On the unreasonable effectiveness of the post-Newtonian approximation in gravitational physics

Clifford M. Will¹

McDonnell Center for the Space Sciences, Department of Physics, Washington University, St. Louis, MO

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2007.

Contributed by Clifford M. Will, February 25, 2011 (sent for review February 17, 2011)

The post-Newtonian approximation is a method for solving Einstein's field equations for physical systems in which motions are slow compared to the speed of light and where gravitational fields are weak. Yet it has proven to be remarkably effective in describing certain strong-field, fast-motion systems, including binary pulsars containing dense neutron stars and binary black hole systems inspiraling toward a final merger. The reasons for this effectiveness are largely unknown. When carried to high orders in the post-Newtonian sequence, predictions for the gravitational-wave signal from inspiraling compact binaries will play a key role in gravitational-wave detection by laser-interferometric observatories.

general relativity | gravitational radiation

The first detection of gravitational radiation, anticipated to occur during this decade, will undoubtedly be a triumph of experimental physics. It will initiate a new kind of astronomy. But it will also shine a spotlight on the theory of general relativity itself. Gravitational radiation is a crucial prediction of Einstein's theory; indeed it is a natural prediction, given that the theory is built on a foundation of Lorentz invariance, which carries with it the concept of a limiting speed for interactions.

On the other hand, general relativity is a notoriously complicated, nonlinear, tensorial theory of the gravitational field. Almost no physically useful exact solutions of the theory are known, and those that are known are endowed with such high degrees of symmetry that their realm of validity is limited. To be sure, two of these solutions, due to Schwarzschild and Kerr, have proven to be of enormous importance, describing as they do the spacetime of isolated black holes, now widely accepted as being ubiquitous throughout the universe. Part of the Schwarzschild solution also describes the exterior geometry of any static spherical star or planet. Another extremely useful, albeit special solution is the Friedman–Robertson–Walker metric of the standard model of homogeneous and isotropic big-bang cosmology.

But by its very nature, gravitational radiation involves spacetimes that are highly nonsymmetrical and highly dynamical. No exact solution of Einstein's equations is known that describes the emission and propagation of gravitational waves from a source, and the reaction of the source to the emission of those waves.

As a result, most of our understanding of gravitational radiation has come from *approximations* to Einstein's equations. One class of approximations assumes that the gravitational fields in and around the source are suitably weak (the fields of the propagating waves weaken progressively as they leave the source), and that the motions within the source are suitably slow compared to the speed of light. This class includes schemes known as post-Newtonian theory, which will be the main subject of this paper, and a related scheme known as post-Minkowski theory. The underlying idea is to treat spacetime as being that of flat Minkowski spacetime as the zeroth approximation, and to modify it by successive corrections.

Another class of approximations takes a known exact solution of Einstein's equations, such as the black hole solutions of Schwarzschild or Kerr, and introduces small perturbations of those spacetimes, induced, for example, by a small particle orbiting the hole. In cosmology, perturbations of the Friedman–Robertson–Walker solution permit treatment of the growth of large-scale structure in the universe and fluctuations in the cosmic background radiation.

A rather different approach to solving Einstein's equations is “numerical relativity,” which endeavors to formulate and solve the exact equations to a precision limited only by available computer resources for highly dynamical, highly asymmetrical situations using accurate and robust numerical computations. In recent years, numerical relativity has significantly enhanced our understanding of colliding black holes and neutron stars and the associated emission of gravitational radiation.

Just as exact solutions of Einstein's equations have limited realms of validity, so too do approximation schemes. Specifically, the post-Newtonian (PN) approximation is formally limited to weak gravitational fields and slow motions. Yet recent experience has shown that the post-Newtonian approximation is “unreasonably effective” in describing physical systems whose characteristics either lie outside the technical realm of validity of the approximation or push up against the boundary of that realm.

The use of the term “unreasonably effective,” and indeed the title of this paper, have been shamelessly appropriated from a famous 1960 paper by Eugene Wigner, entitled “On the unreasonable effectiveness of mathematics in the physical sciences” (1). In that paper, Wigner states “the enormous usefulness of mathematics in the natural sciences is something bordering on the mysterious and ... there is no rational explanation for it.” Einstein presented similar ideas in an address to the Prussian Academy of Sciences in 1921.

While the considerations of Einstein and Wigner deal with deep questions of the nature of knowledge of the physical world and how it is acquired and assembled, the topic of this paper is much more narrow. Our purpose will be to illustrate the various ways in which the post-Newtonian approximation has proven its extraordinary effectiveness in gravitational physics. Nevertheless it is no less mysterious: We have no good understanding of why this approximation to general relativity should be so effective.

Because most of these applications of post-Newtonian theory involve the motions of gravitating bodies in their mutual gravitational fields and the emission of gravitational radiation, we will begin with a history of the “problem of motion,” one of the central challenges in the development of general relativity. The problem has at times been contentious. We will then describe

Author contributions: C.M.W. wrote the paper.

The author declares no conflict of interest.

¹E-mail: cmw@wuphys.wustl.edu.

See Profile on page 5935.

briefly the nature and structure of the post-Newtonian approximation, and will review some of the recent developments in the subject, particularly those involving going to very high orders in the approximation. We will describe the usefulness of the PN approximation for characterizing alternative theories of gravity and as a tool in experimental gravitation. Then we show how the PN approximation effectively and accurately accounts for the observed behavior of binary pulsar systems, despite the presence of strong gravitational fields in the interiors of the orbiting neutron stars. Finally, we will describe how the PN approximation, when carried to high enough orders, effectively describes the motion of and gravitational radiation from inspiralling binary black hole systems, well into the strong-field region, and how its predictions merge smoothly with those from numerical relativity. We will conclude by returning to the mystery of this effectiveness.

