Tests of General Relativity from Timing the Double Pulsar

M. Kramer,[1]* I. H. Stairs,[2] R. N. Manchester,[3] M. A. McLaughlin,[1,4] A. G. Lyne,[1] R. D. Ferdman,[2] M. Burgay,[5] D. R. Lorimer,[1,4] A. Possenti,[5] N. D'Amico,[5,6] J. M. Sarkissian,[3] G. B. Hobbs,[3] J. E. Reynolds,[3] P. C. C. Freire,[7] F. Camilo[8]

The double pulsar system PSR J0737-3039A/B is unique in that both neutron stars are detectable as radio pulsars. They are also known to have much higher mean orbital velocities and accelerations than those of other binary pulsars. The system is therefore a good candidate for testing Einstein's theory of general relativity and alternative theories of gravity in the strong-field regime. We report on precision timing observations taken over the 2.5 years since its discovery and present four independent strong-field tests of general relativity. These tests use the theory-independent mass ratio of the two stars. By measuring relativistic corrections to the Keplerian description of the orbital motion, we find that the "post-Keplerian" parameter *s* agrees with the value predicted by general relativity within an uncertainty of 0.05%, the most precise test yet obtained. We also show that the transverse velocity of the system's center of mass is extremely small. Combined with the system's location near the Sun, this result suggests that future tests of gravitational theories with the double pulsar will supersede the best current solar system tests. It also implies that the second-born pulsar may not have formed through the core collapse of a helium star, as is usually assumed.

Einstein's general theory of relativity (GR) has so far passed all experimental tests with flying colors (*1*), with the most precise tests achieved in the weak-field gravity conditions of the solar system (*2*, *3*). However, it is conceivable that GR breaks down under extreme conditions such as strong gravitational fields where other theories of gravity may apply (*4*). Predictions of gravitational radiation and self-gravitational effects can only be tested using massive and compact astronomical objects such as neutron stars and black holes. Studies of the double–neutron star binary systems PSR B1913+16 and PSR B1534+12 have provided the best such tests so far, confirming GR at the 0.2% and 0.7% level, respectively (*5*–*7*). The recently discovered double pulsar system PSR J0737-3039A/B has much higher mean orbital velocities and accelerations than either PSR B1913+16 or PSR B1534+12 and is unique in that both neutron stars are detectable as radio pulsars (*8*, *9*).

[1]University of Manchester, Jodrell Bank Observatory, Macclesfield SK11 9DL, UK. [2]Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, British Columbia V6T 1Z1, Canada. [3]Australia Telescope National Facility, Commonwealth Scientific and Industrial Research Organisation (CSIRO), P.O. Box 76, Epping, New South Wales 1710, Australia. [4]Department of Physics, West Virginia University, Morgantown, WV 26505, USA. [5]INAF–Osservatorio Astronomica di Cagliari, Loc. Poggio dei Pini, Strada 54, 09012 Capoterra, Italy. [6]Dipartimento di Fisica, Università degli Studi di Cagliari, SP Monserrato-Sestu km 0.7, 09042 Monserrato (CA), Italy. [7]NAIC, Arecibo Observatory, HC03 Box 53995, PR 00612, USA. [8]Columbia Astrophysics Laboratory, Columbia University, 550 West 120th Street, New York, NY 10027, USA.

*To whom correspondence should be addressed. E-mail: mkramer@jb.man.ac.uk

PSR J0737-3039A/B consists of a pulsar with a period of 22 ms, PSR J0737-3039A (henceforth called A), in a 2.4-hour orbit with a younger pulsar with a period of 2.7 s, PSR J0737-3039B (B). Soon after the discovery of A (*8*), it was recognized that the orbit's orientation, measured as the longitude of periastron ω, was changing in time with a very large rate of $\dot{\omega} = d\omega/dt \sim 17^\circ\ \mathrm{year}^{-1}$, which exceeds by a factor of 4 the corresponding value for the Hulse-Taylor binary PSR B1913+16 (*5*). This immediately suggested that the system consists of two neutron stars, a conclusion confirmed by the discovery of pulsations from B (*9*). The pulsed radio emission from B has a strong orbital modulation, both in intensity and in pulse shape. It appears as a strong radio source only for two intervals, each of about 10-min duration, whereas its pulsed emission is rather weak or even undetectable for most of the remainder of the orbit (*9*, *10*).

In double–neutron star systems, especially those having short orbital periods, observed pulse arrival times are modified by relativistic effects that can be modeled in a theory-independent way by means of the so-called "post-Keplerian" (PK) parameters (*11*). These PK parameters are phenomenological corrections and additions to the simple Keplerian description of the binary motion, describing for instance a temporal change in period or orientation of the orbit, or an additional "Shapiro delay" that occurs as a result of the curvature of space-time when pulses pass near the massive companion. The PK parameters take different forms in different theories of gravity, and so their measurement can be used to test these theories (*1*, *12*). For point masses with negligible spin contributions, GR predicts values for the PK parameters that depend only on the two a priori unknown neutron star masses and the precisely measurable Keplerian parameters. Therefore, measurement of three (or more) PK parameters provides one (or more) test of the predictive power of GR. For the double pulsar we can also measure the mass ratio of the two stars, $R \equiv m_A/m_B = x_B/x_A$, where x_A and x_B are the projected semimajor axes of the orbits of A and B. The ability to measure this quantity provides an important constraint because in GR and other theories this simple relationship between the masses and semimajor axes is valid to at least first post-Newtonian (1PN) or $(v/c)^2$ order, where v is the orbital speed of the pulsars and *c* is the speed of light (*12*, *13*).

Observations. Timing observations of PSR J0737-3039A/B were undertaken using the 64-m Parkes radio telescope in New South Wales, Australia; the 76-m Lovell radio telescope at Jodrell Bank Observatory (JBO), UK; and the 100-m Green Bank Telescope (GBT) in West Virginia, USA, between April 2003 and January 2006.

