GENERAL RELATIVITY CONFRONTS EXPERIMENT

Clifford M. Will
McDonnell Center for the Space Sciences
Department of Physics
Washington University
St. Louis, Missouri 63130, U.S.A.

ABSTRACT. We review the status of experimental tests of general relativity. These include tests of the Einstein Equivalence Principle, which requires that gravitation be described by a curved-spacetime, "metric" theory of gravity. General relativity is consistent with all tests to date, including the "classical tests": light deflection using radio interferometers, radar time delay using Viking Mars landers, and the perihelion shift of Mercury; and tests of the strong equivalence principle, such as lunar laser ranging tests of the "Nordtvedt effect", and tests for variations in G. We also review ten years of observations of the Binary Pulsar, in which the first evidence for gravitational radiation has been found.

1. INTRODUCTION

Celestial mechanics is a very old subject, dating back almost three centuries (and further if one is willing to include Ptolemy as a celestial mechanician). Astrometry is an even older subject, going back two millenia, to Hipparchus. By the standards of these two venerable subjects, general relativity is quite young, although it predates most of the participants in this symposium. Yet, paradoxically, the marriage of these subjects, namely relativistic celestial mechanics and relativistic astrometry are easily less than 20 years old, *younger* than most of the participants in this symposium.

To be sure, Einstein's accounting for the perihelion advance of Mercury in 1915 was a first attempt at relativistic celestial mechanics, and the 1919 eclipse expeditions to measure the deflection of starlight were a first attempt at relativistic astrometry. Yet it was not until the middle 1960's that advances in technology provided high-precision tools such as radar and laser tracking systems, radio interferometers, and high-stability atomic clocks, with the ability to determine the spacetime positions of planets, spacecraft and stars with accuracies high enough to be sensitive to a wide range of relativistic effects. Until about 1980, the detection of such relativistic effects was part of a major program to test general relativity. But in the past five

J. Kovalevsky and V. A. Brumberg (eds.), Relativity in Celestial Mechanics and Astrometry, 355–367.

years, as our confidence in Einstein's theory has grown, and as accuracies have continued to improve, the issue of relativistic effects in celestial mechanics and astrometry has taken on practical importance. Establishing accurate ephemerides, determining accurate positions of stars or of locations on the Earth, transferring time accurately from point to point on the Earth, have implications not only for physics, geoscience, and astronomy, but also for navigation and communication. Relativity now must play a key everyday role in these programs. To general relativists, always eager to find practical consequences of their subject, this has been a very welcome development!

The purpose of this paper is not to extol the new "practical relativity", but rather to review the experimental tests and theoretical ideas that have placed general relativity on a firm empirical foundation, and have put it into the position to play this new role.

We begin by describing experiments that test the foundations of gravitation theory and demonstrate that spacetime must be curved (§2). Section 3 describes the parametrized post-Newtonian formalism, which is useful for analysing the weak-field, slow-motion limit of curved spacetime gravitation theories and for treating solar system experiments. The "classical" solar system tests: light deflection, time delay and perihelion shift are reviewed in §4. Section 5 describes tests of the Strong Equivalence Principle and §6 discusses a test for the existence of gravitational radiation that has been provided by the binary pulsar. In §7 we survey the frontiers of experimental gravitation, and in §8 we present concluding remarks. For more details on this subject, see Will (1981 [hereafter referred to as TEGP], 1984).

2. TESTS OF THE EINSTEIN EQUIVALENCE PRINCIPLE

There is a class of experiments that probe the nature of gravity at a very fundamental level. Examples are the Eötvös experiment, the gravitational redshift experiment and the Hughes-Drever experiment. These experiments test what has come to be called the "Einstein Equivalence Principle" (EEP), which states that (i) the trajectories of "test" bodies are independent of their structure and composition (Weak Equivalence Principle or WEP), (ii) in local freely falling frames, the outcomes of non-gravitational experiments are independent of the velocity of the frame (Local Lorentz Invariance, or LLI), (iii) in local freely falling frames, the outcomes of non-gravitational experiments are independent of the location of the frame, in space or in time (Local Position Invariance, or LPI). If EEP is valid, it is then possible to show that the correct classical theory of gravity must be a "metric" theory, one in which all the physical effects of gravitation are produced by spacetime geometry. Although these three sub-principles of EEP may involve very different kinds of experiments, theoretically they are intimately connected, in the sense that if one of them is violated, then one or more of the others is likely to be violated. This bears out a 1960's-conjecture by Schiff that the validity of WEP alone is sufficient for the validity of EEP.

