12 Strong-Field and Dynamical Tests of Relativistic Gravity

We now take the theoretical frameworks for discussing compact body dynamics and gravitational radiation that we developed in the preceding two chapters, and apply them to study current and future tests of general relativity in the strong-field and dynamical regimes. This effort began with the discovery of binary pulsars beginning in 1974, and in Section 12.1, we will discuss these remarkable systems and their implications for testing relativistic gravity. The detection of gravitational waves from an inspiraling binary black hole system in 2015 brought to fruition the ability to test gravitational theories in the strong-field dynamical regime using the gravitational waves themselves. In Section 12.2 we will describe how information about the sources of the waves and about the underlying theory can be extracted from the detected signal, and will discuss the tests that have been carried out since 2015. Another kind of test of strong-field gravity involves exploring the spacetime near otherwise stationary compact objects using probes such as stars, small black holes and gas. Many of these tests are still works in progress; we will give some examples in Section 12.3. Finally, Section 12.4 will close this chapter with a brief description of cosmological tests.

12.1 Binary Pulsars

The summer of 1974 was an eventful one for Joseph Taylor and Russell Hulse. Using the Arecibo radio telescope in Puerto Rico, they had spent the time engaged in a systematic survey for new pulsars. During that survey, they detected 50 pulsars, of which 40 were not previously known, and made a variety of observations, including measurements of their pulse periods to an accuracy of one microsecond. But one of these pulsars, denoted B1913+16,[1] was peculiar: besides having a pulsation period of 59 ms, shorter than that of any known pulsar except the one in the Crab Nebula, it also defied any attempts to measure its period to $\pm 1\,\mu$s, by making "apparent period changes of up to $80\,\mu$s, from day to day, and sometimes by as much as $8\,\mu$s, over 5 minutes" (Hulse and Taylor, 1975). Such behavior is uncharacteristic of pulsars, and Hulse and Taylor rapidly concluded that the observed period changes were the result of Doppler shifts due to orbital motion of the pulsar about a companion (for a popular account of the day-by-day detective work

[1] At the time of the discovery, pulsars were denoted PSR, followed by the right ascension and declination. PSR was ultimately replaced by "B", for pulsars whose coordinates were referred to the 1950.0 epoch. Most pulsars discovered since 1993 are labeled with "J", have coordinates referred to the 2000.0 epoch, and include minutes in the declination. We will use the current conventions.

involved see Will (1986)). By the end of September, 1974, Hulse and Taylor had obtained an accurate "velocity curve" of this "single line spectroscopic binary." The velocity curve was a plot of apparent period of the pulsar as a function of time. By a detailed fit of this curve to a Keplerian two-body orbit, they obtained the following elements of the orbit of the system: K_1, the semiamplitude of the variation of the radial velocity of the pulsar; P_b, the period of the binary orbit; e, the eccentricity of the orbit; ω_0, the longitude of periastron at a chosen epoch (September 1974); $a_1 \sin \iota$, the projected semimajor axis of the pulsar orbit, where ι is the inclination of the orbit relative to the plane of the sky; and $f_1 = (m_2 \sin \iota)^3/(m_1+m_2)^2$, the mass function, where m_1 and m_2 are the mass of the pulsar and companion, respectively. In addition, they obtained the "rest" period P_p of the pulsar, corrected for orbital Doppler shifts at a chosen epoch.

However, at the end of September 1974, the observers switched to an observation technique that yielded vastly improved accuracy (Taylor et al., 1976). That technique measures the absolute arrival times of pulses (as opposed to the period, or the difference between adjacent pulses) and compares those times to arrival times predicted using the best available pulsar and orbit parameters. The parameters are then improved by means of a least-squares analysis of the arrival-time residuals. With this method, it proved possible to keep track of the precise phase of the pulsar over intervals as long as six months between observations. This was partially responsible for the improvement in accuracy. Other improvements over the years included the sophisticated use of multi-wavelength data to suppress the effects of interstellar dispersion on the radio signals, the use of GPS time transfer to improve the timing precision, and a major upgrade of instrumentation at the Arecibo radio telescope between 1993 and 1997. The results of these analyses using data reported in 2016 are shown in column 3 of Table 12.1 (Weisberg and Huang, 2016); in Section 12.1.1 we will define the Keplerian and post-Keplerian parameters.

The discovery of B1913+16 caused considerable excitement in the relativity community, because it was realized that the system could provide a new laboratory for studying relativistic gravity. Post-Newtonian orbital effects would have magnitudes of order $v_1^2 \sim K_1^2 \sim 5 \times 10^{-7}$, or $m/r \sim f_1/a_1 \sin \iota \sim 3 \times 10^{-7}$, a factor of ten larger than the corresponding quantities for Mercury's orbit, and the shortness of the orbital period ($\sim$ 8 hours) would amplify any secular effect such as the periastron shift. This expectation was confirmed by the announcement in December 1974 (Taylor, 1975) that the periastron shift had been measured to be $4.0 \pm 1.5\,^\circ\,\mathrm{yr}^{-1}$ (compare with Mercury's 43 arcseconds per century!), implying (see below) that the total mass of the system was about $2.8\,M_\odot$. Moreover, the system appeared to be a "clean" laboratory, unaffected by complex astrophysical processes such as mass transfer.

The pulsar radio signal was never eclipsed by the companion, placing limits on the geometrical size of the companion, and the dispersion of the pulsed radio signal showed little change over an orbit, indicating an absence of dense plasma in the system, as would occur if there were mass transfer from the companion onto the pulsar. These data effectively ruled out a main-sequence star as a companion: although such a star could conceivably fit the geometrical constraints placed by the eclipse and dispersion measurements, it would produce an enormous periastron shift ($>5000\,^\circ\,\mathrm{yr}^{-1}$) generated by tidal deformations due to the pulsar's gravitational field. Other candidates for the

Table 12.1 Arrival-time parameters for binary pulsars, and their values in B1913+16. Numbers in parentheses denote errors in the last digit. Data taken from Weisberg and Huang (2016).

Parameter	Symbol (units)	Value in B1913+16
(i) Astrometric and pulsar parameters		
Right ascension	α	$19^{\rm h}15^{\rm m}27.^{\rm s}99942(3)$
Declination	δ	$16^\circ 06' 27.''3868(5)$
Pulsar period	$P_{\rm p}$ (ms)	59.030003217813(11)
Derivative of period	$\dot{P}_{\rm p}$	$8.6183(3) \times 10^{-18}$
(ii) Keplerian parameters		
Projected semimajor axis	$a_1 \sin\iota$ (s)	2.341776(2)
Eccentricity	e	0.6171340(4)
Orbital period	$P_{\rm b}$ (day)	0.322997448918(3)
Longitude of periastron	$\omega_0(^\circ)$	292.54450(8)
Julian date of periastron passage	T_0 (MJD)	52144.90097849(3)
(iii) Post-Keplerian parameters		
Mean rate of periastron advance	$\langle\dot{\omega}\rangle$ ($^\circ$ yr^{-1})	4.226585(4)
Redshift/time dilation	γ' (ms)	4.307(4)
Derivative of orbital period	$\dot{P}_{\rm b}$ (10^{-12})	$-2.423(1)$
Range of Shapiro delay	$r\,(\mu{\rm s})$	$9.6\,^{+2.7}_{-3.5}$
Shape of Shapiro delay	$s = \sin\iota$	$0.68\,^{+0.10}_{-0.06}$

companion that were considered early on were a helium main-sequence star, a white dwarf, a neutron star and a black hole. Any of these would be consistent with the evolutionary models for binary systems of two massive stars that were popular at the time. In these models, one massive star evolves more rapidly, undergoing a supernova explosion and leaving a neutron star remnant. Mass transfer from the companion star serves to spin up the neutron star to its present 59 ms rotation period (this is the so-called pulsar recycling model, believed to be responsible for the class of millisecond pulsars). Subsequently, the massive companion star evolved rapidly, possibly undergoing its own explosion, leaving one of the four remnants listed earlier.

The first two companion candidates, the helium star and the white dwarf, fell out of favor because of the complete absence of evidence for orbital perturbations that would be induced by tidal or rotational deformation effects. A black hole companion was disfavored when the observed pericenter advance indicated that the total mass of the system was about $2.8\,M_\odot$; most evolutionary scenarios leading to black holes, and data on X-ray binaries containing black holes suggest that black holes in such systems are significantly more massive than $1.4\,M_\odot$. A neutron star remnant from a supernova explosion of the companion star was viewed as the most compatible with the near equality of the two masses, with the large orbital eccentricity (mass loss from the second supernova explosion can easily lead to disruption of the system), and with the utter orbital "cleanliness" of the system. In this scenario, the companion neutron star is the younger of the two. No evidence was

ever found for pulsed radiation from the companion, so it is either a pulsar whose signal does not intersect the Earth, or more likely is a dead pulsar, having spun down to a stage where it is no longer able to generate the magnetic fields required to emit radio waves. The main pulsar is an old, weakly magnetized, recycled pulsar, with very weak radio emission (indeed it was barely above the detection threshold set by Hulse in 1974), and a very low spin-down rate.

Within weeks of the discovery, it was recognized that the system might be a testing ground for gravitational radiation damping (Wagoner, 1975; Damour and Ruffini, 1974). The observable effect of this damping is a secular decrease in the period of the orbit (see Section 11.5). However, the timescale for this change, according to general relativity, is so long ($\sim 10^9$ yr) that it was thought that 10–15 years of data would be needed to detect it. However, with improved data acquisition equipment and continued ability to "keep in phase" with the pulsar with the arrival-time method, Taylor and his collaborators surpassed all expectations, and in December 1978 they announced a measurement of the rate of change of the orbital period in an amount consistent with the prediction of gravitational radiation damping in general relativity (Taylor et al., 1979; Taylor and McCulloch, 1980). In 1993, Hulse and Taylor were awarded the Nobel Prize in Physics for this discovery.