At Parkes, observations were carried out in bands centered at 680 MHz, 1374 MHz, and 3030 MHz. Although timing observations were frequent after the discovery of the system, later observations at Parkes were typically conducted every 3 to 4 weeks, usually covering two full orbits per session. Observations at GBT were conducted at monthly intervals, with each session consisting of a 5- to 8-hour track (i.e., two to three orbits of the double pulsar). Typically, the observing frequencies were 820 and 1400 MHz for alternate sessions. Occasionally, we also performed observations at 340 MHz. In addition, we conducted concentrated campaigns of five 8-hour observing sessions, all at 820 MHz, in May and November 2005. Observations at JBO used the 76-m Lovell telescope. Most data were recorded at 1396 MHz; some observing sessions were carried out at the lower frequency of 610 MHz. The timing data obtained at JBO represent the most densely sampled data set, but because of the limited bandwidth, longer integration times per timing point were required. The Parkes data set is the longest one available and hence provides an excellent basis for investigation of secular timing terms.

The time-series data of all systems were folded according to the predicted topocentric pulse period. The adopted integration times were 30 s for pulsar A (180 s for JBO data) and 300 s for pulsar B. For A, these integration times reflected a compromise between the need to produce pulse profiles with adequate signal-to-noise ratio and to obtain sufficient sampling of the orbit to detect and resolve phenomena that depend on orbital phase, such as the Shapiro delay. The integration time for B corresponded to about 108 pulse periods and was a compromise between the need to form a stable pulse profile and to resolve the systematic changes seen as a function of orbital phase.

Timing measurements. For each of the final profiles, pulse times-of-arrival (TOAs) were computed by correlating the observed pulse profiles with synthetic noise-free templates (fig. S1) (*14*). A total of 131,416 pulse TOAs were measured for A; 507 TOAs were obtained for B. For A, the same template was used for all observations in a given frequency band, but different templates were used for widely separated bands. We note that our observations still provide no good evidence for secular evolution of A's profile (*15*) despite the predictions of geodetic precession. The best timing precision was obtained at 820 MHz with the Green Bank Astronomical Signal Processor (GASP) back end [see (*16*) for details of this and other observing systems] on GBT, with typical TOA measurement uncertainties for pulsar A of 18 μs for a 30-s integration.

For B, because of the orbital and secular dependence of its pulse profile (*10*), different templates were also used for different orbital phases and different epochs. A matrix of B templates was constructed, dividing the data set into 3-month intervals in epoch and 5-min intervals in orbital phase. The results for the 29 orbital phase bins were studied, and we noticed that although the profile changed quickly during the two prominent bright phases, the profile shape was simpler and more stable at orbital phases when the pulsar is weak. This apparent stability at some orbital phases cannot be attributed to a low signal-to-noise ratio, as secular variations in the pulse shape were still evident. Consequently, the orbital phase was divided into five groups of different lengths to which the same template (for a given 3-month interval) was applied as shown in fig. S2. In the final timing analysis, data from the two groups representing the bright phases (IV and V in fig. S2) were excluded to minimize the systematic errors caused by the orbital profile changes. Also, because of signal-to-noise and radio interference considerations, only data from Parkes and the GBT BCPM (Berkeley-Caltech Pulsar Machine) back end (*16*) were used in the B timing analysis.

All TOAs were transferred to Universal Coordinated Time (UTC) using the Global Positional System (GPS) to measure offsets of station clocks from national standards and Circular T of the Bureau International des Poids et Mesures (BIPM) to give offsets from UTC, and then to the nominally uniform BIPM Terrestrial Time (TT) time scale. These final TOAs were analyzed using the standard software package Tempo (*17*), fitting parameters according to the relativistic and theory-independent timing model of Damour and Deruelle (DD) (*11*, *18*). In addition to the DD model, we also applied the "DD-Shapiro" (DDS) model introduced by Kramer *et al.* (*19*). The DDS model is a modification of the DD model designed for highly inclined orbits. Rather than fitting for the Shapiro parameter s, the model uses the parameter $z_s \equiv -\ln(1 - s)$, which gives a more reliable determination of the uncertainties in z_s and hence in s. We quote the final result for the more commonly used parameter s and note that its value computed from z_s is in good agreement with the value obtained from a direct fit for s within the DD model. Derived pulsar and binary system parameters are listed in Table 1.

In the timing analysis for pulsar B, we used an unweighted fit to avoid biasing the fit toward bright orbital phases. Uncertainties in the timing parameters were estimated using Monte Carlo simulations of fake data sets for a range of TOA uncertainties, ranging from the minimum estimated TOA error to its maximum observed value of about 4 ms. For B, we also fitted for offsets between data sets derived from different templates in the fit because the observed profile changes prevent the establishment of a reliable phase relationship between the derived templates. This precludes a coherent fit across the whole orbit and hence limits the final timing precision for B. It cannot yet be excluded that different parts of B's magnetosphere are active and responsible for the observed emission at different orbital phases.

In the final fit, we adopted the astrometric parameters and the dispersion measure derived for A and held these fixed during the fit, because A's shorter period and more stable profile give much better timing precision than is achievable for B. Except for the semimajor axis—which is observable only as the projection onto the plane of the sky $x_B = (a_B/c)\sin i$, where a_B is

Table 1. Parameters for PSR J0737-3039A (A) and PSR J0737-3039B (B). The values were derived from pulse timing observations using the DD (*11*) and DDS (*19*) models of the timing analysis program Tempo and the Jet Propulsion Laboratory DE405 planetary ephemeris (*41*). Estimated uncertainties, given in parentheses after the values, refer to the least significant digit of the tabulated value and are twice the formal 1σ values given by Tempo. The positional parameters are in the DE405 reference frame, which is close to that of the International Celestial Reference System. Pulsar spin frequencies $\nu \equiv 1/P$ are in barycentric dynamical time (TDB) units at the timing epoch quoted in modified Julian days (MJD). The five Keplerian binary parameters (P_b, e, ω, T_0, and x) are derived for pulsar A. The first four of these (with an offset of 180° added to ω) and the position parameters were assumed when fitting for B's parameters. Five post-Keplerian parameters have now been measured. An independent fit of $\dot{\omega}$ for B yielded a value (shown in square brackets) that is consistent with the much more precise result for A. The value derived for A was adopted in the final analysis (*16*). The dispersion-based distance is based on a model for the interstellar electron density (*26*).

