THE

PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 141, NO. 4 JANUARY 1966

Analysis of the Proposed Planetary Radar Reflection Experiment*

D. K. ROSS† AND L. I. SCHIFF

Institute of Theoretical Physics, Department of Physics, Stanford University, Stanford, California

(Received 24 September 1965)

An analysis is made of an experiment recently proposed by Shapiro as a new test of general relativity theory. This consists in measuring the transit times of radar signals reflected from one of the inner planets. It is found that if the observations are expressed in terms of measurable orbital parameters, as they must be eventually, this experiment has the unexpected property of being sensitive to a nonlinear term in Einstein's theory. Great care must be taken in deciding what is actually measurable, and it is shown that the process of reconciling the results of calculations performed with mathematically different but physically equivalent forms of the metric is of considerable help in making this decision.

I. INTRODUCTION

IT has recently been proposed by Shapiro[1] that a series of measurements of the transit times of radar signals reflected from one of the inner planets, in their dependence on the orbital positions of the earth and the planet, would provide a new test of the general theory of relativity. The geometry of the situation is shown in Fig. 1, in which the sun is at the origin, the earth has instantaneous rectangular coordinates $(-x_e,d)$, and the planet (Venus or Mercury) has coordinates (x_p,d). We shall use units such that the speed of light and the Newtonian gravitational constant are equal to unity. Then if the mass m of the sun (which is of the order of a kilometer in these units) is neglected, the round-trip transit time is equal to $2(x_e+x_p)$. An additional contribution of order m is expected on the basis of general relativity theory, and it is this term that is the subject of the present paper.

We follow Shapiro in ignoring the motions of the earth and the planet during a single transit. These motions are by no means negligible, but may be taken into account in a straightforward way in reducing the observational data. We also follow Shapiro in neglecting the departure of the radar path from the straight line $y=d$. Path curvature is easily seen to produce a contribution to the transit time that is of order m^2, which we neglect since we shall work only to first order in m.

The analysis in this paper has as its primary objective the determination of the extent to which one particular but important aspect of the theory would be tested by the proposed experiment. This aspect has to do with the structure of the space-time metric in the vicinity of the sun. About 1922, Eddington[2] considered what would happen if one were to drop the requirement that the metric tensor be a solution of the field equations of general relativity theory, but retain the requirement that the equations of motion of matter and electromagnetic radiation be determined from the metric in

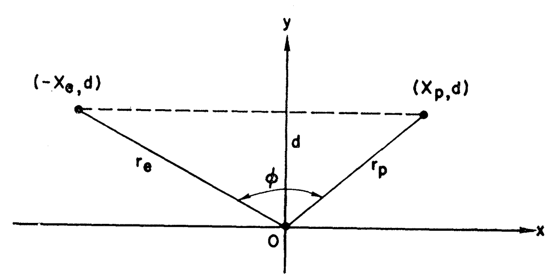

FIG. 1. The unperturbed path with the sun at the origin, the earth at $(-x_e,d)$, and the planet at (x_p,d).

* Supported in part by the U. S. Air Force through Air Force Office of Scientific Research Contract AF 49(638)-1389.
† Woodrow Wilson Fellow, 1964–65.
[1] I. I. Shapiro, Phys. Rev. Letters **13**, 789 (1964); Lincoln Laboratory Report No. 368, 1964 (unpublished). See also D. O. Muhleman and P. Reichley, Jet Propulsion Laboratory Space Programs Summary No. 37-29, 4, 239, 1964 (unpublished).

[2] A. S. Eddington, *The Mathematical Theory of Relativity* (Cambridge University Press, New York, 1957), p. 105. It should be noted that the theory of C. Brans and R. H. Dicke, Phys. Rev. **124**, 925 (1961) can be put in this form.

the same way as in general relativity. This corresponds to modifying Einstein's theory in a special but nevertheless instructive and interesting way, and then determining the effects produced by these modifications on the three tests of the theory. This program of investigating the effects of an altered metric has been carried further by Robertson,[3] and the present paper applies it to the planetary radar reflection experiment.

The most general static spherically symmetric metric may be written in the form

$$ds^2=[1-(2\alpha m/r)+(2\beta m^2/r^2)+\cdots]dt^2 \\ -[1+(2\gamma m/r)+\cdots](dx^2+dy^2+dz^2)\,, \quad (1)$$

where $r^2=x^2+y^2+z^2$, and α, β, γ $\cdots$ are numerical coefficients that are all equal to unity in Einstein's theory. As remarked above, we assume that the equations of motion of matter and electromagnetic radiation are given in terms of this metric tensor by Einstein's equations. In this way we can calculate the dependence of the measured quantities on α, β, γ $\cdots$, which, in turn, measure the time or space distortions that are linear or nonlinear in m and therefore in the gravitational-field strength. While this procedure is to some extent arbitrary, it is well defined and provides a common basis for comparison of various possible experimental tests of general relativity theory.