Figure 1 summarizes these principles and their experimental tests. Any of the non-gravitational interactions strong (S), electromagnetic (E) or weak (W) could violate EEP by means of a "non-metric" coupling to gravitation. Experiments that check EEP thereby set upper limits on certain parameters, denoted α, η and δ that measure the magnitude of possible anomalous effects. These include the Eötvös experiment (e.g., Moscow version 1971: equality of acceleration for Al and Pt to parts in 10^{12}), the gravitational redshift experiment (Vessot-Levine Hydrogen-maser rocket experiment 1976: accuracy two parts in 10^4), and the Hughes-Drever experiment (isotropy of inertial mass, verified to one part in 10^{25} [for a recent experiment see Prestage *et al.* 1985]).

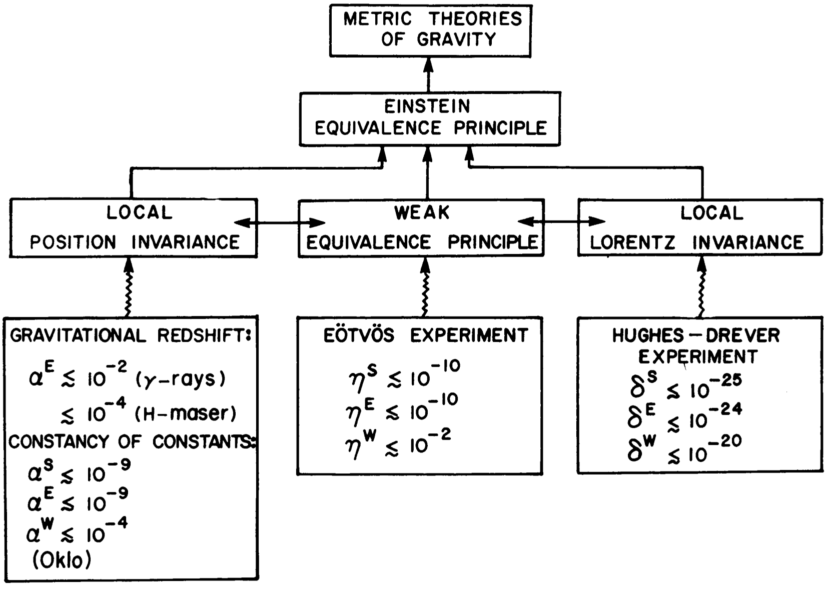

Figure 1. The Einstein Equivalence Principle. The horizontal arrows denote the close connection between the three sub-principles conjectured by Schiff. For detailed discussion see TEGP, §2.

The corresponding limits on the parameters α, η and δ are shown in Figure 1 (for further details, see TEGP, §2). Another possible preferred-location effect that would violate EEP is a cosmological time-dependence of the non-gravitational constants. Current upper limits on such variations came from analyses of fission yields from the Oklo Natural Reactor, a natural, sustained fission reactor believed to have occurred in West Africa around 2×10^9 years ago. The results, shown in Figure 1, represent upper limits on the amount of variation of the constants over one Hubble time (2×10^{10} yrs).

