New Tests of the Strong Equivalence Principle Using Binary-Pulsar Data

Thibault Damour
Institut des Hautes Etudes Scientifiques, 91440 Bures sur Yvette, France
and Département d'Astrophysique Relativiste et de Cosmologie,
Observatoire de Paris–Centre National de la Recherche Scientifique, 92195 Meudon CEDEX, France

Gerhard Schäfer
Institut für Astrophysik, Max-Planck-Institut für Physik und Astrophysik,
D-8046 Garching bei München, Federal Republic of Germany
(Received 20 February 1991)

One of the few experimental handles on the nonlinear properties of the gravitational interaction is to test the "strong equivalence principle," i.e., to test whether the ratio $m_{\text{gravitational}}/m_{\text{inertial}}$ is 1 for self-gravitating bodies. We point out that existing observational data on the class of small-eccentricity long-orbital-period binary pulsars already provide a limit (namely $|m_g/m_i-1|<1.1\times10^{-2}$; 90% C.L.) which goes beyond corresponding solar-system limits in probing strong-gravitational-field effects. Possible observational ways of improving this limit are suggested.

PACS numbers: 04.80.+z, 95.30.Sf, 97.60.Gb, 97.80.Fk

The equivalence principle, i.e., the property that all neutral test masses fall with the same acceleration in an external gravitational field, is a profoundly characteristic feature of the gravitational interaction. It has been verified, with a fractional precision $\delta a/a \sim 10^{-11}$, by the experiments of Roll, Krotkov, and Dicke,[1] and of Braginsky and Panov,[2] as well as, very recently, by new Earth-based experiments[3] motivated by the possible existence of a supplementary finite-range vector or scalar macroscopic interaction.[4] Moreover, a planned satellite experiment[5] aims at improving the precision of the test down to a level $\delta a/a \sim 10^{-17}$. However, it was pointed out by Nordtvedt[6] that the laboratory-size bodies used in such experiments possess a negligible fraction of gravitational self-energy, and therefore that such experiments indicate nothing about the equality of "gravitational," m_g, and "inertial," m_i, masses when including terms of fractional order E_g/mc^2 (where E_g denotes the gravitational self-energy). Nordtvedt further pointed out[7] the possibility to test such a stronger version of the equivalence principle through the analysis of lunar-laser-ranging data. This test has been performed[8-10] and reaches now a precision $|\delta a/a| = \delta(m_g/m_i)| \sim 2\times10^{-12}$, sufficient to put severe limits (currently $|\eta|<0.01$ at the 2σ level[10]) on the parameter quantifying a violation of the strong equivalence principle of the type envisaged by Nordtvedt: $m_g/m_i = 1+\eta E_g/mc^2$.

The purpose of this Letter is to point out that, in view of the smallness of the self-gravity of planetary bodies (e.g., $E_g/mc^2 = -4.6\times10^{-10}$ for the Earth), such solar-system tests of the strong equivalence principle indicate nothing about higher-order gravitational-energy contributions to the ratio m_g/m_i for some body a:

$$(m_g/m_i)_a \equiv 1+\Delta_a = 1+\eta(E_g/mc^2)_a + \eta'[(E_g/mc^2)_a]^2 + \cdots . \quad (1)$$

To test such higher-order effects one needs to consider strongly self-gravitating bodies, such as neutron stars [for which $-E_g/mc^2 \sim 0.15$, so that the higher-order contributions to Δ_a, in Eq. (1), can reach a few percent]. We shall show in this Letter that presently existing binary-pulsar data already contain important information that puts limits on such higher-order violations of the strong equivalence principle. The possibility that such higher-order contributions in Eq. (1) are present (η' of order unity, etc.) *independently* of the magnitude of the lowest-order term ($|\eta|<0.01$) has been recently proven by the investigation of a general class of alternative relativistic field theories of gravity,[11] which can coincide with general relativity in the post-Newtonian limit, and differ arbitrarily from it in the strong-field regime.