With the conventional definition of the gravitational constant, we must choose $\alpha=+1$ in order for planetary orbits to agree with Newtonian theory in lowest order. Then since the first-order gravitational red shift depends only on the α term in the metric, no new information concerning the metric is provided by such a measurement. This conclusion is equivalent to the statement that the first-order red shift follows correctly from any theory of gravitation that is consistent with the equivalence principle.[4] With this choice of α, the gravitational deflection of starlight by the sun is proportional to $1+\gamma$, and the advance of the perihelion of a planetary orbit is proportional to $2(1+\gamma)-\beta$.[3] The precession of the spin axis of a spherical gyroscope that is in a free orbit about a nonrotating attracting body of mass m is proportional to $1+2\gamma$.[3,5] Our primary objective, then, is to derive the dependence of the m-proportional term in the expression for the radar transit time on β and γ.

The secondary objective of this paper is to stress the care that must be taken in comparing the predictions of general relativity theory with observations. This might at first seem to be a trivial matter, since a comparison between theory and experiment in any area of physics is essentially a comparison between two sets of numbers that represent the same physical quantity, one set calculated from the theory and the other from the observations. However, it is not always easy, in dealing with general relativity, to know what is measurable and what is an artifact of the choice of coordinate system, which in this theory is completely arbitrary. Difficulties of this type arose in connection with the gyroscope-spin-precession experiment, which was proposed several years ago as a new test of Einstein's theory[6] and which has been undergoing extensive implementation since then.[7,8] A convenient device for avoiding these difficulties consists in exploiting the coordinate invariance of the physically observable quantities by calculating them in different coordinate systems. Quantities so calculated are of course all obtainable from each other by suitable transformations. However, the process of reconciling them without using these transformations forces the investigator to discard coordinate-dependent artifacts, and retain only physically meaningful relations between measurable quantities.

The radar transit time is calculated in the next section by making use of both standard and isotropic forms of the Schwarzschild metric. The different expressions thus obtained are then reconciled in Sec. III by a careful examination of the measurability of the various parameters that appear in them. Finally the entire calculation is done over in Sec. IV with the generalized metric of Eq. (1), and the dependence of the result on β and γ is obtained.[9]

II. CALCULATION OF THE RADAR TRANSIT TIME

The standard form[10] of the Schwarzschild metric may be written in spherical coordinates as

$$ds^2=[1-(2m/r)]dt^2-[1-(2m/r)]^{-1}dr^2 \\ -r^2(d\theta^2+\sin^2\theta d\phi^2)\,, \quad (2)$$

or in rectangular coordinates as

$$ds^2=[1-(2m/r)]dt^2 \\ -\{\delta_{ij}+(2mx_ix_j/r^3)[1-(2m/r)]^{-1}\}dx^idx^j\,, \quad (3)$$

where $r^2=x^2+y^2+z^2$. We shall work entirely in terms of the universal coordinate time that appears in Eqs. (2) and (3). While not directly measurable, it serves, because of the static character of the metric, to relate events that occur at widely different space-time points.

[3] H. P. Robertson, in *Space Age Astronomy*, edited by A. J. Deutsch and W. E. Klemperer (Academic Press Inc., New York, 1962), p. 228.

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

[5] L. I. Schiff, in *Proceedings on Theory of Gravitation*, edited by L. Infeld (Gauthier-Villars, Paris and PWN-Polish Scientific Publishers, Warszawa, 1964), p. 71.

[6] L. I. Schiff, Proc. Nat. Acad. Sci. **46**, 871 (1960).

[7] W. M. Fairbank and C. W. F. Everitt (private communication).

[8] H. W. Knoebel (private communication).

[9] A summary of the results of this paper, together with an application of the same ideas to Einstein's three tests, the gyroscope spin precession, and the second-order gravitational red shift, has been presented by one of us (L.I.S.) at the American Mathematical Society 1965 Summer Seminar on Relativity and Astrophysics. An abbreviated version of these lectures will be publsihed in the Proceedings of the Seminar, and is available as ITP-181, 1965 (unpublished).

[10] P. G. Bergmann, *Introduction to the Theory of Relativity* (Prentice-Hall, Inc., New York, 1946), pp. 203, 212.