3. METRIC THEORIES OF GRAVITY AND THE PARAMETRIZED POST-NEWTONIAN FORMALISM

The experimental evidence shown in Figure 1 that supports EEP gives confidence that the correct theory of gravity must be a metric theory (TEGP §3). In the weak-field, slow-motion, or "post-Newtonian" limit appropriate to the study of experiments in the solar system, most metric theories of gravity can be analysed in terms of a "theory of theories of gravity" that classifies them in terms of a set of arbitrary dimensionless parameters whose values vary from theory to theory (one exception is the nonsymmetric gravitation theory of Moffat [Will 1984, §2.4]). This theory of theories is known as the Parametrized Post-Newtonian (PPN) formalism (TEGP, §4). One version of the PPN formalism restricts attention to theories of gravity that possess momentum and energy conservation laws, called "semi-conservative" theories. In this version there are five PPN parameters, γ, β, ξ, α_1, and α_2, whose significance and values in general relativity are shown in Table I. In the Brans-Dicke scalar-tensor theory, for example, $\gamma = (1+\omega)/(2+\omega)$, where ω ranges from 0 to ∞, and the other PPN parameters are the same as in general relativity. For a survey of other metric theories of gravity, see TEGP, §5.

TABLE I: THE PPN PARAMETERS AND THEIR SIGNIFICANCE

PPN Parameter	Significance	Value in General Relativity
γ	How much spatial curvature does mass produce?	1
β	How "nonlinear" is gravity?	1
ξ	Are there gravitational preferred-location effects?	0
α_1	Are there gravitational	0
α_2	preferred-frame effects?	0

One can now regard solar-system tests of post-Newtonian effects as measurements of the "correct" values of these parameters. It is convenient to separate solar-system experiments into several classes, including the "Classical Tests", and tests of the "strong Equivalence Principle".

4. THE CLASSICAL TESTS

Three solar-system experiments: light deflection, time delay, and perihelion shift can be called the three "classical" tests (see TEGP §7). This terminology differs from previous usage in which the term "classical tests" referred to gravitational redshift, light deflection,

https://doi.org/10.1017/S0074180900148387 Published online by Cambridge University Press

and perihelion shift. But the gravitational redshift is a test of EEP, *not* of general relativity itself, and so should not be included in this class. Furthermore, the "time delay" test, discovered by Shapiro in 1964, is, in its theoretical interpretation, on an equal footing with the light-deflection, and indeed has yielded the most precise results of the three.

The light-deflection and time-delay tests are related in the sense that they measure effects on the propagation of photons in the curved spacetime around the Sun; in fact, they depend on the PPN parameter γ in the same way. A ray of light that passes near the Sun at a distance d (in units of the solar radius) is deflected by an angle

$$\delta\theta = \frac{1}{2}(1+\gamma)\, 1\overset{''}{.}75/d \quad , \tag{1}$$

and a ray of light that similarly passes near the Sun on a round trip, say from the Earth to Mars, suffers a delay that leads to an excess round trip travel time

$$\delta t = \frac{1}{2}(1+\gamma)\, 250(1-0.16 \ln d)\ \mu s \quad . \tag{2}$$

The most precise light-deflection experiment are those that employ radio interferometry to measure the deflection of radio waves from quasars; since 1969 a series of such measurements have yielded values of the coefficient $\frac{1}{2}(1+\gamma)$ in agreement with general relativity at levels of accuracy approaching 1.5 per cent (see Figure 2). Improvements in the accuracy of radio VLBI and of optical astrometry to the milliarcsecond level and beyond have made it necessary to take the deflection of light into account over the entire celestial sphere (Robertson and Carter 1984). Measurements of the time delay using radar ranging to planets and spacecraft have been carried out since 1967; data from ranging to the Viking orbiters and landers on Mars have yielded results in agreement with general relativity with errors of a part in 1000. (Figure 3).

The third of the classical tests is the perihelion shift, but here the situation is more complicated. In terms of the PPN formalism and Newtonian gravitation, the predicted advance of Mercury's perihelion due to the Sun is, in arcseconds per century,

$$\dot{\omega} = 42\overset{''}{.}98\ \lambda_p \quad ,$$
$$\lambda_p = [\frac{1}{3}(2+2\gamma-\beta) + 3\times10^{-4}(J_2/10^{-7})] \quad , \tag{3}$$

where J_2 is the quadrupole moment of the Sun. From radar ranging to the inner planets, the measured perihelion shift agrees with $42\overset{''}{.}98$ per century yielding $\lambda_p = 1.003 \pm 0.005$. However, measurements of the visual solar oblateness by Dicke and Goldenberg in 1966 were interpreted as corresponding to a value $J_2 = (2.5 \pm 0.2)\times10^{-5}$ which would contribute an anomalous 3" per century to the shift. Later measurements by Hill and colleagues yielded an upper limit $J_2 < 5\times10^{-6}$.