Motion and Radiation in General Relativity: A History

At the most primitive level, the problem of motion in general relativity is relatively straightforward and was an integral part of the theory as proposed by Einstein*. A “test particle”—that is, a particle whose mass is sufficiently small that its own contribution to the curvature of spacetime can be ignored—moves on a geodesic of the curved spacetime in which it finds itself. Underlying this concept is the “weak equivalence principle,” which states that the acceleration of a suitably small body in an external gravitational field is independent of its internal structure or composition, a principle verified by modern experiments to parts in 10^{13} . Using the geodesic equation and an approximate solution for the spacetime metric around the sun, Einstein was able in 1915 to obtain the relativistic contribution to the perihelion advance of Mercury, in agreement with observations.

The first attempts to treat the motion of multiple bodies, each with a finite mass, were made in the period 1916–17 by de Sitter, Lorentz, and Droste (34). They derived the metric and equations of motion for a system of N bodies, in what today would be called the first post-Newtonian approximation of general relativity (de Sitter’s equations turned out to contain some important errors). In 1916, Einstein took the first crack at a study of gravitational radiation, deriving the energy emitted by a body such as a rotating rod or dumbbell, held together by nongravitational forces (5). He made some unjustified assumptions as well as a trivial numerical error (later corrected by Eddington) (6), but the underlying conclusion that dynamical systems would radiate gravitational waves was correct.

The next significant advance in the problem of motion came 20 years later. In 1938, Einstein, Infeld, and Hoffman published the now legendary “EIH” paper, a calculation of the N -body equations of motion using only the vacuum field equations of general relativity (7). They treated each body in the system as a spherically symmetric object whose nearby vacuum exterior geometry approximated that of the Schwarzschild metric of a static spherical star. They then solved the vacuum field equations for the metric between each body in the system in a weak-field, slow-motion approximation. Then, using a primitive version of what today would be called “matched asymptotic expansions” they showed that, in order for the nearby metric of each body to match smoothly to the interbody metric at each order in the expansion, certain conditions on the motion of each body had to be met. Together, these conditions turned out to be equivalent to the Droste–Lorentz N -body equations of motion. Remarkably, the internal structure of each body was irrelevant, apart from the requirement that its nearby field be approximately spherically symmetric.

Around the same time, there occurred an unusual detour in the problem of motion. Using equations of motion based on de

Sitter’s paper, specialized to two bodies, Levi-Civita (8) showed that the center of mass of a binary star system would suffer an acceleration in the direction of the pericenter of the orbit, in an amount proportional to the difference between the two masses, and to the eccentricity of the orbit. Such an effect would be a violation of the conservation of momentum for isolated systems caused by relativistic gravitational effects. Levi-Civita even went so far as to suggest looking for this effect in selected nearby close binary star systems. However, Eddington and Clark (9) quickly pointed out that Levi-Civita had based his calculations on de Sitter’s flawed work; when correct two-body equations of motion were used, the effect vanished, and momentum conservation was upheld. Robertson confirmed this using the EIH equations of motion (10). Ironically, the acceleration of the center of mass of a binary system is today a subject of great theoretical and astrophysical interest, albeit for a very different reason. We will return to this subject later.

Roughly 20 more years would pass before another major attack on the problem of motion. This period was the continuation of a time of relative dormancy for the entire subject of general relativity that lasted from the 1920s until the 1960s. This dormancy resulted in part from the lack of experimental or observational relevance for the theory, in part from the perceived complexity of the theory, and in part from the emergence of new fields of physics such as nuclear and particle physics in the middle part of the twentieth century.

But in the middle 1960s, when a revival of general relativity was in its early phase, Fock in the USSR and Chandrasekhar in the US independently developed and systematized the post-Newtonian approximation in a form that laid the foundation for modern post-Newtonian theory (11, 12). They developed a full post-Newtonian hydrodynamics, with the ability to treat realistic, self-gravitating bodies of fluid, such as stars and planets. In the suitable limit of “point” particles, or bodies whose size is small enough compared to the interbody separations that finite-size effects such as spin and tidal interactions can be ignored, their equations of motion could be shown to be equivalent to the EIH and the Droste–Lorentz equations of motion. Chandrasekhar and his students also began extending the theory to higher orders in the post-Newtonian approximation.

An important byproduct of the Fock–Chandrasekhar work was the discovery by Nordtvedt (13) that, in theories of gravity alternative to general relativity, the motion of self-gravitating bodies *could* depend on their internal structure, in contrast to general relativity, where a body’s internal structure is irrelevant to its motion (barring tidal or spin effects). As a result of this “Nordtvedt effect,” the earth and the moon could fall toward the sun with slightly different accelerations, because of the small difference in their internal gravitational binding energy per unit mass. Nordtvedt’s discovery led to an important new test of general relativity using Lunar laser ranging, and to the development by Nordtvedt and by Will of the parametrized post-Newtonian (PPN) framework for treating alternative theories and experimental tests. This will be the subject of a later section.

The next important period in the history of the problem of motion was 1974–79, initiated by the 1974 discovery of the binary pulsar PSR 1913+16 by Hulse and Taylor (14). Around the same time there occurred the first serious attempt to calculate the head-on collision of two black holes using purely numerical solutions of Einstein’s equations, by Smarr and collaborators (15), building on the pioneering work by Hahn and Lindquist (16).

The binary pulsar consists of two neutron stars, one an active pulsar detectable by radio telescopes, the other very likely an old, inactive pulsar. Each neutron star has a mass of around 1.4 solar masses. The orbit of the system was seen immediately to be quite relativistic, with an orbital period of only eight hours, and a mean orbital speed of 200 km/s, some four times faster than Mercury

*This history will by necessity be personal and selective. For a detailed technical and historical review of the problem of motion, see Damour (2).

in its orbit. Within weeks of its discovery, numerous authors pointed out that PSR 1913+16 would be an important new testing ground for general relativity. In particular, it could provide for the first time a test of the effects of the emission of gravitational radiation on the orbit of the system.