Timing parameter	PSR J0737-3039A	PSR J0737-3039B
Right ascension α	$07^h37^m51^s.24927(3)$	—
Declination δ	−30°39′40″.7195(5)	—
Proper motion in the RA direction (mas year^{-1})	−3.3(4)	—
Proper motion in declination (mas year^{-1})	2.6(5)	—
Parallax π (mas)	3(2)	—
Spin frequency ν (Hz)	44.054069392744(2)	0.36056035506(1)
Spin frequency derivative $\dot{\nu}$ (s^{-2})	$-3.4156(1) \times 10^{-15}$	$-0.116(1) \times 10^{-15}$
Timing epoch (MJD)	53,156.0	53,156.0
Dispersion measure DM (cm^{-3} pc)	48.920(5)	—
Orbital period P_b (day)	0.10225156248(5)	—
Eccentricity e	0.0877775(9)	—
Projected semimajor axis $x = (a/c)\sin i$ (s)	1.415032(1)	1.5161(16)
Longitude of periastron ω (°)	87.0331(8)	87.0331 + 180.0
Epoch of periastron T_0 (MJD)	53,155.9074280(2)	—
Advance of periastron $\dot{\omega}$ (°/year)	16.89947(68)	[16.96(5)]
Gravitational redshift parameter γ (ms)	0.3856(26)	—
Shapiro delay parameter s	0.99974(−39,+16)	—
Shapiro delay parameter r (μs)	6.21(33)	—
Orbital period derivative $\dot{P}_b$	$-1.252(17) \times 10^{-12}$	—
Timing data span (MJD)	52,760 to 53,736	52,760 to 53,736
Number of time offsets fitted	10	12
RMS timing residual σ (μs)	54	2169
Total proper motion (mas year^{-1})	4.2(4)	
Distance d(DM) (pc)	~500	
Distance d(π) (pc)	200 to 1,000	
Transverse velocity (d = 500 pc) (km s^{-1})	10(1)	
Orbital inclination angle (°)	88.69(−76,+50)	
Mass function ($M_\odot$)	0.29096571(87)	0.3579(11)
Mass ratio R	1.0714(11)	
Total system mass ($M_\odot$)	2.58708(16)	
Neutron star mass ($m_\odot$)	1.3381(7)	1.2489(7)

the semimajor axis of B's orbit and i is the orbital inclination angle—we also adopted A's Keplerian parameters (with 180° added to ω_A) and kept these fixed. We also adopted the PK parameter $\dot{\omega}$ (the rate of periastron advance) from the A fit because logically this must be identical for the two pulsars; this equality therefore does not implicitly make assumptions about the validity of any particular theory of gravity (see below). The same applies for the orbital decay parameter $\dot{P}_b$. In contrast, the PK parameters γ (the gravitational redshift and time dilation parameter) and s and r (the Shapiro-delay parameters) are asymmetric in the masses, and their values and interpretations differ for A and B. In practical terms, the relatively low timing precision for B does not require the inclusion of γ, s, r, or $\dot{P}_b$ in the timing model. We can, however, independently measure $\dot{\omega}_B$, obtaining a value of 16.96° ± 0.05° year^{-1}, consistent with the more accurately determined value for A.

Because the overall precision of our tests of GR is currently limited by our ability to measure x_B and hence the mass ratio $R \equiv m_A/m_B = x_B/x_A$ (see below), we adopted the following strategy to obtain the best possible accuracy for this parameter. We used the whole TOA data set for B in order to measure B's spin parameters P and $\dot{P}$, given in Table 1. These parameters were then kept fixed for a separate analysis of the concentrated 5-day GBT observing sessions at 820 MHz. On the time scale of the long-term profile evolution of B, each 5-day session represents a single-epoch experiment and hence requires only a single set of profile templates. The value of x_B obtained from a fit of this parameter only to the two 5-day sessions is presented in Table 1.

Because of the possible presence of unmodeled intrinsic pulsar timing noise and because not all TOA uncertainties are well understood, we adopt the common and conservative pulsar-timing practice of reporting twice the parameter uncertainties given by tempo as estimates of the 1σ uncertainties. Although we believe that our real measurement uncertainties are actually somewhat smaller than quoted, this practice facilitates comparison with previous tests of GR by pulsar observation. The timing model also includes timing offsets between the data sets for the different instruments represented by the entries in table S1. The final weighted root mean square post-fit residual is 54.2 μs. In addition to the spin and astrometric parameters, the Keplerian parameters of A's orbit, and five PK parameters, we also quote a tentative detection of a timing annual parallax that is consistent with the dispersion-derived distance. Further details are given in (*16*).