In 1990, two more binary pulsars were detected, one eerily similar to B1913+16 in its orbital parameters. Today, more than 220 binary pulsars have been detected, of which 70 have orbital periods shorter than one day (the shorter the period, the more relativistic the system). These include the famous "double pulsar," where, for a time, pulses from *both* neutron stars were also detected; several pulsars with a white dwarf companion; and most recently, a pulsar in a triple system, orbited by two white dwarfs. There are also a number of pulsars with planets revolving around them; while these are fascinating in their own right, they are less interesting for relativity. We will describe the more important denizens of this zoo of binary pulsars and their implications for testing relativistic gravity in Section 12.1.2, but first we turn to the arrival-time analysis for studying these systems.

12.1.1 Arrival-time analysis for binary pulsars

The analysis and interpretation of data from binary pulsars is based on an "arrival-time" framework, originally carried out by Blandford and Teukolsky (1976) and extended by Epstein (1977), Haugan (1985), Damour and Deruelle (1986) and Damour and Taylor (1992). Here we present a simplified version that illustrates the basic parameters that can be measured, and how they can be interpreted in order to test theories of gravity.

We begin by setting up a suitable coordinate system (see Figure 12.1). We choose quasi-Cartesian coordinates $(t, \boldsymbol{x})$ in which the physical metric is of post-Newtonian order everywhere, except in the neighborhood and interior of the pulsar and its companion (if the latter is also compact), and is asymptotically flat. The origin of the coordinate system coincides with a suitably chosen center of mass of the binary system, assumed to be at rest with respect to the solar system. The reference $X-Y$ plane is the plane of the sky, with the X-axis conventionally pointing toward the North celestial pole. The observer is chosen to be at great distance on the negative Z-axis. The orbit of the pulsar is characterized by the standard osculating orbit elements described in Section 6.4.2, and illustrated in

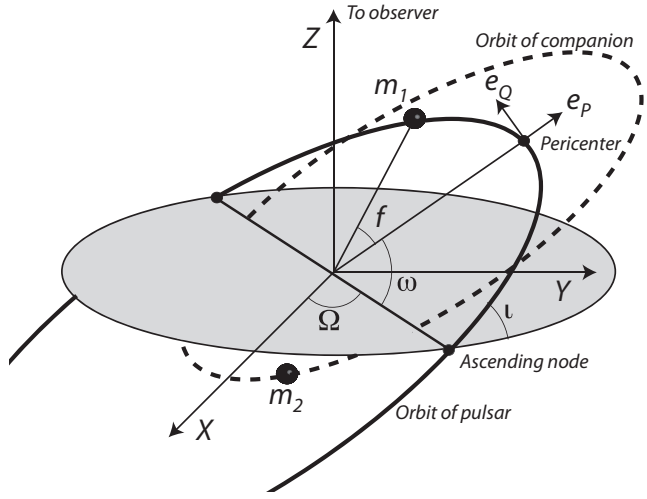

Fig. 12.1 Orbits of a binary pulsar system.

Figure 12.1. However, instead of using the definitions of Eq. (6.73) based on the true anomaly f, we use definitions based on the so-called "eccentric anomaly" u (see PW, section 3.2.4 for detailed discussion), which is more useful in an arrival-time analysis. It is directly related to time by the equation

$$u - e\sin u = n(t - T_0)\,, \tag{12.1}$$

where n is the "mean anomaly," defined by $n \equiv (\mathcal{G}m/a^3)^{1/2} = 2\pi/P_{\rm b}$, where a is the semimajor axis, $P_{\rm b}$ is the orbital period of the binary, and T_0 is the time of pericenter passage. We include the modified EIH parameter $\mathcal{G}$, because that parameter (including possible contributions from the sensitivities of the compact bodies) will govern the quasi-Newtonian limit used to describe the osculating elements in a given theory. The true and eccentric anomalies are related by the equations

$$\cos f = \frac{\cos u - e}{1 - e\cos u}\,, \quad \sin f = \frac{\sqrt{1-e^2}\sin u}{1 - e\cos u}\,. \tag{12.2}$$

With the center of mass $\boldsymbol{X} = (m_1\boldsymbol{x}_1 + m_2\boldsymbol{x}_2)/m$ chosen to be at rest at the origin, the orbits of the pulsar (1) and companion (2) are given by

$$\boldsymbol{x}_1 = \frac{m_2}{m}\boldsymbol{x}\,, \quad \boldsymbol{x}_2 = -\frac{m_1}{m}\boldsymbol{x}\,, \tag{12.3}$$

where $\boldsymbol{x} = \boldsymbol{x}_1 - \boldsymbol{x}_2$ is given by

$$\boldsymbol{x} = a\left[(\cos u - e)\boldsymbol{e}_{\rm P} + \sqrt{1-e^2}\,\sin u\,\boldsymbol{e}_{\rm Q}\right], \tag{12.4}$$

where $\boldsymbol{e}_{\rm P}$ and $\boldsymbol{e}_{\rm Q}$ are unit vectors defined by Eq. (8.29); they point in the direction of the pericenter of body 1 and perpendicular to that vector within the orbital plane (Figure 12.1). Note that the unit vector in the direction of the orbital angular momentum is given by $\hat{\boldsymbol{h}} = \boldsymbol{e}_{\rm P} \times \boldsymbol{e}_{\rm Q}$. The distance between the two bodies is given by $r = a(1 - e\cos u)$.

Any perturbation of the two-body system will induce periodic and secular changes in the orbit elements that can be calculated using the Lagrange planetary equations (6.74).

We next consider the emission of the radio signals by the pulsar. Let τ be proper time as measured by a hypothetical clock in an inertial frame that is momentarily comoving with respect to the pulsar. The time of the Nth rotation of the pulsar is given in terms of the rotation frequency ν of the neutron star by

$$N = N_0 + \nu\tau + \frac{1}{2}\dot{\nu}\tau^2 + \frac{1}{6}\ddot{\nu}\tau^3 + \dots , \tag{12.5}$$

where N_0 is an arbitrary integer, and $\dot{\nu} = d\nu/d\tau|_{\tau=0}$, $\ddot{\nu} = d^2\nu/d\tau^2|_{\tau=0}$. We will ignore the possibility of discontinuous jumps, or "glitches," in the rotation frequency of the pulsar, or of drifts of the phase of the radio beam with respect to the neutron star's rotation phase. These are complicated issues that are still not fully understood. We ultimately wish to determine the arrival time of the Nth pulse on Earth.

Outside the pulsar and its companion, the metric in our chosen coordinate system is given by the modified EIH metric, Eq. (10.28). Because we are interested in the propagation of the radio signal away from the system, we will ignore the possibility of large strong-field corrections to the metric in the close vicinity of the pulsar or its companion. The main result of such effects will be either to add a constant term in the emission time formula (12.5) that can be absorbed into the arbitrary value of N_0, or to multiply terms by a constant factor, such as the redshift at the surface of the neutron star, that can be absorbed into the unknown intrinsic value of ν. Modulo such factors, proper time τ in the local comoving inertial frame, evaluated at the pulsar's center of mass, can be related to coordinate time t by

$$d\tau = dt\left[1 - \alpha_2^*\frac{m_2}{r} - \frac{1}{2}v_1^2 + O(\epsilon^2)\right] , \tag{12.6}$$

where we have dropped the constant contribution of the pulsar's gravitational potential at the point in the inertial frame chosen to be the point of "emission" of the signal, $\alpha_1^* m_1/|\boldsymbol{x}_{\rm em} - \boldsymbol{x}_1|$, and have ignored the difference between the velocity of the emission point and the pulsar's center of mass $\boldsymbol{x}_1$. The two correction terms in Eq. (12.6) are the gravitational redshift and the second-order Doppler shift, or time dilation.

Using Eqs. (12.1) and (12.4), it is simple to show that $v_1^2 = \mathcal{G}(m_2^2/m)(2/r - 1/a)$, and thus that

$$\frac{d\tau}{dt} = 1 - \frac{m_2}{a}\frac{\alpha_2^* + \mathcal{G}m_2/m}{1 - e\cos u} + \frac{\mathcal{G}m_2^2}{2ma} . \tag{12.7}$$

Integrating with respect to t, we obtain

$$\tau = \left[1 - \frac{m_2}{a}\left(\alpha_2^* + \mathcal{G}\frac{m_2}{2m}\right)\right]t_e - \frac{m_2}{a}\left(\alpha_2^* + \mathcal{G}\frac{m_2}{m}\right)\frac{P_{\rm b}}{2\pi}e\,(\sin u - \sin u_0) , \tag{12.8}$$

where u and u_0 are the values of the eccentric anomaly at $t = t_e$ and $t = 0$, respectively. The factor multiplying t_e can be absorbed into the definition of the frequency ν, and the constant term involving u_0 can be dropped. Although these constants may actually undergo secular or periodic variations in time due to orbital perturbations or other effects, these will

be small perturbations of post-Newtonian corrections, and will have negligible effect. Thus we obtain

$$\tau = t_e - \frac{m_2}{a}\left(\alpha_2^* + \mathcal{G}\frac{m_2}{m}\right)\frac{P_{\rm b}}{2\pi}e\sin u\,. \tag{12.9}$$

After emission, the pulsar signal travels along a null geodesic. We can therefore use the method in Sections 6.1 and 7.2 to calculate the coordinate time taken for the signal to travel from the pulsar to the solar system barycenter $\boldsymbol{x}_0$, with the result

$$t_{\rm arr} - t_e = |\boldsymbol{x}_0(t_{\rm arr}) - \boldsymbol{x}_1(t_e)| + (\alpha_2^* + \gamma_2^*)m_2 \ln\left[\frac{2r_0(t_{\rm arr})}{r(t_e) + \boldsymbol{n}\cdot\boldsymbol{x}(t_e)}\right], \tag{12.10}$$

where $r_0 = |\boldsymbol{x}_0|$, $\boldsymbol{n} = \boldsymbol{x}_0/r_0$ and we have used the fact that $r_0 \gg r_1$. The second term in Eq. (12.10) is the Shapiro time delay of the pulsar signal in the gravitational field of the companion. The time delay due to the pulsar's own field is constant to the required accuracy and has been dropped. The effect of the companion's motion during the propagation of the signal across the orbit is a higher post-Newtonian effect, and thus has been ignored.