https://doi.org/10.1017/S0074180900148387 Published online by Cambridge University Press

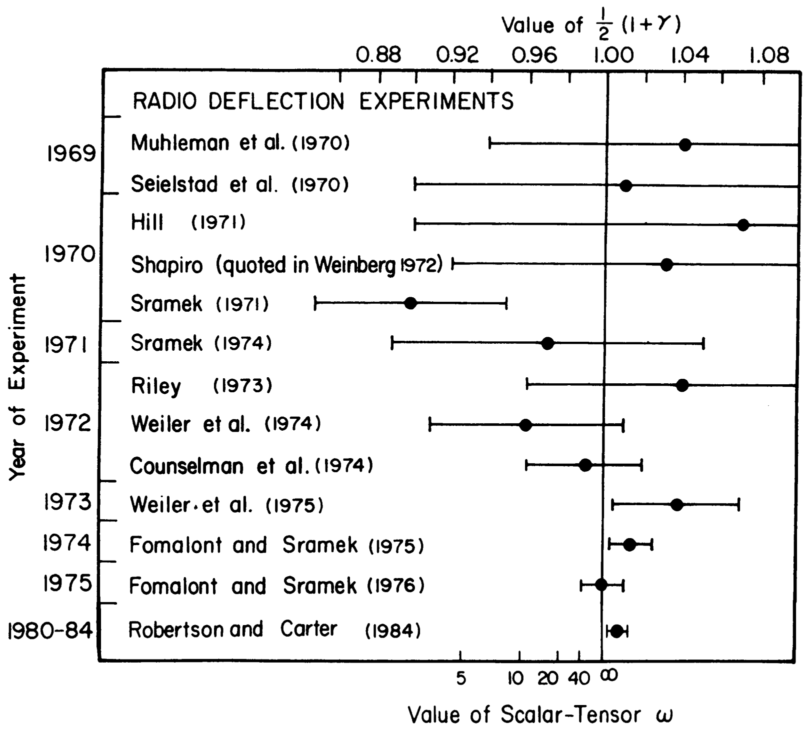

Figure 2. Results of radio-wave deflection measurements 1969-84. The horizontal scale at the top represents values of the PPN parameter combination $(1+\gamma)/2$, whose value in general relativity is unity. The bottom scale gives the value of this combination for the corresponding value of the Brans-Dicke constant ω. In the limit $\omega \to \infty$ Brans-Dicke theory becomes indistinguishable from general relativity.

More recently, values of J_2 ranging between 6×10^{-6} and 10^{-7} have been inferred from differentially rotating solar models constructed to be compatible with the observed multiplet splitting of acoustic normal modes of oscillation of the Sun. Thus there remains some uncertainty in the interpretation of perihelion shift measurements as tests of relativistic gravity.

Figure 4 illustrates the situation: plotted are measured values of J_2 as a function of time. Also shown are the range of values of J_2 expected in a uniformly rotating solar model and the maximum values of J_2 that would be compatible with general relativity within the 1σ and 2σ upper bounds on λ_p.

https://doi.org/10.1017/S0074180900148387 Published online by Cambridge University Press