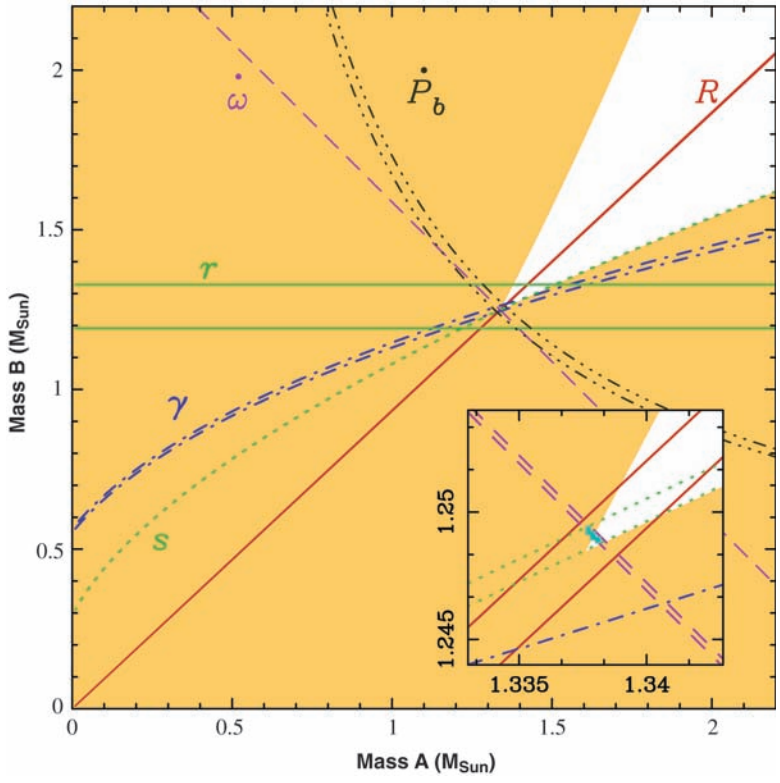

Fig. 1. Graphical summary of tests of GR parameters. Constraints on the masses of the two stars (A and B) in the PSR J0737-3039A/B binary system are shown; the inset is an expanded view of the region of principal interest. Shaded regions are forbidden by the individual mass functions of A and B because sin i must be ≤1. Other constraining parameters are shown as pairs of lines, where the separation of the lines indicates the measurement uncertainty. For the diagonal pair of lines labeled as R, representing the mass ratio derived from the measured semimajor axes of the A and B orbits, the measurement precision is so good that the line separation becomes apparent only in the inset. The other constraints shown are based on the measured PK parameters interpreted within the framework of general relativity. The PK parameter $\dot{\omega}$ describes the relativistic precession of the orbit, γ combines gravitational redshift and time dilation, and $\dot{P}_b$ represents the measured decrease in orbital period due to the emission of gravitational waves. The two PK parameters s and r reflect the observed Shapiro delay, describing a delay that is added to the pulse arrival times when propagating through the curved space-time near the companion. The intersection of all line pairs is consistent with a single point that corresponds to the masses of A and B. The current uncertainties in the observed parameters determine the size of this intersection area, which is marked in blue and reflects the achieved precision of this test of GR and the mass determination for A and B.

Tests of general relativity. Previous observations of PSR J0737-3039A/B (*8*, *9*) resulted in the measurement of R and four PK parameters: $\dot{\omega}$, γ, r, and s. Relative to these earlier results, the measurement precision for these parameters from PSR J0737-3039A/B has increased by up to two orders of magnitude. Also, we have now measured the orbital decay $\dot{P}_b$. Its value, measured at the 1.4% level after only 2.5 years of timing, corresponds to a shrinkage of the pulsars' separation at a rate of 7 mm per day. Therefore, we have measured five PK parameters for the system in total. Together with the mass ratio R, we have six different relationships that connect the two unknown masses for A and B with the observations. Solving for the two masses using R and one PK parameter, we can then use each further PK parameter to compare its observed value with that predicted by GR for the given two masses, providing four independent tests of GR. Equivalently, one can display these tests elegantly in a "mass-mass" diagram (Fig. 1). Measurement of the PK parameters gives curves on this diagram that are, in general, different for different theories of gravity but should intersect in a single point (i.e., at a pair of mass values) if the theory is valid (*12*).

As shown in Fig. 1, we find that all measured constraints are consistent with GR. The most precisely measured PK parameter currently available is the precession of the longitude of periastron, $\dot{\omega}$. We can combine this with the theory-independent mass ratio R to derive the masses given by the intersection region of their curves: $m_A = 1.3381 \pm 0.0007\ M_\odot$ and $m_B = 1.2489 \pm 0.0007\ M_\odot$, where $M_\odot$ is the mass of the Sun (*20*). Table 2 lists the resulting four independent tests that are currently available. All of them rely on comparison of our measured values of s, r, γ, and $\dot{P}_b$ with predicted values based on the masses defined by the intersection of the allowed regions for $\dot{\omega}$ and R in the m_A-m_B plane. The calculation of the predicted values is somewhat complicated by the fact that the orbit is nearly edge-on to the line of sight, so that the formal intersection region

actually includes parts of the plane disallowed by the Keplerian mass functions of both pulsars (see Fig. 1). To derive legitimate predictions for the various parameters, we used the following Monte Carlo method. A pair of trial values for $\dot{\omega}$ and x_B (and hence R and the B mass function) is selected from Gaussian distributions based on the measured central values and uncertainties. (The uncertainty on x_A is very small and is neglected in this procedure.) This pair of trial values is used to derive trial masses m_A and m_B, using the GR equation $\dot{\omega} = 3(P_b/2\pi)^{-5/3}(T_\odot M)^{2/3}(1 - e^2)^{-1}$, where e is the orbital eccentricity and $M = m_A + m_B$ and $T_\odot \equiv GM_\odot/c^3 = 4.925490947$ μs, and the mass-ratio equation $m_A/m_B = x_B/x_A$. If this trial mass pair falls in either of the two disallowed regions (based on the trial mass function for B), it is discarded. This procedure allows for the substantial uncertainty in the B mass function. Allowed mass pairs are then used to compute the other PK parameters, assuming GR. This procedure is repeated until large numbers of successful trials have accumulated. Histograms of the PK predictions are used to compute the expectation value and 68% confidence ranges for each of the parameters. These are the values given in Table 2.