In practice, one must take into account the fact that the measured arrival time is that at the Earth and not at the barycenter of the solar system, and will therefore be affected by the Earth's position in its orbit and by its own gravitational redshift and Doppler-shift corrections. In fact, it is the effect of the Earth's orbital position on the arrival times that permits accurate determinations of the pulsar's position on the sky. It is also necessary to take into account the effects of interstellar dispersion on the radio signal. These effects can be handled in a standard manner and will not be treated here.

Now, because $r_0 \gg r$, we may write

$$|\boldsymbol{x}_0(t_{\rm arr}) - \boldsymbol{x}_1(t_e)| = r_0(t_{\rm arr}) - \boldsymbol{x}_1(t_e)\cdot\boldsymbol{n} + O(r/r_0)\,. \tag{12.11}$$

Combining Eqs. (12.10) and (12.11), resetting the arrival time by the constant offset r_0, evaluating r and $\boldsymbol{x}_1$ at $t_{\rm arr}$ instead of at t_e, and dropping higher-order terms, we obtain

$$t_e = t_{\rm arr} + (\boldsymbol{x}_1(t_{\rm arr})\cdot\boldsymbol{n})\,(1 + \boldsymbol{v}_1(t_{\rm arr})\cdot\boldsymbol{n}) - \Delta t_{\rm S}(t_{\rm arr})\,, \tag{12.12}$$

where $\Delta t_{\rm S}$ is the Shapiro term in Eq. (12.10).

We can now combine Eqs. (12.9) and (12.12) and substitute the result into Eq. (12.5) to relate $t_{\rm arr}$ to the pulse number N. Here we must include another effect that can occur in theories of gravity that violate SEP (Eardley, 1975), whereby the local gravitational constant at the location of the pulsar may depend on the gravitational potential of the companion, that is,

$$G_L = G_0\left(1 - \eta_2^*\frac{m_2}{r}\right), \tag{12.13}$$

where η^* is a parameter that depends on the theory, and possibly on the sensitivities of the two bodies. As G_L varies during the orbital motion, the structure of the pulsar, its moment of inertia, and thus its intrinsic rotation rate will vary, according to

$$\frac{\Delta\nu}{\nu} = -\frac{\Delta I}{I} = \kappa_{(1)}\frac{\Delta G_L}{G_L} = -\kappa_{(1)}\eta_2^*\frac{m_2}{r}\,, \tag{12.14}$$

where $\kappa_{(1)}$ is the sensitivity of the moment of inertia of body 1 to variations in G_L. Thus Eq. (12.5) should be rewritten

$$N = N_0 + \nu\tau + \int \Delta\nu d\tau + \frac{1}{2}\dot{\nu}\tau^2 + \frac{1}{6}\ddot{\nu}\tau^3 + \dots , \tag{12.15}$$

where to the necessary order, and modulo constants,

$$\int \Delta\nu d\tau = -\nu\kappa_{(1)}\eta_2^* \frac{m_2}{a}\frac{P_\mathrm{b}}{2\pi} e \sin u , \tag{12.16}$$

and where we have not applied a similar correction to the smaller terms involving $\dot{\nu}$ and $\ddot{\nu}$. Substituting Eqs. (12.9), (12.12), and (12.16) into (12.15), and dropping the constant N_0 we obtain the timing formula

$$t_\mathrm{arr} = \nu^{-1}N - \Delta_\mathrm{R}(u) - \Delta_\mathrm{E}(u) + \Delta_\mathrm{S}(u) - \frac{1}{2}\dot{\nu}\nu^{-3}N^2 - \frac{1}{6}\ddot{\nu}\nu^{-4}N^3 + \dots , \tag{12.17}$$

where the three "delay" terms are given by

$$\begin{aligned} \Delta_\mathrm{R}(u) &= xF(e,\omega,u)\left[1 + x\dot{F}(e,\omega,u)\right] , \\ \Delta_\mathrm{E}(u) &= \gamma' \sin u , \\ \Delta_\mathrm{S}(u) &= 2r \ln\left[1 - e\cos u - sF(e,\omega,u)\right] , \end{aligned} \tag{12.18}$$

where

$$\begin{aligned} x &\equiv a_1 \sin\iota = \frac{m_2}{m} a \sin\iota , \\ \gamma' &\equiv e\, m_2 \left(\frac{P_\mathrm{b}}{2\pi\mathcal{G}m}\right)^{1/3} \left(\alpha_2^* + \mathcal{G}\frac{m_2}{m} + \kappa_{(1)}\eta_2^*\right) , \\ r &\equiv \frac{1}{2}(\alpha_2^* + \gamma_2^*) m_2 , \\ s &\equiv \sin\iota . \end{aligned} \tag{12.19}$$

The function $F(e,\omega,u)$ is given by

$$F(e,\omega,u) \equiv \sin\omega(\cos u - e) + \cos\omega(1-e^2)^{1/2}\sin u , \tag{12.20}$$

where $\dot{F} = dF/dt$, and u is related to t_arr by

$$u - e\sin u = n(t_\mathrm{arr} - T_0) . \tag{12.21}$$

Equation (12.17) gives t_arr in terms of the pulse number N, a set of orbital and relativistic parameters, and the intrinsic spin parameters ν, $\dot{\nu}$ and $\ddot{\nu}$. From an initial guess for the values of these parameters, a prediction for the arrival time of the Nth pulse can be made. The difference between the predicted arrival time and the observed arrival time is then used to correct the parameters using a suitable parameter estimation technique, such as least-squares.

The term Δ_R in Eq. (12.17) is called the "Roemer" delay. It is simply the shift in arrival time caused by the change in the pulsar's location relative to the center of mass. The amplitude is controlled by the parameter x, the projected semimajor axis of the pulsar on

the plane of the sky. The evolution of this term with respect to $t_{\rm arr}$ is governed by the orbital period $P_{\rm b}$, the eccentricity e, the pericenter angle ω and the pericenter time T_0 . Because this is a purely geometric effect coupled with Newtonian gravity, these five parameters are known as the "Keplerian" parameters of the system, as listed in Table 12.1. This is the time-domain version of the standard technique for determining orbits in Newtonian single-line spectroscopic binary systems whereby one measures the Doppler shift of the pulse period or spectral line. Notice that the combination

$$f_1 \equiv x^3 \left(\frac{2\pi}{P_{\rm b}}\right)^2 = \mathcal{G}\frac{(m_2 \sin\iota)^3}{m^2}\,, \tag{12.22}$$

is the standard "mass function" of spectroscopic binaries.

The Roemer delay is the dominant correction, of order $r/\tau \sim v \sim \epsilon^{1/2}$. The Einstein and Shapiro terms are of order ϵ. Thus, if the orbit elements a, e, ω, ι experience any secular changes, they will be observable in the Roemer term, if at all (a change in the angle of nodes Ω will be unobservable, as it merely rotates the system about the line of sight). It is conventional to define

$$\begin{aligned} \omega &= \omega_0 + \langle\dot\omega\rangle(t-t_0) + \dots\,, \\ P_{\rm b} &= P_{\rm b0} + \frac{1}{2}\dot P_{\rm b}(t-t_0) + \dots\,, \end{aligned} \tag{12.23}$$

where ω_0 and $P_{\rm b0}$ and their time derivatives are defined at a chosen epoch t_0. Note that the factor $1/2$ in the expression for $P_{\rm b}$ comes from the formal definition of $P_{\rm b}$ in terms of the semimajor axis a. One then substitutes Eqs. (12.23) into $\Delta_{\rm R}$ and includes $\dot\omega$ and $\dot P_{\rm b}$ as parameters to be estimated. The parameters $\dot\omega$ and $\dot P_{\rm b}$ are two of the post-Keplerian parameters listed in Table 12.1.

In principle one could include secular variations in e and ι; such variations have recently been detected in B1913+16 (Weisberg and Huang, 2016), resulting from a spin-orbit induced precession of the orbital plane. Furthermore, as we saw in Sections 8.2 and 8.4, and will reiterate below, it has been possible to search for periodic variations in these parameters in some systems, and to place important limits on violations of SEP.

The term $\Delta_{\rm E}$ is the Einstein term, also called the Redshift/time dilation term, with amplitude γ'. The parameter γ' (not to be confused with the PPN parameter γ) is another post-Keplerian parameter. Note that, once e and $P_{\rm b}$ are determined from the Roemer term, it depends only on the masses of the bodies and possibly on their sensitivities, through α_2^*, $\mathcal{G}$, κ_1, and η_2^*.

