock is somewhat better than one part in 10^{18}.

Note added in proof. Most of the accompanying paper by Professor Dicke [Am. J. Phys. **28**, 344 (1960), this issue] has no bearing on this paper. However, two points he raises are worthy of comment. (1) I did not mean my paper to constitute an indictment of the satellite clock experiments. Terrestrial, balloon, or satellite experiments on the gravitational red shift had not been accomplished when this paper was submitted for publication (although there has been much progress since then), and the astronomical measurements are of limited accuracy. Thus I believe that such experiments are well worthwhile. (2) The Eötvös experiments show with considerable accuracy that the gravitational and inertial masses of normal matter are equal. This means that the ground state eigenvalue of the Hamiltonian for this matter appears equally in the inertial mass and in the interaction of this mass with a gravitational field. It would be quite remarkable if this could occur without the entire Hamiltonian being involved in the same way, in which case a clock composed of atoms whose motions are determined by this Hamiltonian would have its rate affected in the expected manner by a gravitational field. Nevertheless, as stated in the foregoing, I believe that a direct demonstration that the equivalence principle is valid for clocks would be useful. On the other hand, it is evident that experiments of this type could not verify any features of general relativity theory other than the first-order change in the time scale.

[3] A. Einstein, Sitzber. kgl. preuss. Akad. Wiss. 831 (1915).

[4] A. S. Eddington, *The Mathematical Theory of Relativity* (Cambridge University Press, New York, 1924), p. 105.