However, the discovery revealed an ugly truth about the “problem of motion.” As Ehlers et al. pointed out in an influential 1976 paper (17), the general relativistic problem of motion and radiation was full of holes large enough to drive trucks through. They pointed out that most treatments of the problem used “delta functions” as a way to approximate the bodies in the system as point masses. As a consequence, the “self-field,” the gravitational field of the body evaluated at its own location, becomes infinite. While this is not a major issue in Newtonian gravity or classical electrodynamics, the nonlinear nature of general relativity requires that this infinite self-field contribute to gravity. In the past, such infinities had been simply swept under the rug. Similarly, because gravitational energy itself produces gravity it thus acts as a source throughout spacetime. This means that, when calculating radiative fields, integrals for the multipole moments of the source that are so useful in treating radiation begin to diverge. These divergent integrals had also been routinely swept under the rug. Ehlers et al. further pointed out that the true boundary condition for any problem involving radiation by an isolated system should be one of “no incoming radiation” from the past. Connecting this boundary condition with the routine use of retarded solutions of wave equations was not a trivial matter in general relativity. Finally they pointed out that there was no evidence that the post-Newtonian approximation, so central to the problem of motion, was a convergent or even asymptotic sequence. Nor had the approximation been carried out to high enough order to make credible error estimates.

During this time, some authors even argued that the “quadrupole formula” for the gravitational energy emitted by a system (see below), while correct for a rotating dumbbell as calculated by Einstein, was actually *wrong* for a binary system moving under its own gravity. The discovery in 1979 that the rate of decay of the orbit of the binary pulsar was in agreement with the standard quadrupole formula made some of these arguments moot. Yet the question raised by Ehlers et al. was still relevant: Is the quadrupole formula for binary systems an actual prediction of general relativity?

Motivated by the Ehlers et al. critique, numerous workers began to address the holes in the problem of motion, and by the late 1990s most of the criticisms had been answered, particularly those related to divergences. The one question that remains open is the nature of the post-Newtonian sequence; we still do not know if it converges, diverges or is asymptotic. Despite this, it has proven to be remarkably effective.

The final important development in this history of the problem of motion was the proposal to build large-scale laser interferometric gravitational-wave observatories, both on the ground and in space (18–20). It was quickly realized that, in order to maximize the likelihood of detecting the leading candidate source of gravitational waves—the final inspiral and merger of binary systems of neutron stars or black holes—*extremely* accurate theoretical predictions of the gravitational waveform signal would be needed. This meant that calculations of the equations of motion and gravitational waves from binary systems would have to be carried out to many orders beyond the simple post-Newtonian approximation. The completion of this ambitious program by many groups worldwide has led to the discovery of the unreasonable effectiveness of post-Newtonian theory in the extreme realm of the merger of compact astronomical bodies.

The Post-Newtonian Approximation

The post-Newtonian approximation is based on the assumption that gravitational fields inside and around bodies are weak and

that characteristic motions of matter are slow compared to the speed of light. This means that one can characterize the system in question by a small parameter ϵ , where

$$\epsilon \sim (v/c)^2 \sim GM/rc^2 \sim p/\rho c^2, \quad [1]$$

where v , M and r denote the characteristic velocity, mass, and size or separation within the system; p and ρ are the characteristic pressure and density within the bodies; G and c are Newton’s gravitational constant and the speed of light, respectively.

One then incorporates this approximation into methods for solving Einstein’s equations. Those equations, $G_{\mu\nu} = 8\pi(G/c^4)T_{\mu\nu}$, are elegant and deceptively simple, showing how geometry (in the form of the Einstein tensor $G_{\mu\nu}$, which is a function of spacetime curvature) is generated by matter (in the form of the material energy-momentum tensor $T_{\mu\nu}$). However, this is not the most useful form for actual calculations. For post-Newtonian calculations, a far more useful form is the set of so-called “relaxed” Einstein equations:

$$\square h^{\alpha\beta} = -16\pi(G/c^4)\tau^{\alpha\beta}, \quad [2]$$

where $\square \equiv -\partial^2/\partial(ct)^2 + \nabla^2$ is the flat-spacetime wave operator, $h^{\alpha\beta}$ is a “gravitational tensor potential” related to the deviation of the spacetime metric $g_{\alpha\beta}$ from its flat-spacetime Minkowski form $\eta_{\alpha\beta}$ by the formula $h^{\alpha\beta} \equiv \eta^{\alpha\beta} - (-g)^{1/2}g^{\alpha\beta}$, where g is the determinant of $g_{\alpha\beta}$, and where a particular coordinate system has been specified by the de Donder or harmonic gauge condition $\partial h^{\alpha\beta}/\partial x^\beta = 0$ (summation on repeated indices is assumed, with $x^0 = ct$). Because we assume that gravity is weak everywhere, the field $h^{\alpha\beta}$ is “small.” This form of Einstein’s equations bears a striking similarity to Maxwell’s equations for the vector potential A^α in Lorentz gauge: $\square A^\alpha = -4\pi J^\alpha$, $\partial A^\alpha/\partial x^\alpha = 0$. The key difference is that the source on the right hand side of Eq. 2 is given by the “effective” energy-momentum pseudotensor,

$$\tau^{\alpha\beta} = (-g)T^{\alpha\beta} + (16\pi)^{-1}\Lambda^{\alpha\beta}, \quad [3]$$

where $\Lambda^{\alpha\beta}$ is the nonlinear “field” contribution given by terms quadratic (and higher) in $h^{\alpha\beta}$ and its derivatives (see ref. 21, equations 20.20, 20.21 for formulae). In general relativity, the gravitational field itself is a source of gravity, a reflection of the nonlinearity of Einstein’s equations, and in contrast to the linearity of Maxwell’s equations.