The final relativistic term is the Shapiro delay $\Delta_{\rm S}$, dependent upon the "range" parameter r, and the "shape" parameter s. The former depends on the mass of the companion and possibly on sensitivities, while the latter is simply $\sin\iota$.

The final two post-Keplerian parameters in our discussion are $\langle\dot\omega\rangle$ and $\dot P_{\rm b}$, whose predicted forms are given by Eqs. (10.49) and (11.85). Box 12.1 summarizes these predictions for the post-Keplerian parameters.

Box 12.1 **Predictions for post-Keplerian parameters**

Here we summarize the predictions for the post-Keplerian parameters in alternative theories of gravity.

Redshift/time dilation:

$$\gamma' = e\, m_2 \left(\frac{P_\mathrm{b}}{2\pi \mathcal{G} m}\right)^{1/3} \left(\alpha_2^* + \mathcal{G}\frac{m_2}{m} + \kappa_{(1)}\eta_2^*\right) ,$$

Shapiro delay range:

$$r = \frac{1}{2}(\alpha_2^* + \gamma_2^*) m_2 \,.$$

Shapiro delay shape:

$$s = \sin\iota \,.$$

Rate of advance of periastron

$$\langle \dot{\omega} \rangle = \frac{6\pi}{P_\mathrm{b}(1-e^2)} \left(\frac{2\pi \mathcal{G} m}{P_\mathrm{b}}\right)^{2/3} \mathcal{P}\mathcal{G}^{-2} .$$

Derivative of orbital period

$$\dot{P}_\mathrm{b} = -\frac{192\pi}{5} \left(\frac{2\pi \mathcal{G}\mathcal{M}}{P_\mathrm{b}}\right)^{5/3} F(e) - 2\pi \kappa_D \eta \mathcal{S}^2 \left(\frac{2\pi \mathcal{G} m}{P_\mathrm{b}}\right) G(e) \,. \tag{12.24}$$

The coefficients α_2^* and γ_2^* appear in the post-Newtonian metric of body 2 [Eqs. (10.28)], while the parameter η_2^* and the sensitivity $\kappa_{(1)}$ refer to the effect of the field of body 2 on the moment of inertia of body 1 [Eqs. (12.13)–(12.16)]. The quantity $\mathcal{G}$ appears in the quasi Newtonian limit of the modified EIH equations of motion, $\boldsymbol{a} = -\mathcal{G} m \boldsymbol{x}/r^3$, while $\mathcal{P}$ is given by Eq. (10.48). The functions $F(e)$ and $G(e)$ are given by Eq. (11.84); κ_D is the dipole radiation parameter of the theory, and $\mathcal{S}$ is related to the difference in sensitivities between the two bodies.

In general relativity, the post-Keplerian parameters simplify to

$$\begin{aligned}
\gamma' &= e m_2 \left(\frac{P_\mathrm{b}}{2\pi m}\right)^{1/3} \left(1 + \frac{m_2}{m}\right) , \\
r &= m_2 \,, \\
s &= \sin\iota \,, \\
\langle \dot{\omega} \rangle &= \frac{6\pi}{P_\mathrm{b}(1-e^2)} \left(\frac{2\pi m}{P_\mathrm{b}}\right)^{2/3} , \\
\dot{P}_\mathrm{b} &= -\frac{192\pi}{5} \left(\frac{2\pi \mathcal{M}}{P_\mathrm{b}}\right)^{5/3} F(e) \,,
\end{aligned} \tag{12.25}$$

where

$$F(e) = (1-e^2)^{-7/2} \left(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4\right) . \tag{12.26}$$

The formulae for $\dot{P}_{\rm b}$ include gravitational radiation contributions only through quadrupole order. They ignore other sources of energy loss, such as tidal dissipation, mass loss from the system, energy loss from the pulsar emission or via magnetic interactions between the two bodies. To date, there has not been an example of a binary pulsar system where these mechanisms come close to the energy loss via gravitational radiation. If such an system were found, it would very likely be deemed too "dirty" to provide quantitative tests of relativistic gravity, though it might well yield interesting physics of other types.

However there is a contribution to $\dot{P}_{\rm b}$ that cannot be ignored. If the center of mass of the binary system is accelerating relative to that of the solar system, then both the orbital and pulsar periods will change at a rate given by

$$\frac{\dot{P}_{\rm b}}{P_{\rm b}} = \frac{\dot{P}_{\rm p}}{P_{\rm p}} = \ddot{r}_0 = \boldsymbol{a}\cdot\boldsymbol{n} + \frac{1}{r_0}\left[v^2 - (\boldsymbol{v}\cdot\boldsymbol{n})^2\right], \tag{12.27}$$

where $\boldsymbol{v}$ and $\boldsymbol{a}$ are the relative velocity and acceleration, respectively, between the binary system and the solar system. The first term is the projection of the acceleration along the line of sight, while the second, called the Shklovskii effect, represents the effect of variation of the line of sight. If we assume that the binary system (b) and the solar system ($\odot$) are on circular orbits around the galaxy with angular velocities $\Omega_{\rm b}$ and $\Omega_\odot$, distances from the galactic center $r_{\rm b}$ and $r_\odot$, and longitudes relative to the galactic center $\phi_{\rm b}$ and $\phi_\odot$, then Eq. (12.27) takes the form

$$\frac{\dot{P}_{\rm b}}{P_{\rm b}} = \frac{\dot{P}_{\rm p}}{P_{\rm p}} = (\Omega_{\rm b} - \Omega_\odot)^2\,\frac{r_{\rm b} r_\odot}{r_0}\left(\cos\phi - \frac{r_{\rm b} r_\odot}{r_0^2}\sin^2\phi\right), \tag{12.28}$$

where $\phi \equiv \phi_{\rm b} - \phi_\odot$. We will see that this effect is significant for some binary pulsars, such as B1913+16, but not important for others, such as the double pulsar, depending on their distance from the solar system and precise location in the galaxy.

12.1.2 A zoo of binary pulsars

Here we describe some of the most interesting binary pulsar systems from the point of view of testing relativistic gravity. For the most part, we will discuss the implications of the system for general relativity. In Section 12.1.3, we will discuss bounds placed by binary pulsar observations on various alternative theories of gravity.

The Hulse-Taylor binary pulsar B1913+16

The measured values of $\langle\dot{\omega}\rangle$, γ' and $\dot{P}_{\rm b}$ shown in Table 12.1 provide three constraints on the two masses m_1 and m_2, given by Eqs. (12.25) in Box 12.1. It is conventional to plot these three constraints, including their uncertainties, on an $m_1 - m_2$ plane. If general relativity is correct, they must overlap at a single point, within the uncertainties. The plot is shown in the left panel of Figure 12.2. The $\langle\dot{\omega}\rangle$ constraint fixes the total mass of the system to be $2.8284\,M_\odot$, accurate to about a part in 10^6; this is the black line in Figure 12.2, with the width of the line much larger than the actual uncertainty. The γ' constraint fixes the quantity $(m_2/m)(1 + m_2/m)$; this is the dark grey line in Figure 12.2. From the intersection of these

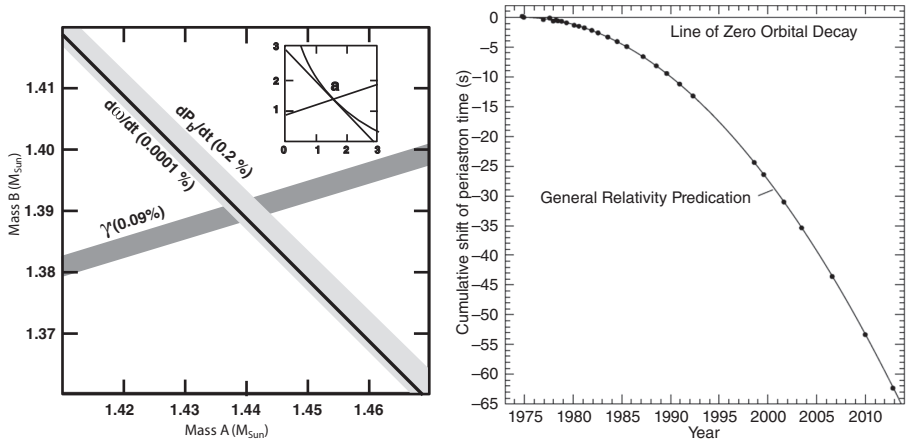

Fig. 12.2 Left: $m_1 - m_2$ plot for B1913+16. Right: Time of periastron passage as a function of time from 1975 to 2013. The curve is the prediction of general relativity for $\dot{P}_\mathrm{b}$ given by Eq. (12.25); the points are the measured values with error bars multiplied by a factor of 400. The large gap in data during the middle 1990s occurred during the upgrade of the Arecibo Radio Telescope. Image reproduced with permission from Weisberg and Huang (2016), copyright by AAS.

two constraints, we obtain the values for the individual masses, $m_1 = 1.438 \pm 0.001\, M_\odot$ and $m_2 = 1.390 \pm 0.001\, M_\odot$ (the error is dominated by the uncertainty in γ').