In the presence of a violation of the strong equivalence principle, Eq. (1), the equations of motion for the relative position[12] $\mathbf{r}=\mathbf{x}_1-\mathbf{x}_2$ between, say, the pulsar (inertial mass m_1) and its companion (inertial mass m_2) have the form

$$d^2\mathbf{r}/dt^2 + \mathcal{G}M\mathbf{r}/r^3 = \mathbf{R}+\mathbf{F}\,, \quad (2)$$

where $M\equiv m_1+m_2$, $\mathcal{G}$ denotes the effective gravitational constant for the interaction between m_1 and m_2, $\mathbf{R}$ denotes the orbital relativistic contributions $[\propto (v^{\text{orbital}}/c)^2]$, and $\mathbf{F}\equiv\Delta\mathbf{g}$, with $\Delta\equiv\Delta_1-\Delta_2$ denoting the supplementary "force" (per unit mass) due to the differential acceleration of free fall in the gravitational field $\mathbf{g}$ of the galaxy. Equation (2) has the same form as the usual "lunar" Nordtvedt effect, with the physical differences that, in our "pulsar" case, the eccentricity of the unperturbed Keplerian orbit is, in general, not small, the force $\mathbf{F}$ is not nearly parallel to the orbital plane and is essentially constant in magnitude and direction (so that we are considering the gravitational analog of the Stark effect), and the basic observable quantity is not the "range" $|\mathbf{r}|$ but the component of $\mathbf{r}$ along the line of sight. The solution of Eq. (2) comprises both short-period and secular effects. The short-period (i.e., with period of the order of the orbital period P_b) effects induced by $\mathbf{F}$ can be shown to contribute to the timing formula of a binary pulsar terms of order

$$F_{\perp}P_b^2/4\pi^2c = 1.75\Delta(g_{\perp}/g_0)[P_b/(10^6\,\mathrm{s})]^2\times 10^{-8}\,\mathrm{s}\,,$$

where $\perp$ denotes the projection onto the orbital plane, and g_0 the value of the galactic acceleration at the solar circle. With such terms seeming unmeasurably small, we conclude that it is sufficient to study the secular effects induced by $\mathbf{F}$.

Averaging over one orbital period the time derivatives of the energy, the angular momentum, and the Lagrange-Laplace (-Runge-Lenz) vector, one finds the following equations for the secular evolution of the Keplerian elements of a binary system:

$$\langle da/dt\rangle = 0,\quad \langle d\mathbf{e}/dt\rangle = \mathbf{f}\times\boldsymbol{l} + \omega_R\mathbf{c}\times\mathbf{e},\quad \langle d\boldsymbol{l}/dt\rangle = \mathbf{f}\times\mathbf{e}\,. \tag{3}$$

In Eq. (3) $\mathbf{e}\equiv e\mathbf{a}$ (eccentricity vector), $\boldsymbol{l}\equiv(1-e^2)^{1/2}\mathbf{c}$, and $\mathbf{f}\equiv\frac{3}{2}\mathbf{F}/na$, $(\mathbf{a},\mathbf{b},\mathbf{c})$ being an orthonormal triad with $\mathbf{a}$ along the apsidal line (towards the periastron), and $\mathbf{c}$ along the orbital angular momentum, while, as usual, a is the relative semimajor axis, e the eccentricity, and $n\equiv 2\pi/P_b = (\mathcal{G}M/a^3)^{1/2}$. Finally, ω_R denotes the average angular velocity of the *relativistic* advance of the periastron, which, in a general alternative relativistic theory of gravity reads $\omega_R = nk$, with $k = \mathcal{F}\times 3\mathcal{G}M/c^2a(1-e^2)$, the factor $\mathcal{F}$ being unity in Einstein's theory and a function of m_1 and m_2 in alternative theories. Equations (3) show that both the shape and the spatial orientation of the Keplerian binary ellipse slowly change under the influence of $\mathbf{F}$. In high-eccentricity binary pulsars such as PSR 1913+16 this leads to new measurable effects, notably a secular change of $x\equiv(m_2/M)a\sin i/c$ (through a change of the inclination angle i), and a secular change of the eccentricity. However, the time scale for these changes is $\sim|\mathbf{f}|^{-1} = (0.1/\Delta)\times 6.7\times 10^8$ yr, i.e., longer (as soon as $\Delta < 0.18$) than the gravitational-wave damping time scale in this system. As it has not yet been possible to measure the gravitational damping effects on x and e in PSR 1913+16,[13] we must turn our attention to other kinds of binary pulsar systems.