Eq. 2 is an exact restatement of Einstein’s equations and depends only on the assumption that spacetime can be covered by harmonic coordinates. It is called “relaxed” because it can be solved formally as a functional of source variables without specifying the motion of the source, in the form

$$h^{\alpha\beta}(t, \mathbf{x}) = \frac{4G}{c^4} \int_{\mathcal{C}} \frac{\tau^{\alpha\beta}(t - |\mathbf{x} - \mathbf{x}'|/c, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d^3x', \quad [4]$$

where the integration is over the past flat-spacetime null cone $\mathcal{C}$ of the field point $(t, \mathbf{x})$. The motion of the source is then determined either by the equation $\partial\tau^{\alpha\beta}/\partial x^\beta = 0$ (which follows from the harmonic gauge condition), or from the usual covariant equation of motion $\nabla_\beta T^{\alpha\beta} = 0$, where ∇_β denotes a covariant derivative. This formal solution can then be iterated in a weak-field ($\|h^{\alpha\beta}\| \ll 1$) approximation. One begins by substituting $h_0^{\alpha\beta} = 0$ into the source $\tau^{\alpha\beta}$ in Eq. 4, and solving for the first iterate $h_1^{\alpha\beta}$, and then repeating the procedure sufficiently many times to achieve a solution of the desired accuracy. This procedure is often called post-Minkowski theory. If one further imposes the slow motion ($v \ll c$) assumption, one obtains post-Newtonian theory. For example, to obtain the equations of motion at first post-Newtonian (1PN) order, *two* iterations are needed (i.e., $h_2^{\alpha\beta}$ must be calculated). To obtain the leading gravitational wave-

form and energy flux far from a binary system, two iterations are needed, while to obtain the leading contributions to gravitational radiation damping of the source, *three* iterations are necessary (see ref. 22 for a discussion).

However, because the source $\tau^{\alpha\beta}$ contains $h^{\alpha\beta}$ itself, it is not confined to a compact region but extends over all spacetime. As a result, as emphasized by Ehlers et al. (17), there is a danger that the integrals involved in various expansions of the solutions for $h^{\alpha\beta}$ will diverge or be ill-defined. Numerous approaches were developed to handle this difficulty. The “post-Minkowski” method of Blanchet, Damour, and Iyer solves Einstein’s equations by two different techniques, one in the near zone (within one gravitational wavelength of the source) and one in the far zone, and uses the method of singular asymptotic matching to join the solutions in an overlap region. The method provides a natural “regularization” technique to control potentially divergent integrals (see ref. 23 for a thorough review of this method). The “Direct Integration of the Relaxed Einstein Equations” (DIRE) approach of Will, Wiseman, and Pati (see ref. 24) retains Eq. 4 as the global solution but splits the integration into one over the near zone and another over the far zone and uses different integration variables to carry out the explicit integrals over the two zones. In the DIRE method, all integrals are finite and convergent.

The problem of “delta-functions” was handled in a variety of ways. One was to import from quantum field theory a set of powerful techniques called “dimensional regularization,” that, subject to some benign assumptions about analyticity, could be used to control the potential infinities order by order in the PN expansion. Another adapted the EIH matching method to higher-order PN calculations. A third approach treated each body as a real, nearly spherical fluid ball and sorted contributions depending on how they scaled with the size of the ball. In every case where different methods made a prediction about the equations of motion or the gravitational-wave signal, they were in complete agreement.

Among the results of these approaches are formulae for the equations of motion and gravitational waveform of binary systems of compact objects, carried out to high orders in a PN expansion. Here we shall only state a few formulae for the purpose of illustration. For example, the relative two-body equation of motion has the form

$$\frac{d\mathbf{v}}{dt} = \frac{Gm}{r^2} \left\{ -\hat{\mathbf{n}} + \frac{1}{c^2} \mathbf{A}_{1\text{PN}} + \frac{1}{c^4} \mathbf{A}_{2\text{PN}} + \frac{1}{c^5} \mathbf{A}_{2.5\text{PN}} + \frac{1}{c^6} \mathbf{A}_{3\text{PN}} + \frac{1}{c^7} \mathbf{A}_{3.5\text{PN}} + \dots \right\}, \quad [5]$$

where $m = m_1 + m_2$ is the total mass, $r = |\mathbf{x}_1 - \mathbf{x}_2|$, $\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2$, and $\hat{\mathbf{n}} = (\mathbf{x}_1 - \mathbf{x}_2)/r$. The notation $\mathbf{A}_{n\text{PN}}$ indicates that the term is $O(c^n)$ relative to the leading Newtonian term $-\hat{\mathbf{n}}$. Explicit and unambiguous formulae for nonspinning bodies through 3.5PN order have been calculated by various authors, and a number of spin-orbit and spin-spin contributions have been obtained (see ref. 23 for a review). Here we quote only the first PN correction and the leading radiation-reaction term at 2.5PN order:

$$\mathbf{A}_{1\text{PN}} = \left\{ (4 + 2\eta) \frac{Gm}{r} - (1 + 3\eta)v^2 + \frac{3}{2}\eta^2 \right\} \hat{\mathbf{n}} + (4 - 2\eta)\dot{r}\mathbf{v}, \quad [6]$$

$$\mathbf{A}_{2.5\text{PN}} = -\frac{8}{15}\eta \frac{Gm}{r} \left\{ \left(9v^2 + 17\frac{Gm}{r} \right) \dot{r}\hat{\mathbf{n}} - \left(3v^2 + 9\frac{Gm}{r} \right) \mathbf{v} \right\}, \quad [7]$$

where $\eta = m_1 m_2 / (m_1 + m_2)^2$ and $\dot{r} = dr/dt$. These terms are

sufficient to analyze the orbit and evolution of binary pulsars. For example, the 1PN terms are responsible for the periastron advance of an eccentric orbit, given by $d\omega/dt = 6\pi f_b Gm / [ac^2(1 - e^2)]$, where a and e are the semimajor axis and eccentricity, respectively, of the orbit, and f_b is the orbital frequency, given to the needed order by Kepler’s third law $2\pi f_b = (Gm/a^3)^{1/2}$.