The Shapiro delay shape illustrated in Fig. 2 gives the most precise test, with $s_{observed}/s_{predicted} = 0.99987 \pm 0.00050$ (*21*). This is by far the best available test of GR in the strong-field limit, having a higher precision than the test based on the observed orbit decay in the PSR B1913+16 system with a 30-year data span (*22*). As for the PSR B1534+12 system (*6*), the PSR J0737-3039A/B Shapiro-delay test is complementary to that of B1913+16 because it is not based on predictions relating to emission of gravitational radiation from the system (*23*). Most important, the four tests of GR presented here are qualitatively different from all previous tests because they include one constraint (R) that is independent of the assumed theory of gravity at the 1PN order. As a result, for any theory of gravity, the intersection point is expected to lie on the mass ratio line in Fig. 1. GR also passes this additional constraint.

In estimating the final uncertainty of x_B and hence of R, we have considered that geodetic precession will lead to changes to the system geometry and hence changes to the aberration of the rotating pulsar beam. The effects of aberration on pulsar timing are usually not separately measurable but are absorbed into a redefinition of the Keplerian parameters. As a result, the observed projected sizes of the semimajor axes, $x^{obs}_{A,B}$, differ from the intrinsic sizes, $x^{int}_{A,B}$, by a factor $(1 + \varepsilon^A_{A,B})$. The quantity ε_A depends for each pulsar A and B on the orbital period, the spin frequency, the orientation of the pulsar spin, and the system geometry (*12*). Although aberration should eventually become detectable in the timing, allowing the determination of a further PK parameter, at present it leads to an undetermined deviation of x^{obs} from x^{int}, where the latter is the relevant quantity for the mass ratio. The parameter $\varepsilon^A_{A,B}$ scales with pulse period and is therefore expected to be two orders of magnitude smaller for A than for B. However, because of the high precision of the A timing parameters, the derived value x^{obs}_A may already be significantly affected by aberration. This has (as yet) no consequences for the mass ratio $R = x^{obs}_B/x^{obs}_A$, as the uncertainty in R is dominated by the much less precise x^{obs}_B. We can explore the likely aberration corrections to x^{obs}_B for various possible geometries. Using a range of values given by studies of the double pulsar's emission properties (*24*), we estimate $\varepsilon^A_A \sim 10^{-6}$ and $\varepsilon^A_B \sim 10^{-4}$. The contribution of aberration therefore is at least one order of magnitude smaller than our current timing precision. In the future this effect may become important, possibly limiting the usefulness of R for tests of GR. If the geometry cannot be independently determined, we could use the observed deviations of R from the value expected within GR to determine ε^A_B and hence the geometry of B.

Space motion and inclination of the orbit. Because the measured uncertainty in $\dot{P}_b$ decreases approximately as $T^{-2.5}$, where T is the data span, we expect to improve our test of the radiative aspect of the system to the 0.1% level or better in about 5 years' time. For the PSR B1913+16 and PSR B1534+12 systems, the precision of the GR test based on the orbit-

Table 2. Four independent tests of GR provided by the double pulsar. Observed PK parameters were obtained by fitting a DDS timing model to the data. Values expected from GR take into account the masses determined from the intersection point of the mass ratio R and the periastron advance $\dot{\omega}$. Uncertainties refer to the last significant digits and were determined using Monte Carlo methods.

PK parameter	Observed value	Expected value from GR	Ratio of observed to expected value
$\dot{P}_b$	1.252(17)	1.24787(13)	1.003(14)
γ (ms)	0.3856(26)	0.38418(22)	1.0036(68)
s	0.99974(−39,+16)	0.99987(−48,+13)	0.99987(50)
r (μs)	6.21(33)	6.153(26)	1.009(55)

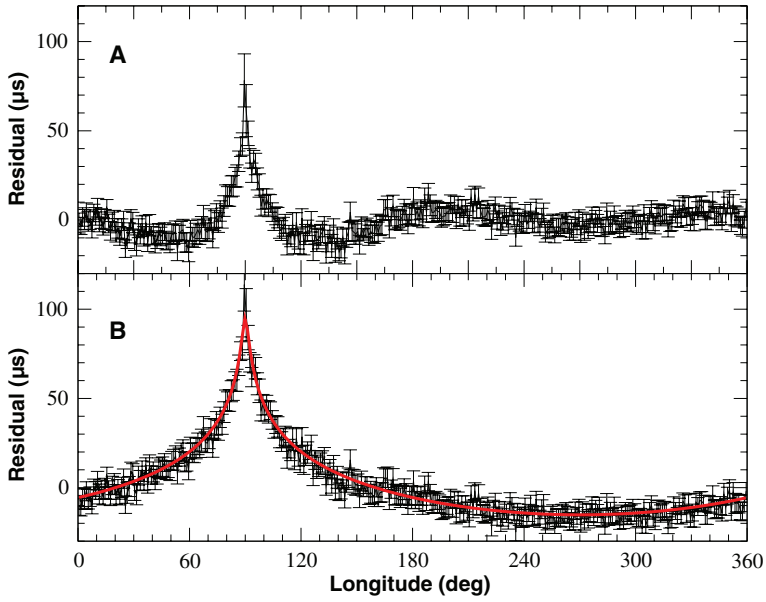

Fig. 2. Measurement of a Shapiro delay demonstrating the curvature of space-time. Timing residuals (differences between observed and predicted pulse arrival times) are plotted as a function of orbital longitude and illustrate the Shapiro delay for PSR J0737-3039A. (**A**) Observed timing residuals after a fit of all model parameters given in Table 1 except the Shapiro-delay terms r and s, which were set to zero and are not included in the fit. Although a portion of the delay is absorbed in an adjustment of the Keplerian parameters, a strong peak at 90° orbital longitude remains clearly visible. This is the orbital phase of A's superior conjunction (i.e., when it is positioned behind B as viewed from Earth), so that its pulses experience a delay when moving through the curved space-time near B. The clear detection of structure in the residuals over the whole orbit confirms the detection of the Shapiro delay, which is isolated in (**B**) by holding all parameters to their best-fit values given in Table 1, except the Shapiro delay terms (which were set to zero). The red line shows the predicted delay at the center of the data span. In both cases, residuals were averaged in 1° bins of longitude.