The $\dot{P}_\mathrm{b}$ constraint fixes the chirp mass of the system $\mathcal{M} = \eta^{3/5} m$; because m is known, this constrains η. However, in this case, the galactic acceleration effect is important. Using data on the location and proper motion of the pulsar, combined with the best information available on galactic rotation; the current value of this effect is $\dot{P}_\mathrm{b}^\mathrm{gal} = -(0.025 \pm 0.004) \times 10^{-12}$. Subtracting this from the measured post-Keplerian parameter $\dot{P}_\mathrm{b}$ shown in Table 12.1 gives the corrected value $\dot{P}_\mathrm{b}^\mathrm{corr} = -(2.398 \pm 0.004) \times 10^{-12}$. This value and its uncertainty are used to plot the dP_b/dt curve in Figure 12.2. It is the hyperbolic curve shown in the inset in Figure 12.2 which ranges from zero to three solar masses, and the light grey band parallel to the $\langle\dot{\omega}\rangle$ constraint in the blow-up of the intersection point "a," which ranges over about $0.05\, M_\odot$. The three curves overlap at a common point, within the uncertainties.

Another way to check the agreement with general relativity is to use the masses inferred from the intersection of the $\langle\dot{\omega}\rangle$ and γ' constraints along with the expression for $\dot{P}_\mathrm{b}$ in Eq. (12.25) to predict the value $\dot{P}_\mathrm{b}^\mathrm{GR} = -(2.40263 \pm 0.00005) \times 10^{-12}$. This agrees with the measured value after the galactic correction, in other words,

$$\frac{\dot{P}_\mathrm{b}^\mathrm{coor}}{\dot{P}_\mathrm{b}^\mathrm{GR}} = 0.9983 \pm 0.0016\,. \tag{12.29}$$

Although the uncertainties in the measured post-Keplerian parameter $\dot{P}_\mathrm{b}$ continue to decrease, in part because of the decrease of statistical errors with observation time T as $T^{-3/2}$, the uncertainties in the parameters that go into the galactic correction are now the limiting factor in the accuracy of the test of gravitational wave damping in general relativity. In fact, if one assumes that general relativity is correct, then the binary pulsar is providing improved data on the galactic rotation curve.

A third way to display the agreement with general relativity is to compare the observed phase of the orbit with a theoretical template phase as a function of time. If $P_{\rm b}$ varies slowly in time, then to first order in a Taylor expansion, the orbital phase is given by $\Phi_{\rm b}(t) = 2\pi t/P_{\rm b0} - \pi \dot{P}_{\rm b0} t^2/P_{\rm b0}^2 + \ldots$. The time of periastron passage T_0 is given by $\Phi(T_0) = 2\pi N$, where N is an integer. Consequently, the periastron time will not grow linearly with N. Thus the cumulative difference between periastron time T_0 and $NP_{\rm b0}$, the quantities actually measured in practice, should vary according to

$$T_0 - NP_{\rm b0} \approx \frac{\dot{P}_{\rm b0}}{2P_{\rm b0}^3} N^2 \approx \frac{1}{2}\frac{\dot{P}_{\rm b0}}{P_{\rm b0}} t^2 \,. \tag{12.30}$$

The right panel of Figure 12.2 shows the comparison between the measured times and the predicted times of periastron passage. Even after a nearly five-year gap in observations while the Arecibo radio telescope was undergoing a major upgrade, the measured periastron times landed right on top of the predicted curve.

The consistency among the constraints provides a test of the assumption that the two bodies behave as "point" masses, without complicated tidal effects, obeying the general relativistic equations of motion including gravitational radiation. This supports the evolutionary model whereby the pulsar is an old recycled pulsar and the companion is a young but dead pulsar. It is also a test of the adherency of general relativity to the SEP in the presence of strong gravity, in that the highly relativistic internal structure of the neutron stars does *not* influence their orbital motion.

Observations indicate that the pulse profile is varying with time (Kramer, 1998; Weisberg and Taylor, 2002), which suggests that the pulsar is undergoing geodetic precession on a 300-year timescale as it moves through the curved spacetime generated by its companion (see Section 9.1.1). The amount is consistent with GR, assuming that the pulsar's spin is suitably misaligned with the orbital angular momentum. Unfortunately, the evidence suggests that the pulsar beam may precess out of our line of sight by 2025.

The precession of the pulsar's spin is accompanied by a precession of the orbital plane, since the total angular momentum of the system is constant up to the changes induced by gravitational radiation reaction. One consequence of this is that the orbital inclination has increased enough that, with the aid of improved measurement accuracy, the Shapiro delay has become measurable (see Table 12.1). Another consequence is that variations in the projected semimajor axis x and eccentricity e must now be included in the analysis, leading to somewhat larger errors in the redshift/time dilation parameters γ' than those presented in earlier analyses (see, e.g., Weisberg et al. (2010)).

The "double" pulsar J0737-3039A,B

This binary pulsar system, discovered by Burgay et al. (2003), was already remarkable for its extraordinarily short orbital period (0.1 days) and large periastron advance ($16.8995^\circ\ {\rm yr}^{-1}$), but then the companion was also detected as a pulsar (Lyne et al., 2004). Because two projected semimajor axes could be measured, the mass ratio was obtained directly and to high precision from the ratio of the values of the Keplerian parameters $x_1 = a_1 \sin\iota$ and $x_2 = a_2 \sin\iota$, since $x_1/x_2 = m_2/m_1$ modulo post-Newtonian corrections

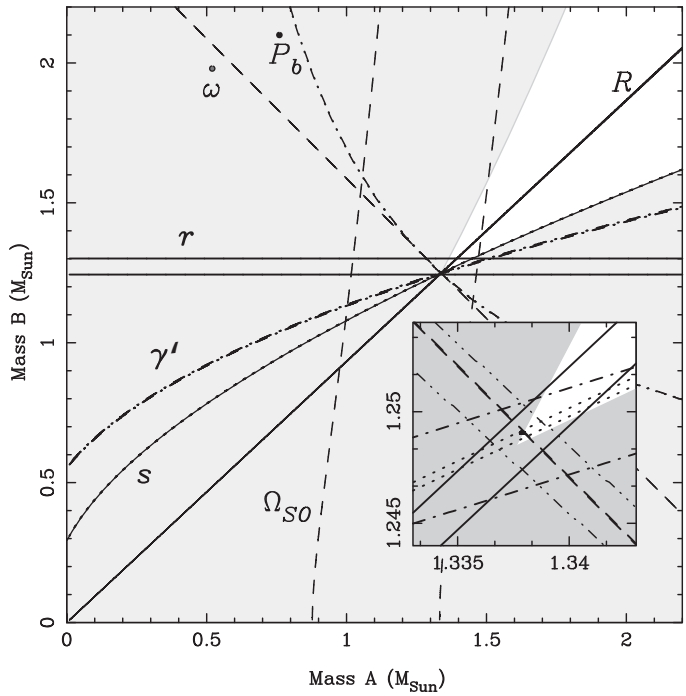

Fig. 12.3 Mass-mass plots for the double pulsar. Image courtesy of Michael Kramer.

(see Table 12.2). In Figure 12.3, the line labeled "R" is that mass ratio, and the white area is the allowed region for the masses, set by the measurement of the two Keplerian mass functions $f_1 = x_1^3(2\pi/P_b)^2$ and $f_2 = x_2^3(2\pi/P_b)^2$ combined with the condition that $|\sin\iota| \leq 1$. Then, the individual masses were obtained by combining the mass ratio with the periastron advance, assuming general relativity to be valid. The results are $m_A = 1.3381 \pm 0.0007\,M_\odot$ and $m_B = 1.2489 \pm 0.0007\,M_\odot$, where A denotes the primary (first) pulsar. From these values, one finds that the orbit is nearly edge-on, with $\sin\iota = 0.9997$, a value which is completely consistent with that inferred from the Shapiro delay shape parameter $s = \sin\iota$. In fact, the five measured post-Keplerian parameters plus the ratio of the projected semimajor axes give six constraints on the masses. As seen in Figure 12.3, all six overlap within their measurement errors (Kramer et al., 2006). The fact that the overlap region, shown in detail in the inset, is so close to the vertex of the allowed region is another indication that the orbit is nearly edge-on. Note that Figure 12.3 is based on more recent data than that quoted in Kramer et al. (2006), described in this discussion and listed in Table 12.2.

Because of the location of the system, galactic proper-motion effects play a significantly smaller role in the interpretation of $\dot{P}_b$ measurements than they did in B1913+16; this and the reduced effect of interstellar dispersion means that the accuracy of measuring the gravitational-wave damping will eventually beat that from the Hulse-Taylor system. It may ultimately be necessary for the data analysis to include 2PN corrections, for example in the periastron advance.

Table 12.2 Parameters of other binary pulsars. Values for $\dot{P}_b$ include corrections for galactic kinematic effects. Numbers in parentheses denote errors in the last digit. See the text for references.

Parameter	J0737–3039(A,B)	J1738+0333	J1141–6545
(i) Keplerian			
$a_1 \sin \iota$ (s)	1.415032(1)/1.516(2)	0.34342913(2)	1.858922(6)
e	0.0877775(9)	$(3.4 \pm 1.1) \times 10^{-7}$	0.171884(2)
P_b (day)	0.10225156248(5)	0.354790739872(1)	0.1976509593(1)
(ii) Post-Keplerian			
$\langle \dot{\omega} \rangle$ ($^\circ$ yr^{-1})	16.8995(7)		5.3096(4)
γ' (ms)	0.386(3)		0.77(1)
$\dot{P}_b$ (10^{-12})	−1.25(2)	−0.0259(32)	−0.401(25)
r (μs)	6.2(3)		
$s = \sin \iota$	0.9997(4)		

The geodetic precession of pulsar B's spin axis has also been measured by monitoring changes in the patterns of eclipses of the signal from pulsar A as it passes through the magnetosphere surrounding pulsar B, with a result in agreement with general relativity to about 13 percent (Breton et al., 2008). The constraint on the masses from that effect, denoted Ω_{SO} is also shown in Figure 12.3. In fact, pulsar B has precessed so much that its signal no longer sweeps by the Earth, so it has gone "silent." For a recent overview of the double pulsar, see Burgay (2012).