Another product is a formula for the gravitational field far from the system, whose spatial components h^{ij} (often called the gravitational “waveform”) are sufficient to determine the signal detected by a laser interferometer, written schematically in the form

$$h^{ij}(t, \mathbf{x}) = \frac{2Gm}{Rc^4} \left\{ Q^{ij} + \frac{1}{c} Q_{0.5\text{PN}}^{ij} + \frac{1}{c^2} Q_{1\text{PN}}^{ij} + \frac{1}{c^3} Q_{1.5\text{PN}}^{ij} + \frac{1}{c^4} Q_{2\text{PN}}^{ij} + \frac{1}{c^5} Q_{2.5\text{PN}}^{ij} + \dots \right\}, \quad [8]$$

where R is the distance from the source, and the variables are to be evaluated at retarded time $t - R/c$. The leading term is the so-called quadrupole formula, given explicitly by

$$h^{ij}(t, \mathbf{x}) = \frac{2G}{Rc^4} \ddot{I}^{ij}(t - R/c), \quad [9]$$

where I^{ij} is the quadrupole moment of the source, and overdots denote time derivatives. For a binary system this leads to $Q^{ij} = 2\eta(v^i v^j - Gm\hat{n}^i \hat{n}^j / r)$. For binary systems, explicit formulae for the waveform through 3.5PN order have been derived (see ref. 23 for a full review).

Given the gravitational waveform, one can compute the rate at which energy is carried off by the radiation. The lowest-order quadrupole formula leads to the gravitational-wave energy flux

$$\dot{E} = \frac{8}{15}\eta^2 \frac{G^2 m^4}{r^4 c^5} (12v^2 - 11\dot{r}^2). \quad [10]$$

This has been extended to 3.5PN order beyond the quadrupole formula (23). Formulae for fluxes of angular and linear momentum can also be derived. The 2.5PN radiation-reaction terms in the equation of motion (Eq. 5) result in a decrease of the orbital energy at a rate that precisely balances the energy flux (Eq. 10) determined from the waveform. Averaged over one orbit, this results in a rate of increase of the binary’s orbital frequency given by

$$\frac{df_b}{dt} = \frac{192\pi}{5} f_b^2 \left(\frac{2\pi G \mathcal{M} f_b}{c^3} \right)^{5/3} F(e), \quad [11]$$

$$F(e) = (1 - e^2)^{-7/2} \left(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4 \right),$$

where $\mathcal{M}$ is the so-called “chirp” mass, given by $\mathcal{M} = \eta^{3/5} m$. Notice that by making precise measurements of the phase $\Phi(t) = 2\pi \int^t f(t') dt'$ of either the orbit or the gravitational waves (for which $f = 2f_b$ for the dominant component) as a function of the frequency, one in effect measures the “chirp” mass of the system.

The Parametrized Post-Newtonian Framework and Tests of General Relativity

The post-Newtonian approximation has been remarkably effective as a tool for interpreting experimental tests of general relativity. This is because, in a broad class of alternative metric theories of gravity, it turns out that only the values of a set of numerical coefficients in the post-Newtonian expression for the spacetime metric vary from theory to theory. Thus one can encompass a wide range of alternative theories by simply introdu-

damping of the orbital period is now limited by uncertainties in our knowledge of the relative acceleration between the solar system and the binary system as a result of galactic differential rotation. In the recently discovered “double pulsar” J 0737-3039, all five post-Keplerian parameters are measured, together with the mass ratio m_1/m_2 derived directly from the ability to observe the motion of both pulsars. All six constraints on the masses overlap one another on the m_1 - m_2 plane, again consistently with general relativity.

However, there is something potentially wrong here. All the post-Newtonian effects discussed above are calculated using post-Newtonian theory. Yet the neutron stars that compose these systems have *very* strong internal gravity. This gravitational binding energy reduces the total mass of each body by 10 to 20% compared to the total rest mass of its constituent particles. By contrast, the orbital energy is only 10^{-6} of the mass-energy of the system. Because general relativity is a nonlinear theory, surely there is some mixing between the strong internal gravity and the weak interbody gravity. So how can post-Newtonian theory possibly give valid predictions for such systems?

The reason is a remarkable property of general relativity called the Strong Equivalence Principle (SEP). A consequence of this principle is that the internal structure of a body is “effaced,” so that the orbital motion and gravitational radiation emitted by a system of well separated bodies depend *only* on the total mass of each body and not on its internal structure, apart from standard tidal and spin-coupling effects. In other words, the motion of a normal star or a neutron star or a black hole depends on the body’s total mass and not on the strength of its internal gravitational fields. This behavior was already implicit in the work of Einstein, Infeld, and Hoffman, where only the exterior nearby field of each body was needed, and has been verified theoretically to at least second post-Newtonian order by more modern methods.

By contrast, in alternative theories of gravity, SEP is not valid in general, and internal-structure effects can lead to significantly different behavior, such as the Nordtvedt effect, a possible difference in acceleration of the earth compared to the moon in the solar gravitational field, or such as the emission of dipole gravitational radiation from systems of bodies with dissimilar internal structure. The close agreement of binary-pulsar data with the predictions of general relativity constitutes a kind of “null” test of the effacement of strong-field effects in that theory. It also constitutes a verification of the unreasonable effectiveness of post-Newtonian theory in this class of strong-gravity systems.