There exists, besides the high-mass, high-eccentricity, small-orbital-period class of binary pulsars (of which PSR 1913+16 is the paradigm), a class of low-companion-mass, small-eccentricity, long-orbital-period binary pulsars which turns out, surprisingly, to provide a useful testing ground for the violation of the equivalence principle considered here. Taking advantage of the fact that these binary pulsars have a very small eccentricity, we see that the evolution equations (3) essentially decouple: The last equation implies that the orbital plane is fixed, $\langle d\mathbf{c}/dt\rangle = O(ef)\simeq\mathbf{0}$, and then the penultimate equation gives for $\mathbf{e}(t)$ a linear (vectorial) evolution equation with constant coefficients and a constant forcing term. The general solution of this evolution equation can then be written as the following vectorial superposition:

$$\mathbf{e}(t) = \mathbf{e}_F + \mathbf{e}_R(t),\quad \mathbf{e}_F \equiv \mathbf{f}_{\perp}/\omega_R = \tfrac{3}{2}\Delta\mathbf{g}_{\perp}/\omega_R na\,. \tag{4}$$

In Eq. (4) $\mathbf{e}_R(t)$ (general homogeneous solution) represents a usual relativistic periastron advance phenomenon (the eccentricity vector $\mathbf{e}_R$ turning in the orbital plane with angular velocity ω_R), and $\mathbf{e}_F$ (inhomogeneous solution) represents a constant eccentricity vector directed along the projection of the external force onto the orbital plane (i.e., a constant, $\mathbf{F}$-induced, "polarization" of the orbit). To make the link with the usual (lunar) Nordtvedt effect, one can consider the evolution equation for the eccentricity, $e\equiv|\mathbf{e}|$, as obtained from Eqs. (3): $de/dt = (1-e^2)^{1/2}\mathbf{f}\cdot\mathbf{b}$. The important factor in this equation is the presence of a slowly changing angle, say $\theta(t)$, between $\mathbf{f}$ and $\mathbf{b}$. In the lunar case the time dependence of θ is due to the rotation of the external force $\mathbf{f}$, while in our case it is due to the relativistic rotation of the orbit-based direction $\mathbf{b}$. This argument shows directly that the limit of validity of our result, $\mathbf{e}_F = \mathbf{f}_{\perp}/\omega_R$, is that ω_R should be appreciably faster than ω_0, the angular velocity of rotation of the galaxy with which $\mathbf{F}$ rotates. As one finds $\omega_0/\omega_R = [P_b/(1364\ \mathrm{d})]^{5/3}(M/1.7M_{\odot})^{-2/3}$, this condition will be well satisfied by the two binary pulsars that we shall consider below.