The final results must of course be converted to proper time as measured on the earth, since all observations are made there. However, this common transformation can be made in a straightforward way, and need not be dwelt on further; it does not affect the relationships between measured quantities that are derived below.

While we shall make use of the spherical form (2) in the next section, it is more convenient to calculate the transit time from Eq. (3) since the unperturbed path is most simply expressed in rectangular coordinates (see Fig. 1). As remarked in Sec. I, the deviation of this path from the straight line $y=d$ can be neglected, so that $dy=dz=0$ to first order in m. The equation of motion of the electromagnetic signal is $ds=0$, so that Eq. (3) may be written to first order as

$$dt=[1+(m/r)+(mx^2/r^3)]dx, \qquad (4)$$

where now $r^2=x^2+d^2$. Integration of Eq. (4) gives for the round-trip coordinate transit time t_S when earth motion is neglected:

$$t_S=2(x_e+x_p)+4m\ln[(r_p+x_p)/(r_e-x_e)] \\ -2m[(x_e/r_e)+(x_p/r_p)]. \qquad (5)$$

This is equivalent to the result obtained by Shapiro.[1]

The isotropic form[11] of the Schwarzschild metric in spherical coordinates is

$$ds^2=\left[\frac{1-(m/2r)}{1+(m/2r)}\right]^2 dt^2 \\ -[1+(m/2r)]^4(dr^2+r^2d\theta^2+r^2\sin^2\theta d\phi^2);$$

in rectangular coordinates, the last parenthesis is replaced by $(dx^2+dy^2+dz^2)$, where again $r^2=x^2+y^2+z^2$. Equation (4) is now replaced by

$$dt=[1+(2m/r)]dx, \qquad (6)$$

which integrates to

$$t_I=2(x_e+x_p)+4m\ln[(r_p+x_p)/(r_e-x_e)]. \qquad (7)$$

This result, also obtained in Ref. 1, is significantly different from Eq. (5). Moreover the discrepancy persists when both are converted to earth proper times since this conversion is the same, to first order in m, in both forms of the metric.

III. EXPRESSION IN TERMS OF MEASURABLE QUANTITIES

The discrepancy between Eqs. (5) and (7) can be reconciled through a careful examination of the measurability of the parameters that appear in them. Since we are interested in a difference between terms that are of first order in m, it follows that first-order corrections to m, r_e, x_e, r_p, and x_p will not help matters insofar as they are applied to the first-order terms. On the other hand, it is necessary to determine whether or not the zero order term $2(x_e+x_p)$ has the same meaning, to first order in m, in the two expressions.

It is apparent that x_e+x_p cannot be measured directly. Standard measuring rods are of no use, and light signals are equivalent to the radar signals with which we are concerned. A plausible way in which to proceed is first to express x_e+x_p in terms of r_e, r_p, and ϕ (see Fig. 1). Then since r_e and r_p are themselves not directly measurable to sufficient accuracy, they may be expressed in terms of the observed orbital periods, eccentricities, and elapsed times since perihelion for the earth and the planet. The angle ϕ can also be expressed in terms of these quantities, together with the elapsed time since conjunction. In this way x_e+x_p is ultimately given in terms of measured times and orbital eccentricities. As remarked in Sec. II, it is convenient to express these measured times in terms of the universal coordinate time, which to first order in m is the same for both forms of the metric. All of these times, as well as t_S or t_I, are readily convertible to earth proper times.

The first step consists in expressing x_e+x_p in terms of r_e, r_p, and ϕ. Since $r^2=x^2+y^2+z^2$ in both forms of the metric, the ordinary formulas of Euclidean geometry are applicable so far as relations between these coordinate (but not proper) quantities are concerned. Thus we may use the law of cosines for this purpose:

$$x_e+x_p=(r_e^2+r_p^2-2r_er_p\cos\phi)^{1/2}. \qquad (8)$$

Next we must find the relation between r_p, say, and the measurable orbital parameters. This is accomplished by solving the equations of motion of the planet with the standard and itotropic metrics in turn; for this problem the spherical forms of the metric are more convenient than the rectangular forms.

It is sufficient for the present discussion to consider only circular orbits, although finite eccentricity can be taken into account with some additional complication. Then $r=r_p$, $\theta=\pi/2$, and $\dot{\phi}\equiv d\phi/dt$ is constant and equal to $2\pi/T_p$, where T_p is the orbital period expressed in universal coordinate time. Of the three equations of motion with the standard form of the metric,[12] the θ and ϕ equations are then satisfied identically, and the r equation becomes

$$\dot{\phi}^2=m/r_p^3, \qquad (9)$$

which is exactly the same as in Newtonian theory. We thus obtain Kepler's third law for the relation between coordinate radius and period:

$$r_p=(mT_p^2/4\pi^2)^{1/3}. \qquad (10)$$

When the isotropic form of the metric is used, the equation analogous to Eq. (9) is

$$\dot{\phi}^2=(m/r_p^3)[1-(3m/r_p)], \qquad (11)$$

[11] R. Adler, M. Bazin, and M. Schiffer, *Introduction to General Relativity* (McGraw-Hill Book Company, New York, 1965), p. 176.