Gravitational Waves and Inspiralling Compact Binaries

Possibly the most remarkable example of the unreasonable effectiveness of post-Newtonian theory is that of the inspiral and merger of binary systems of compact objects such as neutron stars and black holes. The decay of the orbit of a compact binary system through gravitational-wave emission will ultimately bring the two bodies together in a final and catastrophic inspiral, followed by a merger and the likely formation of a terminal black hole. This process will emit a characteristic gravitational-wave signal with rising frequency and amplitude (often called a “chirp”) that should be detectable by the worldwide network of ground-based laser-interferometric observatories that is expected to be once again operational following major upgrades, by 2015 (18, 19). In the low-frequency end of the gravitational-wave spectrum, the merger of supermassive black holes in the centers of galaxies will be detectable from cosmological distances by the proposed space-based interferometer LISA, currently being planned for a launch after 2020 (20), and by next-generation arrays of radio telescopes doing pulsar timing (34).

The most effective technique for detecting potential binary inspiral gravitational-wave signals embedded in the noisy output of these interferometers is the method of *matched filtering*, where-

by a theoretically generated gravitational-wave signal appropriate for a given source is cross correlated against the output of the detector. Because the noise is a random process, such a cross-correlation will yield a positive signature if there is a signal that precisely matches the template over the hundreds to thousands of cycles of signal that are expected to lie within the detectable band, even if the signal is formally weaker than the noise. With a bank of template waveforms that depend on the source parameters such as the two masses, spins, sky location, orbital eccentricity, etc., it will be possible both to detect signals and to measure the properties of the source (35). To be most effective, this method requires very accurate theoretical templates. For the inspiral part of the signal, these templates have been calculated by many groups using post-Newtonian theory, with equations of motion (Eq. 5) and gravitational waveforms (Eq. 8) calculated to 3.5PN order beyond the leading terms (23).

Eventually, however, the inspiral will reach the state where the orbital velocities are high and the gravitational fields are strong, so that the post-Newtonian approximation is no longer valid. Given our lack of knowledge of the convergence properties of the approximation, it is not known a priori where this should occur.

Meanwhile, it has become increasingly clear that the signal from the final few inspiral orbits, from the merger of the two bodies, and even from the final vibrations of the newly formed black hole will make important contributions to the detectability of the waves by the interferometers. Luckily, many years after its primitive beginnings in the 1970s, numerical relativity finally reached a stage, following critical breakthroughs in 2005 (36), where researchers could reliably and robustly simulate that final part of the inspiral process.

It therefore came as a complete surprise, when gravitational waveforms from post-Newtonian theory were compared with numerical relativity waveforms for those final orbits, to discover that the agreement was unreasonably good. The amplitudes and phases of the waves calculated by the two methods agreed remarkably well cycle by cycle over many cycles, and this was in a regime where $Gm/rc^2 \sim 0.2$ and $v/c \sim 0.4$, where one had no right to expect the post-Newtonian approximation to be valid, even with many high-order correction terms in the PN formulae (37–39).

This unreasonable agreement was crucial, because it permitted the development of techniques for “stitching together” post-Newtonian and numerical relativity waveforms to obtain templates that are accurate and valid over the entire inspiral and merger process. Selecting the best stitching method involves taking into account the noise characteristics of the interferometers whose data is to be analyzed, and to find stitchings that optimize all the data analysis protocols, such as false alarm thresholds, detection confidence criteria, and so on, that are part and parcel of all signal detection strategies. Because the post-Newtonian waveforms are analytic expansions, they can be resummed using such tricks as Padé approximants to suggest alternative ways to match numerical results. This ongoing work involves a unique collaboration among post-Newtonian theorists, numerical relativists, and interferometer data analysts (see, e.g., ref. 40), but it would have been moot had not post-Newtonian theory been so effective in overlapping with numerical relativity in the strong-gravity high-speed regime.

Another example of the unreasonable effectiveness of post-Newtonian theory relates to the “kick” given to a black hole formed from the merger of two compact objects. In contrast to the erroneous 1937 claim by Levi-Civita, this is a real effect. If a system emits gravitational waves anisotropically, then the waves carry linear momentum away in addition to energy and angular momentum, and, by virtue of the overall conservation of momentum, the source must recoil in the opposite direction. It turns out that the gravitational-wave recoil imparted to a final black hole could have important astrophysical consequences,

especially for the mergers of supermassive black holes, possibly ejecting the black hole from the host galaxy, or displacing it sufficiently from the center to cause interactions with surrounding gas or stars, thus generating an electromagnetic counterpart signal (41).

Using formulae for the radiated momentum flux valid to 2PN order, Blanchet et al. (42) calculated the kick imparted to a black hole from the merger of two nonspinning black holes as a function of their mass difference (for equal masses, the effect vanishes by symmetry). The resulting kick velocity turned out to be in remarkable agreement with kicks determined subsequently using numerical relativity, but only up to the point where the two black holes were about to merge. The agreement was surprising, because the dominant contribution to the recoil comes from a regime where the post-Newtonian approximation should have failed.

The numerical simulations also showed that, following the merger, there was a small “antickick,” reducing the final recoil velocity by around 30% (43). Le Tiec et al. (44) then used a hybrid calculation that combined formulae for the metric surrounding two closely spaced black holes accurate to 2PN order as initial data, combined with the method of black hole perturbation theory, to study whether the antickick was produced by the linear momentum radiated during the ringdown phase of the vibrating final black hole. The analysis found that the combination of the 2PN inspiral kick plus the ringdown kick was in agreement with the kicks obtained by numerical relativity (Fig. 2).