Equation (4) tells us that the observable eccentricity vector $\mathbf{e}(t)$ lies on a circle of radius $|\mathbf{e}_R|$, centered at $\mathbf{e}_F$. If $|\mathbf{e}_F|\gg|\mathbf{e}_R|$, the observed $e = |\mathbf{e}_F+\mathbf{e}_R|\simeq|\mathbf{e}_F|$ is a direct estimate of $|\mathbf{e}_F|$, while if $|\mathbf{e}_F|\ll|\mathbf{e}_R|$, the observed $e = |\mathbf{e}_F+\mathbf{e}_R|\simeq|\mathbf{e}_R|\gg|\mathbf{e}_F|$ gives an upper bound to $|\mathbf{e}_F|$. In both cases the observation of a binary pulsar system having a very small eccentricity directly yields an upper limit to the Δ-induced polarization $\mathbf{e}_F$. However, this is no longer the case if $|\mathbf{e}_F|\simeq|\mathbf{e}_R|$, where a vectorial compensation between $\mathbf{e}_F$ and $\mathbf{e}_R(t)$ could happen at the time of observation. To get secure limits on Δ let us henceforth put ourselves in the worst case, where there could occur, sometimes, an exact cancellation between $\mathbf{e}_F$ and $\mathbf{e}_R(t)$, i.e., let us assume $|\mathbf{e}_F| = |\mathbf{e}_R|$.[14] We are lucky that there exist two small-eccentricity, long-orbital-period, binary pulsar systems, namely, PSR 1855+09 and PSR 1953+29, which are known to be so old[15] that $\mathbf{e}_R$ has had the time to make many turns ($\omega_R t > 557\pi$ for $t > 10^9$ yr and the least relativistic system 1953+29). We are then entitled to reason probabilistically, by considering that the position of $\mathbf{e}(t)$ on the circle which it describes is a random variable uniformly distributed on the circle. Let us denote by θ $[=\omega_R(t-t_0)]$ the angle on the eccentricity circle, by Ω the longitude of the ascending node, by i the inclination of the orbital plane, and by λ the angle between the Sun and the galactic center as seen from the pulsar, and let us define $\hat{e} \equiv (P_b/2\pi)^2 gc^2/2\mathcal{J}\mathcal{G}M$. With this notation, the absolute value of $\Delta = \Delta_1 - \Delta_2$ can be expressed as

$$|\Delta| = [f_{i,\lambda}(\theta,\Omega)]^{-1} e/\hat{e}\,, \tag{5}$$

where $f_{i,\lambda}(\theta,\Omega) \equiv 2\sin(\theta/2) R_{i,\lambda}(\Omega)$, with

$$R_{i,\lambda}(\Omega) \equiv [1 - (\cos i \cos\lambda + \sin i \sin\lambda \sin\Omega)^2]^{1/2}.$$

Apart from θ and Ω which are not observable (and that we shall treat as independent random variables, uniformly distributed between 0 and 2π), the other quantities entering Eq. (5) are, in principle, observable in the timing of binary pulsars. For instance, to get λ and g (galactic acceleration at the location of the pulsar) in terms of observable quantities, we use the results of Ref. 12 expressing λ and g as functions of the galactic longitude l and of $\delta \equiv$(Earth-pulsar distance)/(galactic radius). In view of our probabilistic assumptions we can define confidence levels for $|\Delta| \propto [f(\theta,\Omega)]^{-1}$ by considering the probability measure of the regions of the θ-Ω torus where $f(\theta,\Omega)$ is smaller than some value $f_{\text{C.L.}}$, depending on the chosen confidence level. We shall take 90%-confidence-level regions.

Finally, we find that the observation of an (old) binary pulsar system having a (small) observed eccentricity e allows one to put an upper limit on $|\Delta|$ given by

$$|\Delta| < (10/\pi) I_{i,\lambda} e/\hat{e} \quad (90\%\ \text{C.L.})\,, \tag{6}$$

where $I_{i,\lambda} = (2\pi)^{-1}\int_0^{2\pi} d\Omega / R_{i,\lambda}(\Omega)$ is a complete elliptic integral of the first kind [arising when approximating $2\sin(\theta/2)$ by θ or $2\pi - \theta$].

A survey of existing small-eccentricity, long-orbital-period, binary pulsars (using P_b^2/e as figure of merit) shows that the three best systems for constraining $|\Delta|$ are PSR 1855+09, PSR 1953+29, and PSR 0820+02. We must, however, discard the latter system because its age is unknown apart from the fact that it is much younger than the other two, as shown by the observation of a hot white dwarf companion.[16] Concerning PSR 1855+09 the beautiful recent results of Ryba and Taylor[17] give us all the quantities we need to estimate the right-hand side of Eq. (6): $e = 2.167\times10^{-5}$, $P_b = 1.0650676\times10^6$ s, $l = 42.3°$, $\delta = (1.1\ \text{kpc})/(7.7\ \text{kpc}) = 0.143$, $i = 88.28°$, $m_2 = 0.233 M_\odot$, and $M = m_1 + m_2 = 1.50 M_\odot$. Note that, as we are using these data to estimate a small upper limit to Δ, we are entitled to neglect the violations of the strong equivalence principle appearing in the deduction of m_2 and M from the timing data, i.e., we use general relativity to deduce the masses (we also approximate $\mathcal{J}\mathcal{G} \simeq G_{\text{Newton}}$ in $\hat{e}$). Another convenient feature of the class of binary pulsars we consider is that $\Delta_2 \ll \Delta_1$ (white dwarf compared to neutron star) so that we get a direct limit on $\Delta_1 = \Delta(m_1) = (m_g/m_i - 1)_{m_1}$. The limit we finally get is $|\Delta(m_1 = 1.27)| < 5.6\times10^{-2}$ (90% C.L., PSR 1855+09 data).