[12] C. Møller, *The Theory of Relativity* (Oxford University Press, New York, 1952), p. 349.

through terms of second order in m. Thus with sufficient accuracy for our purpose, Eq. (10) is replaced by

$$r_p=(mT_p^2/4\pi^2)^{1/3}-m. \tag{12}$$

Thus the transition from the isotropic to the standard form of the metric is accomplished by replacing r_e by r_e-m and r_p by r_p-m. This has the effect of changing x_e+x_p given by Eq. (8) into

$$x_e+x_p-m[(x_e/r_e)+(x_p/r_p)].$$

It follows that t_I given by Eq. (7) then becomes equal to t_S given by Eq. (5), and the discrepancy has been removed.

We conclude that the round-trip coordinate transit time expected from Einstein's theory is given by Eq. (5) if the orbits are circles, provided that its zero order part, $2(x_e+x_p)$, is computed from ϕ and the orbital periods in accordance with Newtonian theory since the Newtonian equation (10) is valid with the standard form of the metric. In reality the orbits are somewhat elliptical, so that this conclusion may not be strictly correct for the actual situation.

IV. DEPENDENCE ON β AND γ

We are now in a position to determine the extent to which the radar transit time depends on the parameters β and γ that appear in the generalized metric, given by Eq. (1) with $\alpha=1$. The generalized form of Eq. (6) is easily seen to be

$$dt=[1+(1+\gamma)(m/r)]dx,$$

which may be integrated to give for the round-trip coordinate transit time:

$$t_G=2(x_e+x_p)+2(1+\gamma)m\ln[(r_p+x_p)/(r_e-x_e)]. \tag{13}$$

It might be concluded from Eq. (13) that this experiment would only supply information on the value of γ, and this is the natural expectation when an electromagnetic-propagation experiment is under consideration.[13] However, the discussion of the last two sections shows that the first-order part of t_G has no meaning until a prescription is given for specifying the zero-order part. It can shown that Eq. (11) now becomes

$$\phi^2=(m/r_p^3)[1-(\gamma+2\beta)(m/r_p)],$$

so that Eq. (12) becomes

$$r_p=(mT_p^2/4\pi^2)^{1/3}-(m/3)(\gamma+2\beta).$$

with this substitution into the zero-order part of Eq. (13), the expression for t_G becomes

$$2(x_e+x_p)+2(1+\gamma)m\ln[(r_p+x_p)/(r_e-x_e)] \\ -(2m/3)(\gamma+2\beta)[(x_e/r_e)+(x_p/r_p)]. \tag{14}$$

Again, this is valid for circular orbits provided that the zero-order part is computed from ϕ and the orbital periods in accordance with Newtonian theory.

The dependence of the expression (14) on the nonlinear β term in the metric is, as remarked above, unexpected. Indeed, if only first-order electromagnetic propagation were involved, β should not appear. However, the dynamics of the planetary motion also enters the problem, through the calibration of the zero-order distance x_e+x_p in terms of the orbital periods. Furthermore, the relativistic correction to particle motion, even for a circular orbit, is expected to involve γ and β on an equal footing, as is the case with the advance of the perihelion of a planetary orbit.[13] Thus radar ranging may be looked on as a method for exploring details of planetary motion that are not accessible to optical astronomy. Our results, then, endow the proposed radar reflection experiment with the added significance that it and the planetary perihelion advance are the only observations thus far considered that are sensitive to a nonlinear term in Einstein's theory.

Note added in proof. We thank Dr. I. I. Shapiro for pointing out to use that the presence of the last (β-dependent) term in Eq. (14) may be masked by uncertainties in the effective planetary radii. It seems to us that the possibility of such masking adds to the importance of radar reflection experiments with instrumented space probes, since they present relatively well defined targets.

ACKNOWLEDGMENT

One of the authors (L.I.S.) acknowledges with thanks the hospitality of the Institute of Geophysics and Planetary Physics, University of California, San Diego, where part of the work on this paper was carried out.

[13] The physical basis for this expectation has been discussed by L. I. Schiff, J. Soc. Indust. Appl. Math. **10**, 795 (1962).