A final instance of the effectiveness of post-Newtonian theory came from analyses of the initial configurations used in numerical relativity to study compact binary inspiral. Because gravitational radiation tends to circularize binary orbits over time, it is natural to assume that at late times, the inspiralling binary is in a “quasicircular” orbit—that is, an orbit that is circular, apart from the slow shrinkage due to gravitational-wave damping. Using numerical relativity it was possible to solve the so-called initial value equations of Einstein’s theory for such quasicircular orbits for a variety of systems, including double black holes, double neutron stars, or mixed systems, with and without spin. These solutions yielded a set of accurate values for the orbital energy E and angular momentum J as a function of the orbital angular velocity Ω . In Newtonian gravity (which would apply to widely separated binaries), these variables would be given by $E = -\eta m(Gm\Omega)^{2/3}/2$ and $J = G\eta m^2(Gm\Omega)^{-1/3}$. But the systems in these simulations were highly relativistic, corresponding to $Gm/rc^2 \sim 0.1$ and $v/c \sim 0.3$. Using post-Newtonian expressions for $E(\Omega)$ and $J(\Omega)$ valid to 3PN order, it was found that the agreement between the PN and numerical results was remarkably good, at the level of several percent in most cases (45, 46). In fact it was suggested that some of the systematic differences between the PN and the numerical results could be explained if the numerical initial configurations actually corresponded to slightly eccentric orbits (46–48). This would change the relation between E , J and Ω . Because one is in a highly relativistic regime, it is not obvious in solving Einstein’s initial value equations numerically how to choose the initial separation and initial angular velocity so as to guarantee a circular orbit initially. Subsequently it was discovered that the numerical evolution of such initial orbits forward in time *did* indeed display small amounts of eccentricity. Numerical methods were then developed to fine-tune the initial configurations to ensure the desired amount of initial eccentricity.

Fig. 2. Kick velocity vs. reduced mass parameter η for nonspinning black hole mergers. The kick vanishes for equal masses ($\eta = 0.25$). Dashed lines indicate estimated uncertainties in numerical relativity kicks, while error bars indicate estimated uncertainties in the PN-ringdown kicks.

In a related development, Favata recently pointed out the remarkable effectiveness of the PN approximation for determining the final stable circular orbit of two black holes (49).

Concluding Remarks

Wigner remarked that the effectiveness of mathematics in the natural sciences was mysterious. The unreasonable effectiveness of the post-Newtonian approximation in gravitational physics is no less mysterious. There is no obvious reason to expect PN theory to account so well for the late stage of inspiral and merger of two black holes. The SEP of general relativity undoubtedly plays a role, by making the internal structure of the bodies irrelevant until they begin to distort one another tidally. But it does not explain why PN waveforms should agree so well with numerical waveforms when the orbital velocities are almost half the speed of light or why recoil velocities calculated using PN methods should agree so well with those from numerical methods. Our colleague Robert V. Wagoner once speculated during the 1970s that, because the gravitational redshift effect makes processes near black holes appear slower and “weaker” from the point of view of external observers, the PN approximation should somehow work better than expected, even under such extreme conditions. Because of the redshift effect, “strong” gravity is not as strong as one might think. But nobody has been able to translate Wagoner’s musing into anything quantitative or predictive. And yet the unreasonable effectiveness of post-Newtonian theory will likely be an important factor in the anticipated first detection of gravitational waves.

ACKNOWLEDGMENTS. We are grateful to the Gravitation and Cosmology Group (GReCo) of the Institut d’Astrophysique de Paris for their hospitality during the preparation of this paper, and to Eric Poisson, Luc Blanchet, Saul Teukolsky, and Bala Iyer for a critical reading of the manuscript. This work was supported in part by the National Science Foundation, grants PHY06-52448 and PHY09-65133, and the Centre National de la Recherche Scientifique, Programme Internationale de la Coopération Scientifique (CNRS-PICS), Grant 4396. This paper is based in part on a talk given at the Subrahmanyan Chandrasekhar Centennial Symposium at the University of Chicago, October 15–17, 2010.

1. Wigner E (1960) The unreasonable effectiveness of mathematics in the natural sciences. *Commun Pur Appl Math* 13(1):1–14.
2. Damour T (1987) *The Problem of Motion in Newtonian and Einsteinian Gravity*. 300 Years of Gravitation, eds SW Hawking and W Israel (Cambridge Univ Press, Cambridge), pp 128–198.
3. de Sitter W (1916) On Einstein’s theory of gravitation and its astronomical consequences. Second paper. *Mon Not R Astron Soc* 77:155–184.
4. Lorentz HA, Droste J (1917) *Versl K Akad Wet Amsterdam* 26:392, 649.

5. Einstein A (1916) Approximate integration of the field equations of gravitation. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften (Berlin)* 688–696.
6. Eddington AS (1922) *The Mathematical Theory of Relativity* (Cambridge Univ Press, Cambridge).
7. Einstein A, Infeld L, Hoffmann B (1938) The gravitational equations and the problem of motion. *Ann Math* 39:65–100.
8. Levi-Civita T (1937) Astronomical consequences of the relativistic two-body problem. *Am J Math* 59:225–334.