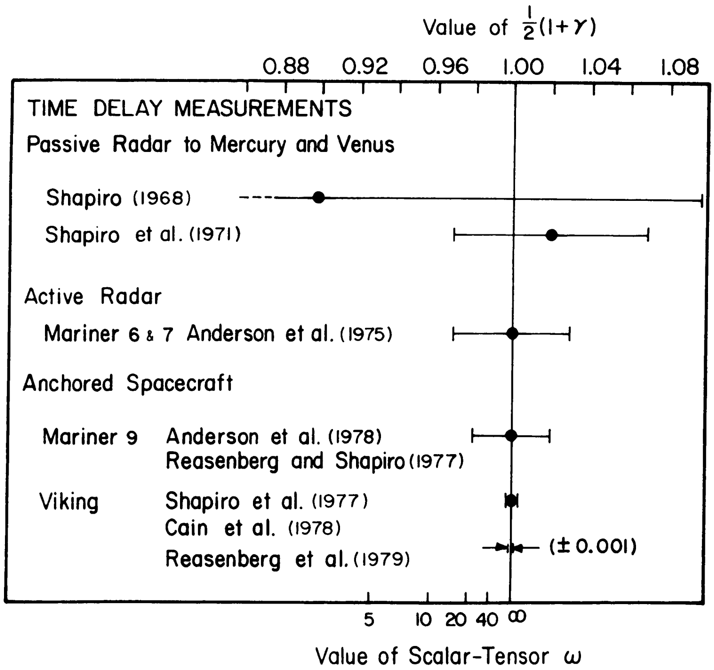

Figure 3. Results of radar time-delay measurements 1968-79.

A direct, unambiguous measurement of J_2 could be provided by a mission that was at one time under study by NASA. Known as Starprobe, it was a spacecraft that would approach the Sun to within four solar radii. Feasibility studies indicated that J_2 could be measured to an accuracy of ten percent of its conventional value of 10^{-7}. Unfortunately, this mission is not a part of NASA's current plans.

5. TESTS OF THE STRONG EQUIVALENCE PRINCIPLE

A number of tests of post-Newtonian gravity can be viewed as tests of the "Strong Equivalence Principle" (SEP). This principle has many of the essential features of EEP, except that it also incorporates the effects of local *gravitational* interactions. For example, for SEP to be valid, "test" bodies, bodies that are small compared to inhomogeneities in external gravitational fields, yet that themselves contain significant self-gravitational binding energy, must fall with the same acceleration (WEP). In addition, local gravitational experiments (such as Cavendish experiments), should show no dependence on the velocity or location of the frame (see Figure 5 and compare Figure 1). Although it is impossible to go beyond this to formulate a precise statement of SEP in parallel with that of EEP, it can be argued, at least heuristically, that SEP implies the presence of *only one* gravitational field, namely the physical metric. Some authors have gone further to argue that this implies general relativity uniquely (TEGP §3.3).

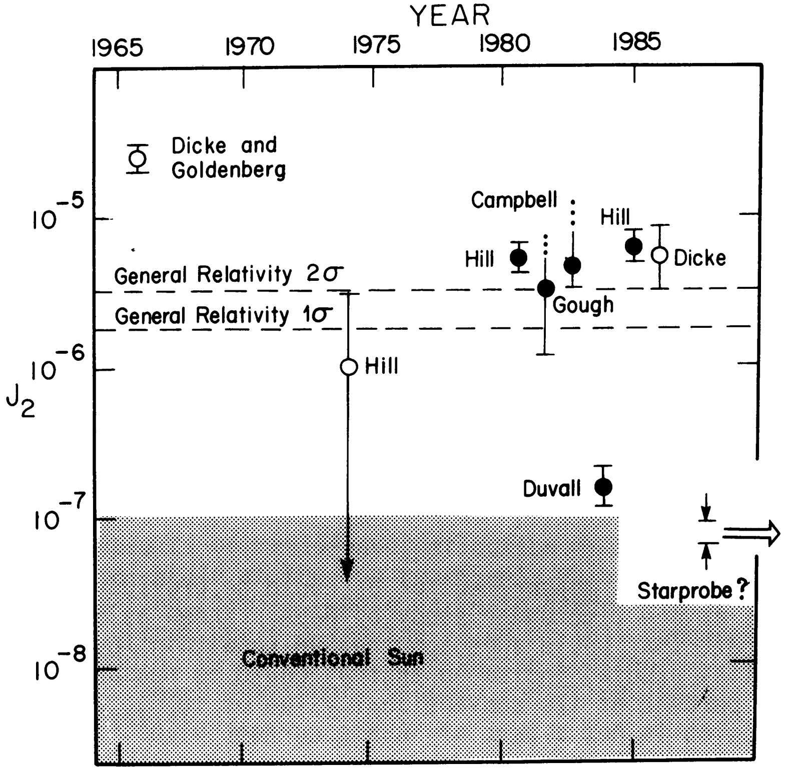

Figure 4. The problem of J_2. Open circles represent visual oblateness measurements, filled circles represent solar oscillation inferences. The "general relativity 1σ (2σ)" lines represent the maximum values of J_2 that would be compatible with general relativity within 1σ (2σ) errors in the radar determinations of Mercury's perihelion shift. The shaded area represents values of J_2 that would be expected from a conventional, uniformly rotating solar model.