Concerning PSR 1953+29 the observational results[18] are less rich because neither the masses nor the inclination is measured. However, the formation of such systems is sufficiently well understood to allow one to deduce m_2 from the observed orbital period.[19,20] This yields $m_2 = 0.31 \pm 0.04 M_\odot$,[19] or, consistently, $m_2 \simeq 0.30$.[20] (It is important to note that the corresponding estimates for PSR 1855+09 have been independently confirmed by the observations of Ryba and Taylor.[17]) We shall adopt $m_2 = 0.3 M_\odot$ and (based on the other observations of pulsar masses and on the fact that only a small fraction of a solar mass is required to spin up a millisecond pulsar[21]) $m_1 = 1.4 M_\odot$, so that $M = 1.7 M_\odot$. Then the inclination is determined from the observed mass function. As concerns the galactic-reduced distance, $\delta = d/R_0$, we can estimate it from the results of Damour and Taylor,[12] who found a flat recalibration factor of 0.55, for galactic longitudes $47° < l < 70°$, between the distance estimates of Refs. 22 and 23. Using d(LMT 85) = 2.7 kpc for PSR 1953+29,[18] and remembering that Ref. 23 uses $R_0 = 8.5$ kpc, we get $\delta = 2.7/(0.55\times8.5)$. Finally, the quantities we need in the case of PSR 1953+29 are $e = 3.304\times10^{-4}$, $P_b = 1.013896\times10^7$ s, $l = 65.84°$, $\delta = 0.58$, $i = 41.5°$, $m_2 = 0.3 M_\odot$, and $M = 1.7 M_\odot$. They yield $|\Delta(m_1 = 1.4)| < 1.1\times10^{-2}$ (90% C.L., PSR 1953+29 data).

As we see from Eq. (1), this limit, with $|\eta| < 0.01$ and $|E_g/mc^2| \sim 0.15$, begins to put a (modest) constraint on the possible presence of higher-order self-gravity terms. However, a precise discussion of this constraint must take into account, within a specified class of alternative theories, the exact physical structure of the higher-order terms, and will be left to a separate study.[11]

The final question one might address is whether there are ways to tighten the above limits. One evident way is to hope for the discovery of a new binary pulsar system with a bigger value for the figure of merit: $\inf[P_b^2, (1364\ \text{d})^2]/e$. Another way would be to get direct observational limits on the secular variation of the eccentricity *vector*.[24] Indeed, we have $d\mathbf{e}/dt = \omega_R \mathbf{c}\times\mathbf{e}_R$, so that a measurement of (respectively, upper limit on) $d\mathbf{e}/dt$ gives a direct measurement of (respectively, upper limit on) $\mathbf{e}_R$. If in this way one can get the information that e_R/e is smaller than about $(10/\pi) I_{i,\lambda}$ (=3.87 for PSR 1855+09) then we can render more secure (by making independent from probabilistic considerations), and maybe tighten, the limits obtained above.

G. S. acknowledges support by the Deutsche Forschungsgemeinschaft (Heisenberg Foundation).

[1] P. G. Roll, R. Krotkov, and R. H. Dicke, Ann. Phys. (N.Y.) **26**, 442 (1964).

[2] V. B. Braginsky and V. I. Panov, Zh. Eksp. Teor. Fiz. **61**, 873 (1972) [Sov. Phys. JETP **34**, 463 (1972)].

[3] See E. G. Adelberger *et al.*, Phys. Rev. D **42**, 3267 (1990),

for the most precise equivalence-principle results obtained during the "fifth force" searches, and references therein to the other results.