J1738+0333: A white-dwarf companion

This is an ultra-low-eccentricity, 8.5-hour period system in which the white-dwarf companion is bright enough to permit detailed spectroscopy, allowing the companion mass to be determined directly to be $0.181\,M_\odot$. The mass ratio is determined from Doppler shifts of the pulsar signal and of the spectral lines of the companion, giving the pulsar mass $1.46\,M_\odot$. Ten years of observation of the system yielded both a measurement of the apparent orbital period decay, and enough information about its parallax and proper motion to account for the substantial galactic effect to give a value of the intrinsic period decay of $\dot{P}_b = (-25.9 \pm 3.2) \times 10^{-15}$ in agreement with the predicted effect of general relativity (Freire et al., 2012). But because of the asymmetry in the sensitivities of the bodies in the system, the result also places a significant bound on the existence of dipole radiation, predicted by many alternative theories of gravity. Data from this system were also used to place a tight bound on the PPN parameter α_1 (see Section 12.1.3 for discussion of these tests of alternative theories).

J1141–6545: A white-dwarf companion

This system is similar in some ways to the Hulse-Taylor binary: short orbital period (0.2 days), significant orbital eccentricity (0.172), rapid periastron advance (5.3° yr^{-1}) and

massive components ($m_p = 1.27 \pm 0.01\,M_\odot$, $m_c = 1.02 \pm 0.01\,M_\odot$). The key difference is that the companion is again a white dwarf. The intrinsic $\dot{P}_\mathrm{b}$ has been measured in agreement with general relativity to about 6 percent, again placing limits on dipole gravitational radiation (Bhat et al., 2008).

J0348+0432: The most massive neutron star

Discovered in 2011 (Lynch et al., 2013; Antoniadis et al., 2013), this is another neutron-star white-dwarf system, in a very short-period (0.1 day), low-eccentricity (2×10^{-6}) orbit. Timing of the neutron star and spectroscopy of the white dwarf have led to mass values of $0.172\,M_\odot$ for the white dwarf and $2.01 \pm 0.04\,M_\odot$ for the pulsar, making it the most massive accurately measured neutron star yet. This supported an earlier discovery of a $2\,M_\odot$ pulsar (Demorest et al., 2010); such large masses rule out a number of heretofore viable soft equations of state for nuclear matter, assuming general relativity to be correct. The orbit period decay agrees with the general relativistic prediction within 20 percent and is expected to improve steadily with time.

J0337+1715: Two white-dwarf companions

This remarkable system was reported by Ransom et al. (2014). It consists of a 2.73 millisecond pulsar of $1.4378(13)\,M_\odot$, with extremely good timing precision, accompanied by two white dwarfs in coplanar circular orbits. The inner white dwarf ($m = 0.19751(15)\,M_\odot$) has an orbital period of 1.629 days, with $e = 6.9177(2) \times 10^{-4}$, and the outer white dwarf ($m = 0.4101(3)\,M_\odot$) has a period of 327.26 days, with $e = 3.5356196(4) \times 10^{-2}$. This is an ideal system for testing the Nordtvedt effect in the strong-field regime. Here the inner system is the analogue of the Earth–Moon system, and the outer white dwarf plays the role of the Sun. Because the outer semimajor axis is about $1/3$ of an astronomical unit, the basic driving perturbation is comparable to that provided by the Sun on the Earth–Moon system. However, the self-gravitational binding energy per unit mass (sensitivity) of the neutron star is almost a billion times larger than that of the Earth, greatly amplifying the potential size of the Nordtvedt effect. In 2018, Archibald et al (Nature, in press) reported a bound of 2.6 parts per million on any difference in acceleration between the neutron star and the white dwarf, representing an improvement over lunar laser ranging of about a factor of 10.

Other binary pulsars

Two of the earliest binary pulsars, B1534+12 and B2127+11C, discovered in 1990, failed to live up to their early promise despite being similar to the Hulse-Taylor system in most respects (both were believed to be double neutron-star systems). The main reason was the significant uncertainty in the kinematic effect on $\dot{P}_\mathrm{b}$ of local accelerations, that of the galaxy in the case of B1534+12, and that of the host globular cluster in the case of B2127+11C.

A class of wide-orbit binary millisecond pulsar (WBMSP) systems, containing a pulsar and a white dwarf in low-eccentricity orbits has been used to place interesting bounds on the Nordtvedt effect. See Section 8.1 for a discussion.

12.1.3 Tests of alternative theories

Soon after the discovery of the binary pulsar B1913+16 it was widely hailed as a new testing ground for relativistic gravitational effects. As we have seen in the case of general relativity, in most respects, the system has lived up to, indeed exceeded, the early expectations. In many ways, the double pulsar topped the Hulse-Taylor binary in this regard.

In another respect, however, B1913+16 only partially lived up to its promise, namely as a direct testing ground for alternative theories of gravity. The origin of this promise was the discovery (Eardley, 1975; Will, 1977) that alternative theories of gravity generically predict the emission of dipole gravitational radiation from binary star systems. As one fulfillment of this promise, Will and Eardley (1977) worked out in detail the effects of dipole gravitational radiation in the bimetric theory of Rosen (1974), and when the first observation of the decrease of the orbital period was announced in 1979, the Rosen theory suffered a fatal blow. A wide class of alternative theories of that period also failed the binary pulsar test because of dipole gravitational radiation (Will, 1977).

On the other hand, the early observations of B1913+16 already indicated that, in general relativity, the masses of the two bodies were nearly equal, so that, in theories of gravity that are in some sense "close" to general relativity, dipole gravitational radiation would not be a strong effect, because of the apparent symmetry of the system. The Rosen theory, and others like it, are not "close" to general relativity, except in their predictions for the weak-field, slow-motion regime of the solar system. When relativistic neutron stars are present, theories like these can predict strong effects on the motion of the bodies resulting from their internal highly relativistic gravitational structure (violations of SEP). As a consequence, the masses inferred from observations of the periastron shift and γ' may be significantly different from those inferred using general relativity, and may be different from each other, leading to strong dipole gravitational radiation damping.

By contrast, the Brans-Dicke theory is close to general relativity, roughly speaking within $1/\omega_{\rm BD}$ of the predictions of the latter, for large values of the coupling constant $\omega_{\rm BD}$. Thus, despite the presence of dipole gravitational radiation, the Hulse-Taylor binary pulsar provides only a weak test of pure Brans-Dicke theory, not competitive with solar-system tests.

However, the discovery of binary pulsar systems with a white dwarf companion, such as J1738+0333, J1141–6545, and J0348+0432 has made it possible to perform strong tests of the existence of dipole radiation. This is because such systems are necessarily asymmetrical, since the gravitational binding energy per unit mass, or sensitivity of white dwarfs is of order 10^{-4}, much less than that of the neutron star. Already, significant bounds have been placed on dipole radiation using J1738+0333 and J1141–6545.

Because the gravitational-radiation and strong-field properties of alternative theories of gravity can be quite different from those of general relativity and each other, it is difficult to parametrize these aspects of the theories in the manner of the PPN framework. In addition, because of the generic violation of the Strong Equivalence Principle in these theories, the results can be very sensitive to the equation of state and mass of the neutron star(s) in the system. In the end, there is no way around having to analyze every theory in turn. On the other hand, because of their relative simplicity, scalar-tensor theories provide

an illustration of the essential effects, and so we will begin by discussing binary pulsars within this class of theories.

Scalar-tensor theories

We combine the results of Section 10.3.2, notably Eqs. (10.48), (10.49), (10.59), and (10.65) to obtain the metric coefficients α^* and γ^* and the periastron advance rate $\langle\dot{\omega}\rangle$. We then combine Eqs. (11.84), (11.85), and (11.121) to obtain $\dot{P}_{\rm b}$. Recalling that $P_{\rm b} = 2\pi(a^3/\mathcal{G}m)^{1/2}$ where

$$\mathcal{G} = 1 - \zeta + \zeta(1 - 2s_1)(1 - 2s_2)\,, \tag{12.31}$$

and $\zeta = (4 + 2\omega_0)^{-1}$, we obtain expressions for the post-Keplerian parameters in scalar-tensor theories:

$$\begin{aligned}
\gamma' &= e\, m_2 \left(\frac{P_{\rm b}}{2\pi\mathcal{G}m}\right)^{1/3} \left(1 - 2\zeta s_2 + \mathcal{G}\frac{m_2}{m} + 2\zeta(1 - 2s_2)(1 + 2\lambda)\kappa_{(1)}\right)\,.\\
r &= m_2(1 - \zeta)\,,\\
s &= \sin\iota\,,\\
\langle\dot{\omega}\rangle &= \frac{6\pi}{P_{\rm b}(1 - e^2)} \left(\frac{2\pi\mathcal{G}m}{P_{\rm b}}\right)^{2/3} \left[1 + \frac{1}{3}\left(2\bar{\gamma} - \bar{\beta}_+ - \Delta\bar{\beta}_-\right)\right]\,,\\
\dot{P}_{\rm b} &= -\frac{192\pi}{5} \left(\frac{2\pi\mathcal{G}\mathcal{M}}{P_{\rm b}}\right)^{5/3} F(e) - 8\pi\zeta\eta\mathcal{S}^2 \left(\frac{2\pi\mathcal{G}m}{P_{\rm b}}\right) G(e)\,.
\end{aligned} \tag{12.32}$$

In the limit $\zeta \to 0$, we recover general relativity, and all structure dependence disappears. The first term in $\dot{P}_{\rm b}$ is the combined effect of quadrupole and monopole gravitational radiation, post-Newtonian corrections to dipole radiation, and a dipole-octupole coupling term, all contributing at 0PN order, while the second term is the effect of dipole radiation, contributing at the dominant -1PN order.