The "gravitational WEP" has been verified to a few parts in 10^{12} by the lunar laser ranging experiment (the possibility of a violation of WEP here is called the "Nordtvedt effect"), setting a limit on a combination of PPN parameters shown in Figure 5. Geophysical measurements have set limits on preferred-frame and preferred-location effects in the local gravitational constant, and a variety of measurements have limited cosmic variations in the gravitational constant to a factor around unity in one Hubble time (for further details see TEGP §8).

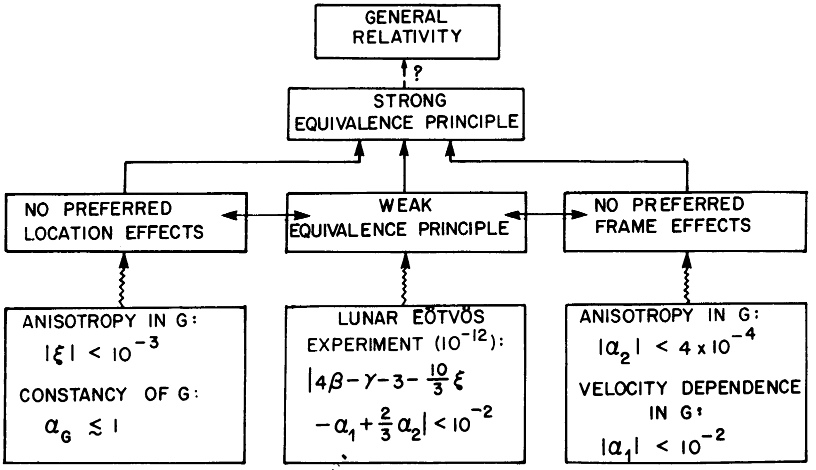

Figure 5. The strong Equivalence Principle. For detailed discussion see TEGP §8.

6. TEST FOR THE EXISTENCE OF GRAVITATIONAL RADIATION

A remarkable new laboratory for studying relativistic gravity known as the Binary Pulsar has provided the first evidence for gravitational waves. Discovered in the summer of 1974 by Hulse and Taylor, it is a pulsar of nominal pulse period 59 ms in a close binary system with an as yet unseen companion. Because the orbit is so close ($\sim$1 $R_\odot$) and because there is no evidence of an eclipse of the pulsar signal or of mass transfer from the companion, it is generally believed that the companion is compact: a white dwarf, a neutron star, or a black hole. Thus the orbital motion is thought to be free of tidal interactions. Furthermore, the data acquisition is "clean" in the sense that the observers can keep track of the pulsar phase with an accuracy of 50 μs, despite gaps of up to six months between observing sessions. The pulsar has shown no evidence of "glitches" in its pulse period. Because of its short orbital period ($\sim$8 hours), and large orbital velocities ($\sim$300 km/s), the effects of relativistic gravity are large, and because of its clean and very stable pulse signal, radio astronomers have been able to measure the parameters of the binary system with extraordinary accuracy and to detect the relativistic effects readily (Table II). Recent results are given by Weisberg and Taylor (1984) (see also TEGP §12). Among the important effects are the periastron shift (4.2263 ± 0.0003 deg/yr) and the effect of the gravitational redshift and special relativistic time dilation of the pulsar's period because of its proximity to and motion about its companion (amplitude of variations in pulse arrival times 4.38 ± 0.12 ms). Making the reasonable assumption that the companion is sufficiently compact to produce no significant tidal interactions which might cause some periastron shift, one can infer unique values for the masses of the two bodies from a comparison of the above measurements with the predictions of general relativity. The results are

$$m_p = 1.42 \pm 0.03 \; m_\odot$$
$$m_c = 1.40 \pm 0.03 \; m_\odot \quad . \tag{4}$$