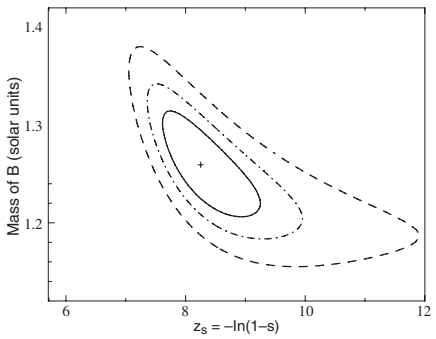

Fig. 3. Contour plots of the χ^2 distribution in the plane of the Shapiro-delay parameter $z_s \equiv -\ln(1 - s)$ and the mass of the B pulsar, m_B. The contours correspond to 68%, 95%, and 99% confidence limits.

decay rate is severely limited by the uncertainty in the differential acceleration of the Sun and the binary system in the galactic gravitational potential as well as by the uncertainty in pulsar distance (*6*, *25*). For PSR J0737-3039A/B, both of these corrections are very much smaller than for these other systems. On the basis of the measured dispersion measure and a model for the galactic electron distribution (*26*), PSR J0737-3039A/B is estimated to be about 500 pc from Earth. From the timing data we have measured a marginally significant value for the annual parallax, 3 ± 2 mas, corresponding to a distance of 200 to 1000 pc (Table 1), which is consistent with the dispersion-based distance that was also used for studies of detection rates in gravitational wave detectors (*8*). The observed proper motion of the system (Table 1) and differential acceleration in the galactic potential (*27*) then imply a kinematic correction to $\dot{P}_b$ at the 0.02% level or less. Independent distance estimates also can be expected from measurements of the annual parallax by very-long-baseline interferometry observations, allowing a secure compensation for this already small effect. A measurement of $\dot{P}_b$ at the 0.02% level or better will provide stringent tests for alternative theories of gravity. For example, limits on some scalar-tensor theories will surpass the best current solar system tests (*28*).

In GR, the parameter s can be identified with $\sin i$, where i is the inclination angle of the orbit. The value of s given in Table 1 corresponds to $i = 88°.69^{+0°.50}_{-0°.76}$. On the basis of scintillation observations of both pulsars over the short time interval when A is close to superior conjunction, Coles *et al*. (*29*) derived a value for $|i - 90°|$ of 0°.29 ± 0°.14. This is consistent with our measurement only at the 3σ level. As mentioned above, we used the DDS model to solve for the Shapiro delay. The resulting χ^2 contours in the z_s-m_B plane are shown in Fig. 3. The value and uncertainty range for s quoted in Table 1 correspond to the peak and range of the 68% contour. Because of the nonlinear relationship between z_s and s, the uncertainty distribution in s (and hence in i) corresponding to these contours is very asymmetric with a very steep edge on the 90° side. Only close to the 99% confidence limit is the timing result consistent with the scintillation-derived value. We note that the scintillation measurement is based on the correlation of the scintillation fluctuations of A and B over the short interval when A is close to superior conjunction (i.e., behind B). In contrast, the measurement of i from timing measurements depends on the detection of significant structure in the post-fit residuals after a portion of the Shapiro delay is absorbed in the fit for x_A (*30*). As shown in Fig. 2, the Shapiro delay has a signature that is spread over the whole orbit and hence can be cleanly isolated. We also examined the effects on the Shapiro delay of using only low- or high-frequency data, and we found values of s consistent within the errors in each case. The scintillation result is based on the plasma properties of the interstellar medium and may also be affected by possible refraction effects in B's magnetosphere. We believe that the timing result is much less susceptible to systematic errors and is therefore more secure.

Scintillation observations have also been used to deduce the system transverse velocity. Ransom *et al*. (*31*) derived a value of 141 ± 8.5 km s^{-1}, whereas Coles *et al*. (*29*) obtained 66 ± 15 km s^{-1} after considering the effect of anisotropy in the scattering screen. Both of these values are in stark contrast to the value of 10 ± 1 km s^{-1} (relative to the solar system barycenter) obtained from pulsar timing (Table 1). We note that the scintillation-based velocity depends on a number of assumptions about the properties of the effective scattering screen. In contrast, the proper motion measurement has a clear and unambiguous timing signature, although the transverse velocity itself scales with the pulsar distance. Even allowing that unmodeled effects of Earth motion could affect the published scintillation velocities by about 30 km s^{-1}, the dispersion-based distance would need to be underestimated by a factor of 3 to 4 to make the velocities consistent. We believe this is very unlikely, particularly as the tentative detection of a parallax gives us some confidence in the dispersion-based distance estimate. Hence, we believe that our timing results for both inclination angle and transverse velocity are less susceptible to systematic errors and are therefore more secure than those based on scintillation.

We note that because the inclination angle is significantly different from 90°, gravitational lensing effects (*32*) can be neglected. The implied low space velocity, the comparatively low derived mass for B, and the low orbit eccentricity are all consistent with the idea that the B pulsar may have formed by a mechanism different from the usually assumed core collapse of a helium star (*33*, *34*). A discussion of its progenitor is presented in (*35*). We also note that, as expected for a double–neutron star system, there is no evidence for variation in dispersion measure as a function of orbital phase.