9. Eddington AS, Clark GL (1938) The problem of n bodies in general relativity. *P R Soc London* 166:465–75.
10. Robertson HP (1938) The two-body problem in general relativity. *Ann Math* 39:101–104.
11. Fock VA (1964) *The Theory of Space, Time and Gravitation* (MacMillan, New York).
12. Chandrasekhar S (1964) The post-Newtonian equations of hydrodynamics in general relativity. *Astrophys J* 142:1488–1512.
13. Nordtvedt K, Jr (1968) Equivalence principle for massive bodies. II. Theory. *Phys Rev* 169:1017–1025.
14. Hulse RA, Taylor JH (1975) Discovery of a pulsar in a binary system. *Astrophys J Lett* 195:L51–L53.
15. Smarr L, Čadež A, Dewitt B, Eppley K (1976) Collision of two black holes: Theoretical framework. *Phys Rev D* 14:2443–2452.
16. Hahn SG, Lindquist RW (1964) The two-body problem in geometrodynamics. *Ann Phys-New York* 29:304–331.
17. Ehlers J, Rosenblum A, Goldberg JN, Havas P (1976) Comments on gravitational radiation damping and energy loss in binary systems. *Astrophys J Lett* 208:L77–L81.
18. LIGO Scientific Collaboration: Abbott B, et al. (2008) LIGO: The laser interferometer gravitational-wave observatory. *Rep Prog Phys* 72 07690119.
19. Virgo Collaboration: Acernese F, et al. (2006) The status of Virgo. *Classical Quant Grav* 23:S63–S70.
20. Danzmann K (1997) LISA—an ESA cornerstone mission for a gravitational-wave observatory. *Classic Quant Grav* 14:1399–1404.
21. Misner CW, Thorne KS, Wheeler JA (1973) *Gravitation* (Freeman, San Francisco).
22. Walker M, Will CM (1980) The approximation of radiative effects in relativistic gravity: Gravitational radiation reaction and energy loss in nearly Newtonian systems. *Astrophys J Lett* 242:L129–133.
23. Blanchet L (2006) Gravitational radiation from post-Newtonian sources and inspiralling compact binaries. *Living Rev Relativ* 9:4, (cited January 1, 2011) <http://www.livingreviews.org/lrr-2006-4>.
24. Pati ME, Will CM (2000) Post-Newtonian gravitational radiation and equations of motion via direct integration of the relaxed Einstein equations: Foundations. *Phys Rev D* 62:124015–1–28.
25. Nordtvedt KL, Jr (1968) Equivalence principle for massive bodies. II. Theory. *Phys Rev* 169:1017–1025.
26. Will CM (1971) Theoretical frameworks for testing relativistic gravity. II. Parametrized post-Newtonian hydrodynamics and the Nordtvedt effect. *Astrophys J* 163:611–628.
27. Will CM, Nordtvedt KL, Jr (1972) Conservation laws and preferred frames in relativistic gravity. I. Preferred-frame theories and an extended PPN formalism. *Astrophys J* 177:757–774.
28. Will CM (1993) *Theory and Experiment in Gravitational Physics* (Cambridge Univ Press, Cambridge).
29. Will CM (1988) Henry Cavendish, Johann von Soldner, and the deflection of light. *Am J Phys* 56:413–415.
30. Shapiro SS, Davis JL, Lebach DE, Gregory JS (2004) Measurement of the solar gravitational deflection of radio waves using geodetic very-long-baseline interferometry data, 1979–1999. *Phys Rev Lett* 92:121101–1–4.
31. Bertotti B, Iess L, Tortora P (2003) A test of general relativity using radio links with the Cassini spacecraft. *Nature* 425:374–376.
32. Will CM (2006) The confrontation between general relativity and experiment. *Living Rev Relativ* 9:3, (cited January 1, 2011) <http://www.livingreviews.org/lrr-2006-3>.
33. Weisberg JM, Nice DJ, Taylor JH (2010) Timing measurements of the relativistic binary pulsar PSR B1913+16. *Astrophys J* 722:1030–1034.
34. Hobbs G, et al. (2010) The international pulsar timing array project: Using pulsars as a gravitational wave detector. *Classical Quant Grav* 27:084013–1–10.
35. Cutler C, et al. (1993) The last three minutes: Issues in gravitational-wave measurements of coalescing compact binaries. *Phys Rev Lett* 70:2984–2987.
36. Pretorius F (2005) Evolution of binary black hole spacetimes. *Phys Rev Lett* 95:121101–1–4.
37. Baker JG, van Meter JR, McWilliams ST, Centrella J, Kelly BJ (2007) Consistency of post-Newtonian waveforms with numerical relativity. *Phys Rev Lett* 99:181101–1–4.
38. Boyle M, et al. (2007) High-accuracy comparison of numerical relativity simulations with post-Newtonian expansions. *Phys Rev D* 76:124038–1–31.
39. Hannam M, Husa S, González JA, Sperhake U, Brügmann B (2008) Where post-Newtonian and numerical-relativity waveforms meet. *Phys Rev D* 77:044020–1–15.
40. Buonanno A, Iyer BR, Ochsner E, Pan Y, Sathyaprakash BS (2009) Comparison of post-Newtonian templates for compact binary inspiral signals in gravitational-wave detectors. *Phys Rev D* 80:084043–1–24.
41. Merritt D, Milosavljevic M, Favata M, Hughes SA, Holz DE (2004) Consequences of gravitational radiation recoil. *Astrophys J Lett* 607:L9–L12.
42. Blanchet L, Qusailah MSS, Will CM (2005) Gravitational recoil of inspiralling black-hole binaries to second post-Newtonian order. *Astrophys J* 635:508–515.
43. Baker JG, et al. (2006) Getting a kick out of numerical relativity. *Astrophys J Lett* 653:L93–L96.
44. Le Tiec A, Blanchet L, Will CM (2010) Gravitational-wave recoil from the ringdown phase of coalescing black hole binaries. *Classical Quant Grav* 27:012001–1–7.
45. Blanchet L (2002) Innermost circular orbit of binary black holes at the third post-Newtonian approximation. *Phys Rev D* 65:124009–1–7.
46. Mora T, Will CM (2002) Numerically generated quasi-equilibrium orbits of black holes: Circular or eccentric? *Phys Rev D* 66:101501–1–5.
47. Mora T, Will CM (2004) Post-Newtonian diagnostic of quasi-equilibrium binary configurations of compact objects. *Phys Rev D* 69:104021–1–25 Erratum- *ibid.* 71:129901(2005).
48. Berti E, Iyer S, Will CM (2008) Post-Newtonian diagnosis of quasidequilibrium configurations of neutron star-neutron star and neutron star-black hole binaries. *Phys Rev D* 77:024019–1–15.
49. Favata M (2011) Conservative corrections to the innermost stable circular orbit (ISCO) of a Kerr black hole: A new gauge-invariant post-Newtonian ISCO condition, and the ISCO shift due to test-particle spin and the gravitational self-force. *Phys Rev D* 83:024028–1–17.