[4]E. Fishbach *et al.*, Phys. Rev. Lett. **56**, 3 (1986).

[5]F. Barlier *et al.*, Satellite Test of the Equivalence Principle, ESA-NASA Assessment Study Report, ESA Document No. SCI(91)4, January, 1991.

[6]K. Nordtvedt, Phys. Rev. **169**, 1014 (1968).

[7]K. Nordtvedt, Phys. Rev. **170**, 1186 (1968).

[8]J. G. Williams *et al.*, Phys. Rev. Lett. **36**, 551 (1976).

[9]I. I. Shapiro *et al.*, Phys. Rev. Lett. **36**, 555 (1976).

[10]I. I. Shapiro *et al.*, reported in I. I. Shapiro, in *General Relativity and Gravitation, 1989,* edited by N. Ashby, D. F. Bartlett, and W. Wyss (Cambridge Univ. Press, Cambridge, 1990), pp. 313–330.

[11]T. Damour and G. Esposito-Farèse (to be published).

[12]T. Damour and J. H. Taylor, Astrophys. J. **366**, 501 (1991), have recently pointed out that the galactic acceleration of the center of mass of the binary pulsar PSR 1913+16 contributes measurably to the observed $\dot{P}_b$. This fact gives us a further handle on possibly nonzero Δ_a's using the center-of-mass motion instead of the relative one considered in the text. However, a meaningful discussion of the corresponding constraint must take into account the many other new contributions to the "theoretical" $\dot{P}_b$ entailed by the use of a non-Einsteinian theory of gravity, and will be left to a separate study.

[13]J. H. Taylor and J. M. Weisberg, Astrophys. J. **345**, 434 (1989).

[14]This case is also the most plausible one from the astrophysical point of view. Indeed, low-companion-mass binary pulsars are currently viewed as having evolved from the detachment of tidally circularized [$e(t_{\rm detach})\simeq 0$] binary systems where a neutron star accretes from a low-mass subgiant companion.

[15]The (negative) observational results of P. J. Callanan *et al.*, Mon. Not. Roy. Astron. Soc. **238**, 25 (1989), imply that the age of the PSR 1855+09 system is in excess of $\sim 1.0\times 10^9$ yr. The "distance" of PSR 1953+29 away from the spin-up line indicates an age for this system of $\sim 2.9\times 10^9$ yr [see, e.g., E. P. J. van den Heuvel, in *The Origin and Evolution of Neutron Stars,* edited by D. J. Helfand and J. H. Huang (Reidel, Dordrecht, 1987), p. 393].

[16]S. R. Kulkarni, Astrophys. J. **306**, L85 (1986).

[17]M. F. Ryba and J. H. Taylor, Astrophys. J. (to be published).

[18]L. A. Rawley, J. H. Taylor, and M. M. Davis, Nature (London) **319**, 383 (1986).

[19]P. C. Joss, S. Rappaport, and W. Lewis, Astrophys. J. **319**, 180 (1987).

[20]G. J. Savonije, Nature (London) **325**, 416 (1987).

[21]M. A. Alpar *et al.*, Nature (London) **300**, 728 (1982).

[22]A. G. Lyne, R. N. Manchester, and J. H. Taylor, Mon. Not. Roy. Astron. Soc. **213**, 613 (1985).

[23]D. A. Frail, Ph.D. thesis, University of Toronto, 1989 (unpublished); see also D. A. Frail and J. M. Weisberg, Astron. J. **100**, 743 (1990).

[24]If one performs the small-eccentricity expansion of the timing formula of T. Damour and N. Deruelle [Ann. Inst. Henri Poincaré **44**, 263 (1986)] using as basic angular variable $l'=(1+k)(nt+\sigma)+\omega_0$, the coefficients of the second harmonics ($\cos 2l'$ and $\sin 2l'$) are just (half) the two components of the eccentricity vector in the orbital plane. One can then replace $\mathbf{e}\rightarrow \mathbf{e}_0+\dot{\mathbf{e}}_0(t-t_0)$ and fit for the two independent components of $\dot{\mathbf{e}}_0$.