Future tests. In contrast to all previous tests of GR, we are now reaching the point with PSR J0737-3039A where expressions of PK parameters to only 1PN order may no longer be sufficient for a comparison of theoretical predictions with observations. In particular, we have measured $\dot{\omega}$ so precisely (i.e., to a relative precision approaching 10^{-5}) that we expect corrections at the 2PN level (*13*) to be observationally significant within a few years. These corrections include contributions expected from spin-orbit coupling (*36*, *37*). A future determination of the system geometry and the measurement of two other PK parameters at a level of precision similar to that for $\dot{\omega}$ would allow us to measure the moment of inertia of a neutron star for the first time (*13*, *38*). Although this measurement is potentially very difficult, a determination of A's moment of inertia to a precision of only 30% would allow us to distinguish between a large number of proposed equations of state for dense matter (*39*, *40*). The double pulsar would then not only provide the best tests of theories of gravity in the strong-field regime, as presented here, but would also give insight into the nature of superdense matter.

References and Notes

1. C. Will, *Living Rev. Relativity* **4**, 4 (2001).
2. B. Bertotti, L. Iess, P. Tortora, *Nature* **425**, 374 (2003).
3. J. G. Williams, S. G. Turyshev, D. H. Boggs, *Phys. Rev. Lett.* **93**, 261101 (2004).
4. T. Damour, G. Esposito-Farèse, *Phys. Rev. D* **58**, 1 (1998).
5. J. H. Taylor, J. M. Weisberg, *Astrophys. J.* **345**, 434 (1989).
6. I. H. Stairs, S. E. Thorsett, J. H. Taylor, A. Wolszczan, *Astrophys. J.* **581**, 501 (2002).
7. Stairs *et al*. (*6*) found agreement of their measured values for PSR B1534+12 with GR at the 0.05% level, but the measurement uncertainty on the most precisely measured parameter in the test, s, was only 0.7%.
8. M. Burgay *et al.*, *Nature* **426**, 531 (2003).
9. A. G. Lyne *et al.*, *Science* **303**, 1153 (2004); published online 8 January 2004 (10.1126/science.1094645).
10. M. Burgay *et al.*, *Astrophys. J.* **624**, L113 (2005).
11. T. Damour, N. Deruelle, *Ann. Inst. H. Poincare Phys. Theor.* **44**, 263 (1986).
12. T. Damour, J. H. Taylor, *Phys. Rev. D* **45**, 1840 (1992).
13. T. Damour, G. Schäfer, *Nuovo Cim.* **101**, 127 (1988).
14. M. Kramer *et al.*, *Astrophys. J.* **526**, 957 (1999).
15. R. N. Manchester *et al.*, *Astrophys. J.* **621**, L49 (2005).
16. See supporting material on *Science* Online.
17. Tempo (www.atnf.csiro.au/research/pulsar/tempo).
18. T. Damour, N. Deruelle, *Ann. Inst. H. Poincare Phys. Theor.* **43**, 107 (1985).
19. M. Kramer *et al.*, *Ann. Phys.* **15**, 34 (2006).
20. The true masses will deviate from these values by an unknown, but essentially constant, Doppler factor, probably of order 10^{-3} or less (*12*). Moreover, what is measured is a product containing Newton's gravitational constant G. The relative uncertainty of G of 1.5×10^{-4} limits our knowledge of any astronomical mass in kilograms, but because the product $T_\odot = GM_\odot/c^3 = 4.925490947$ μs is known to very high precision, masses can be measured precisely in solar units.
21. Note that s has the same relative uncertainty as our determination of the masses.

22. J. M. Weisberg, J. H. Taylor, in *Binary Radio Pulsars*, F. Rasio, I. H. Stairs, Eds. (Astronomical Society of the Pacific, San Francisco, 2005), pp. 25–31.
23. J. H. Taylor, A. Wolszczan, T. Damour, J. M. Weisberg, *Nature* **355**, 132 (1992).
24. M. Lyutikov, *Mon. Not. R. Astron. Soc.* **362**, 1078 (2005).
25. T. Damour, J. H. Taylor, *Astrophys. J.* **366**, 501 (1991).
26. J. M. Cordes, T. J. W. Lazio, http://arxiv.org/abs/astro-ph/0207156 (2002).
27. K. Kuijken, G. Gilmore, *Mon. Not. R. Astron. Soc.* **239**, 571 (1989).
28. T. Damour, G. Esposito-Farèse, in preparation.
29. W. A. Coles, M. A. McLaughlin, B. J. Rickett, A. G. Lyne, N. D. R. Bhat, *Astrophys. J.* **623**, 392 (2005).
30. C. Lange *et al.*, *Mon. Not. R. Astron. Soc.* **326**, 274 (2001).
31. S. M. Ransom *et al.*, *Astrophys. J.* **609**, L71 (2004).
32. R. R. Rafikov, D. Lai, *Phys. Rev. D* **73**, 063003 (2006).
33. E. Pfahl, S. Rappaport, P. Podsiadlowski, H. Spruit, *Astrophys. J.* **574**, 364 (2002).
34. T. Piran, N. J. Shaviv, *Phys. Rev. Lett.* **94**, 051102 (2005).
35. I. H. Stairs, S. E. Thorsett, R. J. Dewey, M. Kramer, C. McPhee, *Mon. Not. R. Astron. Soc.*, in press (2006).
36. T. Damour, R. Ruffini, *C. R. Acad. Sci. Paris Ser. I Math.* **279**, 971 (1974).
37. B. M. Barker, R. F. O'Connell, *Astrophys. J.* **199**, L25 (1975).
38. N. Wex, *Class. Quantum Grav.* **12**, 983 (1995).
39. I. A. Morrison, T. W. Baumgarte, S. L. Shapiro, V. R. Pandharipande, *Astrophys. J.* **617**, L135 (2004).
40. J. M. Lattimer, B. F. Schutz, *Astrophys. J.* **629**, 979 (2005).
41. E. M. Standish, *Astron. Astrophys.* **336**, 381 (1998).
42. We thank T. Damour and N. Wex for useful discussions. The Parkes radio telescope is part of the Australia Telescope, which is funded by the Commonwealth of Australia for operation as a National Facility managed by CSIRO. The National Radio Astronomy Observatory (NRAO) is a facility of the NSF operated under cooperative agreement by Associated Universities Inc. GASP is funded by a Natural Sciences and Engineering Research Council of Canada (NSERC) RTI-1 grant (I.H.S.) and by NSF grants to D. Backer and D. Nice. We thank P. Demorest, R. Ramachandran, and J. van Leeuwen for their contributions to GASP hardware and software development. Supported by an NSERC University Faculty Award (I.H.S.), by the Italian Ministry of University and Research under the national program Cofin 2003 (M.B., A.P., and N.D'A.), and by NSF, NASA, and NRAO (F.C.). Pulsar research at the University of British Columbia is supported by an NSERC Discovery Grant.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1132305/DC1
SOM Text
Figs. S1 to S4
Tables S1 and S2