TABLE II: PARAMETERS OF THE BINARY PULSAR

Parameter	Symbol (units)	Value from Arrival-Time Data 9/1974-8/1983
Pulse period	P_p(s)	0.059029995270̲9̲ ± 20
Derivative of period	$\dot{P}_p$(ss^{-1})	$(8.63 \pm 0.02) \times 10^{-18}$
Projected semi-major axis	$a_1 \sin i$ (light-s)	2.34185 ± 0.00012
Orbital eccentricity	e	0.617127 ± 0.000003
Orbital period	P_b(s)	27906.98163 ± 0.00002
Longitude of periastron (9/74)	ω (deg)	178.8643 ± 0.0009
Periastron advance rate	$\dot{\omega}$ (deg yr^{-1})	4.2263 ± 0.0003
Redshift-doppler parameter	$\mathscr{E}$ (s)	0.00438 ± 0.00012
Sine of inclination angle	sin i	0.76 ± 0.14
Derivative of orbital period	$\dot{P}_b$(ss^{-1})	$(-2.40 \pm 0.09) \times 10^{-12}$

With the masses thus fixed, and the other orbital parameters such as the eccentricity known, the "quadrupole formula" for gravitational radiation reaction in general relativity makes a definite prediction for the rate of orbital period change (essentially, the orbit loses energy to gravitational radiation, so the orbital period must decrease). The predicted rate is

$$(dP/dt)_{GR} = (-2.403 \pm 0.002) \times 10^{-12} \quad . \tag{5}$$

The observed rate is

$$(dP/dt)_{OBS} = (-2.40 \pm 0.09) \times 10^{-12} \tag{6}$$

https://doi.org/10.1017/S0074180900148387 Published online by Cambridge University Press

in complete agreement with the prediction, within the measurement errors. No other plausible source of orbital period decrease has been proposed which could account for all or part of the observed decrease. In addition, the high quality of the recent data has made it possible to detect small, periodic, post-Newtonian effects, including periodic perturbations in the binary orbit, and including the time delay of the pulsar signal as it passes the vicinity of the companion. The predicted size of these effects is uniquely determined by the orbital parameters and the inferred masses and is in agreement within the errors with the data (Weisberg and Taylor 1984). This is an important consistency check on the interpretation of the system as one of two effectively "point" masses, for if there were significant tidal effects on the orbit, the inferred masses would be different, and the predicted post-Newtonian effects would disagree with the observations.

7. FRONTIERS OF EXPERIMENTAL GRAVITATION

Despite the success of general relativity in confronting the experiments described in the previous sections, the subject of experimental gravitation is far from being a closed book. Work continues to improve many of the measurements, for example by continued analysis of Viking radar data to improve both the determinations of PPN parameters and possibly of J_2 (see Hellings 1984), and to improve the limits on a cosmological variation in the gravitational constant (Hellings, *et al.* 1983). Other experiments are underway or are planned that will measure effects that have not been seen before.

One of these is the Stanford Relativity Gyroscope Experiment, under development since 1960 (see Anderson *et al.* 1982 for a review). The goal of the experiment is to measure the precessions of a set of orbiting gyroscopes that result from two effects, the curvature of space around the Earth (net effect $\sim$7" per year) and the dragging of inertial frames by the rotation of the Earth (net effect $\sim 0\overset{''}{.}05$ per year). The gyroscopes are 4 cm diameter quartz spheres coated with a layer of superconducting niobium; at liquid Helium temperatures the sphere develops a magnetic moment parallel to its spin axis whose direction can then be determined by SQUID magnetometers. The precession of the gyroscope axes will be measured relative to the optical axis of a telescope system fixed on a distant star (Rigel). The entire system will be in a drag-compensated satellite. Current plans call for a proving flight to test the components on a 1988 Space Shuttle mission. If all goes as planned, an operational flight could follow in a few years.