Unfortunately, because of the near equality of neutron star masses in typical double neutron star binary pulsars, dipole radiation is somewhat suppressed, and the bounds obtained are typically not competitive with the Cassini bound on scalar-tensor theories. Figure 12.4 uses the α_0-β_0 parametrization of scalar-tensor theories of Damour and Esposito-Farèse, to display bounds from a variety of systems and experiments. The bounds from the three binary neutron star systems B1913+16, J0737–3039, and B1534+12 are not close to being competitive with the Cassini bound on α_0, or with lunar laser ranging (LLR) bounds, except for those generalized scalar-tensor theories with $\beta_0 < 0$, where the strong gravity of the neutron stars induces spontaneous scalarization effects (Damour and Esposito-Farèse, 1998). Recall that $\alpha_0 = (3 + 2\omega_0)^{-1/2} = [\zeta/(1 - \zeta)]^{1/2}$ and $\beta_0 = 2\omega_0'\phi_0/(3 + 2\omega_0)^2 = 2\lambda/(1 - \zeta)$.

On the other hand, a binary pulsar system with dissimilar objects, such as one with a white dwarf or black hole companion, is a more promising testing ground for dipole radiation. As a result, the neutron-star-white-dwarf systems J1141–6545 and J1738+0333 yield much more stringent bounds. Indeed, the latter system surpasses the Cassini bound for $\beta_0 > 1$ and $\beta_0 < -2$, and is close to that bound for the pure Brans-Dicke case $\beta_0 = 0$ (Freire et al., 2012).

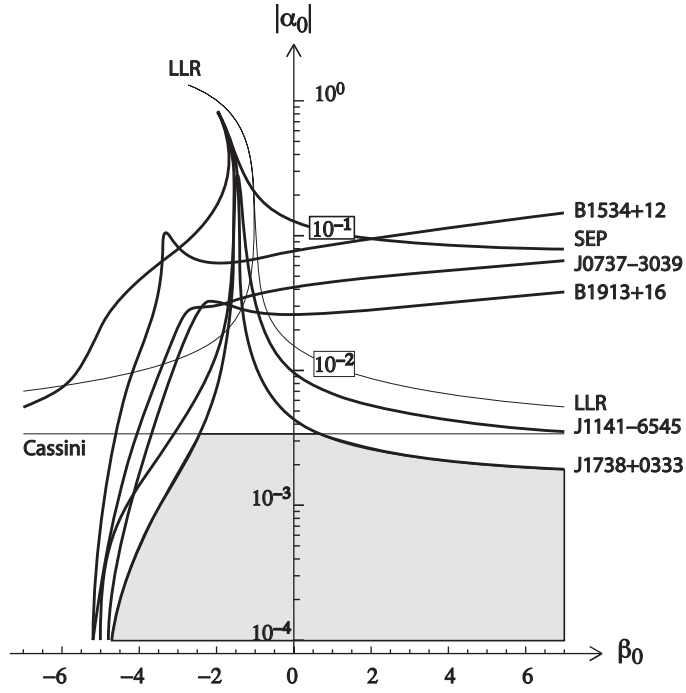

Fig. 12.4 Bounds on scalar-tensor theories from solar-system and binary-pulsar tests. Image reproduced with permission from Freire et al. (2012), copyright by Oxford University Press.

It is worth pointing out that the bounds displayed in Figure 12.4 have been calculated using a specific choice of scalar-tensor theory, in which the function $A(\varphi)$ is given by

$$A(\varphi) = \exp\left[\alpha_0(\varphi - \varphi_0) + \frac{1}{2}\beta_0(\varphi - \varphi_0)^2\right], \tag{12.33}$$

where α_0 and β_0 are arbitrary parameters, and φ_0 is the asymptotic value of the scalar field. In the notation for scalar tensor theories used here, this theory corresponds to the choice

$$\omega(\phi) = -\frac{3}{2} + \frac{1}{2(\alpha_0^2 - \beta_0 \ln \phi)}, \tag{12.34}$$

where $\phi_0 = A(\varphi_0)^{-2} = 1$. The parameters ζ and λ are given by $\zeta = \alpha_0^2/(1+\alpha_0^2)$, and $\lambda = \beta_0/2(1+\alpha_0^2)$. It is useful to note that the PPN parameter combination $4\beta - \gamma - 3$, which governs the Nordtvedt effect and other violations of SEP in the post-Newtonian limit is given by

$$4\beta - \gamma - 3 = 2\zeta(1+2\lambda) = \frac{2\alpha_0^2}{1+\alpha_0^2}\left(1 + \frac{\beta_0}{1+\alpha_0^2}\right), \tag{12.35}$$

which vanishes when $\beta_0 \simeq -1$. This partially explains the distorted "spike" in the curve for lunar laser ranging, where the violation of SEP is suppressed, and no useful bound can be obtained. Something similar is occurring in the other curves, albeit complicated by the strong gravity effects taking place in the neutron stars.

The bounds shown in Figure 12.4 were calculated using a polytropic equation of state, which tends to give lower maximum masses for neutron stars than do more realistic equations of state. It is not known how much the use of stiffer equations of state would alter the bounds on this class of theories.

Other theories

Bounds on various versions of TeVeS theories have also been established, with the tightest constraints again coming from neutron-star-white-dwarf binaries (Freire et al., 2012); in the case of TeVeS, the theory naturally predicts $\gamma = 1$ in the post-Newtonian limit, so the bounds from Cassini are irrelevant here.

Strong constraints on the Einstein-Æther and Khronometric theories were also set using binary pulsar measurements, exploiting both gravitational-wave damping data, and data related to preferred-frame effects (Yagi et al., 2014a, 2014b). For Einstein-Æther theory, for example, the approximate bounds were $c_+ < 0.03$ and $c_- < 0.003$.

12.2 Inspiralling Compact Binaries and Gravitational Waves

A new era for testing general relativity began on September 14, 2015, with the first detection of gravitational waves by LIGO (Abbott et al., 2016c). The signal, denoted GW150914, was the final burst of gravitational waves from the inspiral and merger of two black holes. Even though the signal was detectable above the noise for only 0.2 seconds, three phases could be clearly delineated: a "chirp" phase of increasing amplitude and frequency, corresponding to the late inspiral phase, a "merger" phase, where the two black holes were becoming one, and a "ringdown" phase, where the final, perturbed black emitted exponentially damped radiation from its quasinormal mode oscillations before settling down to a stationary black hole (see Figure 12.5 for an illustration). Analysis of the chirp radiation showed that the black holes had masses 26 and $39\,M_\odot$, while analysis of the ringdown radiation revealed that the final black hole had a mass of $62\,M_\odot$ and a spin parameter $\chi \simeq 0.67$. The three solar masses converted to energy during the process corresponded to an approximate luminosity of 3.6×10^{56} erg s^{-1}, larger than the luminosity of all the stars in the observable universe combined.

Already with the first detection, important tests of general relativity were carried out. One test checked the consistency between the observed signal and theoretical template waveforms based on combining post-Newtonian theory with numerical relativity within general relativity. Another test placed a bound on the mass of the graviton that improved upon the bound derived from solar system dynamics. The fourth detection, GW170814, which included the Virgo detector (Abbott et al., 2017c), resulted in an interesting test of the polarization content of the gravitational waves. The most recent detection GW170817 was a binary neutron-star inspiral (Abbott et al., 2017d). This observation was notable because it was also observed across the entire electromagnetic spectrum, with major implications for gamma-ray bursts, the neutron-star equation of state, and the synthesis of

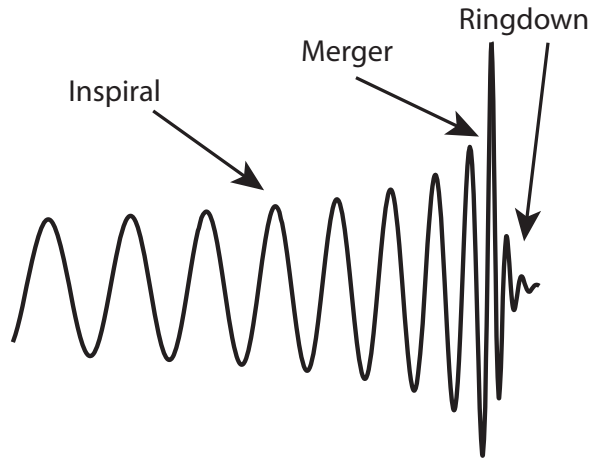

Fig. 12.5 Schematic gravitational wave from binary inspiral, showing the inspiral, merger and ringdown parts of the waveform.

the heaviest elements (Abbott et al., 2017e). It also produced a very precise direct bound on the difference between the speeds of gravitational waves and light. When the analysis is complete, it may also contribute additional tests of relativistic gravity.

In this section we will lay the groundwork for discussing gravitational waves from binary inspiral within general relativity and scalar-tensor theory, and then will discuss a few parametrized frameworks for analysing alternative theories. Finally we will describe the results from specific tests carried out so far. This field is truly just beginning, and so it is likely that much of what is presented here will soon be out of date!