10 July 2006; accepted 7 September 2006
Published online 14 September 2006;
10.1126/science.1132305
Include this information when citing this paper.

Out of the Tropics: Evolutionary Dynamics of the Latitudinal Diversity Gradient

David Jablonski,[1]* Kaustuv Roy,[2] James W. Valentine[3]

The evolutionary dynamics underlying the latitudinal gradient in biodiversity have been controversial for over a century. Using a spatially explicit approach that incorporates not only origination and extinction but immigration, a global analysis of genera and subgenera of marine bivalves over the past 11 million years supports an "out of the tropics" model, in which taxa preferentially originate in the tropics and expand toward the poles without losing their tropical presence. The tropics are thus both a cradle and a museum of biodversity, contrary to the conceptual dichotomy dominant since 1974; a tropical diversity crisis would thus have profound evolutionary effects at all latitudes.

The most striking large-scale pattern in biological diversity is the dramatic increase in the number of species and higher taxa from the poles to the tropics. This taxonomic trend, commonly called the latitudinal diversity gradient (LDG), has been documented in the multicellular biotas of forests, grasslands, wetlands, continental shelves, the open ocean, and even the deep sea; it characterizes plants, fungi, marine and freshwater invertebrates, and all of the vertebrate classes (*1*). The history of the LDG extends back through the Mesozoic into the Paleozoic (*2*–*7*), although the slope of the gradient has varied over time and the trend might even have disappeared for a time if any of the mass extinctions were disproportionately severe in the tropics (*8*).

Although the existence of the LDG has been known for more than a century (*9*, *10*) and has been quantified hundreds of times (*1*), it remains the "major, unexplained pattern of natural history" [Ricklefs in (*11*)], with "an astonishing lack of consensus about the mechanisms leading to this variation in diversity" (*1*). Recent work has focused primarily on ecological explanations for the LDG (*9*, *12*–*15*), and although these analyses have found interesting correlations between diversity and environmental variables, they reveal little about the evolutionary dynamics of the species and lineages that established and maintain the LDG (*16*, *17*). Because virtually all possible combinations of the key evolutionary parameters have been proposed to shape the LDG (table S1), progress in this area depends on empirical data that can falsify alternatives. Here we (i) outline a framework for evaluating the spatial and temporal dynamics that underlie the present-day LDG, (ii) synthesize previous work from this perspective, and (iii) present paleontological analyses that falsify the classic portrayal of the tropics as either a cradle or a museum of biodiversity (*18*).

Cradles and Museums

From an evolutionary perspective, large-scale spatial patterns of biodiversity depend on three variables: origination rates (O), extinction rates (E), and changes in geographic distributions (expressed here as I, for immigration into a latitudinal bin) of taxa. For a simple two-box model, with the tropics and extratropics denoted as subscripts, diversity in the tropics (D_T) is determined by $O_T - E_T + I_T$, and diversity in the extratropics (D_E) by $O_E - E_E + I_E$ (Fig. 1). With this notation, it can easily be seen that a latitudinal gradient in richness, with $D_T > D_E$, can result from many different combinations of these variables. Theoretically, the extinction terms could represent either true global extinction of taxa, local extinction for a particular spatial bin, or a combination of the two. Estimating local extinction rates using paleontological data is generally difficult owing to incomplete spatial sampling, and even more difficult using phylogenetic information. In addition, our empirical results suggest that the effect of local extinction is much smaller than that of range expansion, at least for marine bivalves. Thus, as in most previous studies (table S1), our discussion of the role of extinction in shaping the LDG focuses primarily on global processes.

The simplest evolutionary models for the LDG assume that taxa are static in their geographic distributions ($I_T = I_E = 0$) and treat the greater number of species and higher taxa in the tropics as the result of either a higher rate of origination of species and lineages ($O_T > O_E$) or lower extinction rates as compared to extratropical regions ($E_T < E_E$). For example, Wallace (*19*) attributed high tropical diversity to a more stable climatic history, which allowed more time to accumulate taxa ($E_T < E_E$), and this view has found proponents ever since (*20*) (table S1). Others have argued that extinction rates are high in the tropics but are outstripped by even higher origination rates ($E_T > E_E$, $O_T \gg O_E$) (*21*). The importance of origination and extinction in generating the LDG was highlighted in Stebbins' (*18*) famous metaphor of the tropics as a cradle or a museum, and this memorable dichotomy has been the dominant paradigm ever since.

[1]Department of Geophysical Sciences, University of Chicago, 5734 South Ellis Avenue, Chicago, IL 60637, USA. [2]Section of Ecology, Behavior and Evolution, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093–0116, USA. [3]Department of Integrative Biology, University of California, Berkeley, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed. E-mail: djablons@uchicago.edu