The possibility of measuring the second-order, or post-post-Newtonian contributions to solar-system relativistic effects is being studied by several groups. The ideas include a precision optical interferometer in Space (POINTS) with microarcsecond accuracy, to measure the second-order contributions to the deflection of light (Reasenberg 1980), and the use of ultra-stable Hydrogen maser clocks on a Starprobe-type mission to measure the second-order part of the gravitational redshift (Vessot 1984).

There is always interest in finding new arenas for confronting general relativity with observation. One that has been exploited only partially is cosmology. Our knowledge of the primordial abundances of the light elements, primarily Helium and Deuterium has become sufficiently reliable (five to ten percent accuracy), that it may be fruitful to test alternative theories by comparing their nucleosynthesis predictions with the "observations" (Will 1985). General relativity is in good agreement with the primordial values.

8. CONCLUDING REMARKS

During the past 25 years, an intensive theoretical and experimental effort has tested the predictions of general relativity and of other theories of gravitation in many different arenas, and to high precision. General relativity has passed every test, while numerous theories have fallen by the wayside. Although many opportunities remain for further testing of gravitational theory, we can be sufficiently secure about the empirical underpinnings of general relativity to use it as a practical tool in relativistic astrometry and in relativistic celestial mechanics.

This work was supported in part by the National Science Foundation [PHY 83-13545].

REFERENCES

Anderson, J. T., Cabrera, B., Everitt, C. W. F., Leslie, B. C., and Lipa, J. A. 1982, in *Proc. Second Marcel Grossmann Meeting on General Relativity*, ed. R. Ruffini (Amsterdam: North Holland), p. 939.

Hellings, R. W. 1984, in *General Relativity and Gravitation*, eds. B. Bertotti, F. de Felice, and A. Pascolini (Dordrecht: Reidel), p. 365.

Hellings, R. W., Adams, P. J., Anderson, J. D., Keesey, M. S., Lau, E. L., Standish, E. M., Canuto, V. M., and Goldman, I. 1983, *Phys. Rev. Letters* **51**, 1609.

Prestage, J. D., Bollinger, J. J., Itano, W. M., and Wineland, D. J. 1985, *Phys. Rev. Letters*. **54**, 2387.

Reasenberg, R. D. 1984, *Phil. Trans. Roy. Soc. London*, **310**, 227.

Robertson, D. S., and Carter, W. E. 1984, *Nature*, **310**, 572.

Vessot, R. F. C. 1984, *Contemp. Phys.*, **25**, 355.

Weisberg, J. M., and Taylor, J. H. 1984, *Phys. Rev. Letters*, **52**, 1348.

Will, C. M. 1981, *Theory and Experiment in Gravitational Physics* (Cambridge: Cambridge University Press), referred to as TEGP.

----------- 1984, *Phys. Reports*, **113**, 345.

----------- 1985, in *Inner Space/Outer Space: The Interface of Cosmology and Particle Physics*, eds. E. W. Kolb, M. S. Turner, K. Olive, D. Seckel, and D. Lindley (Chicago: University of Chicago).

DISCUSSION

Hill : there exist other models of binary pulsar. Are you sure that it is a purely General Relativity effect ?

Will : it is not a test of General Relativity, but rather an application of General Relativity to astrophysics. However, there is a good possibility here that the model and the observations agree.

Grishchuk : General Relativity appears whenever the curvature of the space is a suitable tool in order to explain the experimental data. However, it may sometimes be more convenient to use a flat space.

Cannon : what are the estimates of the anisotropy of the 3K background radiation and of preferred frame PPN parameters α_i ?

Will : on the anisotropy of the background radiation, it is estimated to be smaller or equal to 10^{-4} on angular scales ranging from 90° (the quadrupole anisotropy) down to arc-minutes. Concerning the parameters α_i of the PPN formalism, the upper bounds range from 10^{-4} to 10^{-10}, showing that there are no preferred-frame effects.

Grishchuk : do you see other possibilities for testing post- post Newtonian effects ?

Will : very few. The only possibilities that seem reasonable are the second order deflection and perhaps a second-order gravitational redshift experiment on a star-probe mission.