12.2.1 Compact binary inspiral in general relativity

To keep the discussion simple, we will assume that general relativity is correct and that the two compact bodies are in quasi-circular orbits, which means orbits that are circular apart from the adiabatic inspiral induced by gravitational radiation. For compact binaries that have evolved in isolation for considerable time, this turns out to be an excellent approximation. As we learned in Section 11.5.2, the orbital eccentricity decreases with semimajor axis approximately as $e = e_0(a/a_0)^{19/12} = e_0(P_{\rm b}/P_{\rm b0})^{19/18}$. Thus, for example, by the time the Hulse-Taylor binary pulsar B1913+16 has spiralled inward until its orbital period is around 0.1 seconds, so that its gravitational waves will be entering the LIGO-Virgo sensitive band, its eccentricity will be 10^{-6}. On the other hand, for an "extreme mass-ratio inspiral," in which a stellar-mass compact object is injected from a surrounding star cluster into a highly eccentric orbit around a supermassive black hole, there may not be enough time for the orbit to circularize, and thus it may be important to include eccentricity.

We will also assume that the bodies are spinning, but we will assume that the spins are either aligned or anti-aligned with the orbital angular momentum. This means that we will ignore the precessions of the spins and of the orbital plane that result from spin-orbit and spin-spin coupling (see Section 9.1 for discussion). These precessions have the effect of

modulating the amplitude and phase of the gravitational waveform in complex ways. There has not been sufficient information in the short binary black hole waveforms observed so far to detect precession effects, but they will surely be important for binary black hole inspirals in the future. Similarly, it is not yet known whether spin precession played a role in the binary neutron-star inspiral event. For simplicity, we will ignore such complications, and treat only the aligned spin cases.

We will ignore tidal interactions. For binary black holes, these effects are unimportant until the onset of the actual merger of two event horizons. For systems containing neutron stars they will also be unimportant until the final few orbits, but then the signature of tidal distortion and even disruption of the neutron star as imprinted on the gravitational wave signal could carry important information about the equation of state of the matter in the neutron star. In order to test alternative theories of gravity, it is desirable to avoid regimes where such complex phenomena could occur.

Finally, we will treat only the part of the inspiral that is governed by the post-Newtonian approximation.

To obtain the relevant conditions for a circular orbit, we take the PPN two-body equation of motion (6.70) without the J_2 term, choose the PPN parameters for general relativity and impose the constraint $\dot{r} = \ddot{r} = 0$. This yields a condition on the orbital angular velocity ω given by $\omega^2 = (m/r^3)[1-(3-\eta)m/r]$. Including the spin-orbit and spin-spin contributions to the equations of motion, Eqs. (9.4a) and (9.4b), along with 2PN terms in the equations of motion (see, for example, Blanchet et al. (1995a)), we obtain

$$\omega^2 = \frac{m}{r^3}\Bigg[1 - \frac{m}{r}(3-\eta) - \left(\frac{m}{r}\right)^{3/2}\sum_a \left(2\frac{m_a^2}{m^2} + 3\eta\right)\hat{\boldsymbol{h}}\cdot\boldsymbol{\chi}_a + \left(\frac{m}{r}\right)^2\left\{6 + \frac{41}{4}\eta + \eta^2 - \frac{3}{2}\eta\left(\boldsymbol{\chi}_1\cdot\boldsymbol{\chi}_2 - 3\hat{\boldsymbol{h}}\cdot\boldsymbol{\chi}_1\hat{\boldsymbol{h}}\cdot\boldsymbol{\chi}_2\right)\right\}\Bigg], \tag{12.36}$$

where $\boldsymbol{\chi}_a = \boldsymbol{S}_a/m_a^2$, and $\hat{\boldsymbol{h}}$ is the unit vector in the direction of the orbital angular momentum $\boldsymbol{x}\times\boldsymbol{v}$. Similarly, the energy of the circular orbit is given by

$$E = -\eta\frac{m^2}{2r}\Bigg[1 - \frac{1}{4}\frac{m}{r}(7-\eta) + \left(\frac{m}{r}\right)^{3/2}\sum_a\left(2\frac{m_a^2}{m^2} + \eta\right)\hat{\boldsymbol{h}}\cdot\boldsymbol{\chi}_a - \frac{1}{8}\left(\frac{m}{r}\right)^2\left\{7 - 49\eta - \eta^2 - 4\eta\left(\boldsymbol{\chi}_1\cdot\boldsymbol{\chi}_2 - 3\hat{\boldsymbol{h}}\cdot\boldsymbol{\chi}_1\hat{\boldsymbol{h}}\cdot\boldsymbol{\chi}_2\right)\right\}\Bigg]. \tag{12.37}$$

The rate of loss of energy is given by Eq. (11.109), with spin-orbit and spin-spin effects added (Kidder et al., 1993; Kidder, 1995),

$$\frac{dE}{dt} = -\frac{32}{5}\eta^2\left(\frac{m}{r}\right)^5\Bigg[1 - \frac{m}{r}\left(\frac{2927}{336} + \frac{5}{4}\eta\right) + \left(\frac{m}{r}\right)^{3/2}\left\{4\pi - \frac{1}{12}\sum_a\left(73\frac{m_a^2}{m^2} + 75\eta\right)\hat{\boldsymbol{h}}\cdot\boldsymbol{\chi}_a\right\} + \left(\frac{m}{a}\right)^2\left\{\frac{293383}{9072} + \frac{380}{9}\eta - \frac{\eta}{48}\left(223\boldsymbol{\chi}_1\cdot\boldsymbol{\chi}_2 - 649\hat{\boldsymbol{h}}\cdot\boldsymbol{\chi}_1\hat{\boldsymbol{h}}\cdot\boldsymbol{\chi}_2\right)\right\}\Bigg]. \tag{12.38}$$

The orbital angular frequency $\omega = 2\pi f_{\rm b}$, where $f_{\rm b}$ is the binary orbit frequency. But for quadrupole radiation from a circular orbit, the gravitational-wave frequency F is given by $F = 2f_{\rm b}$. Combining Eqs. (12.36), (12.37), and (12.38) we can obtain an equation for the evolution of the gravitational-wave frequency with time as a function of F, given by $dF/dt = (dE/dt)/(dE/dF)$, or

$$\frac{dF}{dt} = \frac{96\pi}{5}F^2(\pi\mathcal{M}F)^{5/3}\left[1 - \left(\frac{743}{336} + \frac{11}{4}\eta\right)(\pi mF)^{2/3} + (4\pi - \beta)(\pi mF) + \left(\frac{34103}{18144} + \frac{13661}{2016}\eta + \frac{59}{18}\eta^2 + \sigma\right)(\pi mF)^{4/3}\right], \quad (12.39)$$

where β and σ are the spin-orbit and spin-spin contributions, given by

$$\beta \equiv \frac{1}{12}\sum_a \left(113\frac{m_a^2}{m^2} + 75\eta\right)\hat{\boldsymbol{h}}\cdot\boldsymbol{\chi}_a\,,$$
$$\sigma \equiv \frac{\eta}{48}\left(-247\boldsymbol{\chi}_1\cdot\boldsymbol{\chi}_2 + 721\hat{\boldsymbol{h}}\cdot\boldsymbol{\chi}_1\hat{\boldsymbol{h}}\cdot\boldsymbol{\chi}_2\right). \quad (12.40)$$

Note that $(\pi mF)^{2/3} \sim v^2 \sim m/r$. If the source is at a sufficiently great distance that cosmological redshifts become significant, then the observed frequency $F_{\rm obs} = F/(1+Z)$ and the observed time interval is $dt_{\rm obs} = (1+Z)dt$. Then Eq. (12.39) applies equally well to the observed quantities provided that we define the "observed" or redshifted masses $\mathcal{M}_{\rm obs} = \mathcal{M}(1+Z)$ and $m_{\rm obs} = m(1+Z)$.

Combining Eqs. (11.47) and (11.108) for a circular orbit, and substituting $r = m^{1/3}(2\pi F)^{-2/3}[1 + O(\epsilon)]$ into the wave amplitude, we can write the response function $S(t)$ of a laser interferometer in the general form

$$S(t) = \frac{\mathcal{M}}{R}Q(\text{angles})(\pi\mathcal{M}F)^{2/3}\cos\Phi(t)\,, \quad (12.41)$$

where $\Phi(t) = 2\phi(t) = 2\pi\int^t F(t')dt'$ is the gravitational-wave phase. Integrating Eq. (12.39) to obtain F as a function of t, and then integrating to obtain the phase, we obtain

$$\Phi(F) = \Phi_c - \frac{1}{16}(\pi\mathcal{M}F)^{-5/3}\left[1 + \frac{5}{3}\left(\frac{743}{336} + \frac{11}{4}\eta\right)(\pi mF)^{2/3} - \frac{5}{2}(4\pi - \beta)(\pi mF) + 5\left(\frac{3058673}{1016064} + \frac{5429}{1008}\eta + \frac{617}{144}\eta^2 - \sigma\right)(\pi mF)^{4/3}\right], \quad (12.42)$$

where Φ_c is a constant. Thus we have an accurate prediction (under the chosen assumptions) for the gravitational-wave signal at the detector. This is essential for confirming a detection and for the measurement of the source parameters (Cutler et al., 1993), which include distance, position in the sky, orientation of the orbital plane, and the masses and spins of the companions. Roughly speaking, the measured signal (which includes detector noise) is passed through a linear filter constructed from the theoretical signal $S(t;\boldsymbol{\theta})$ and the spectral density of the detector noise (Wainstein and Zubakov, 1962). The theoretical signal is expressed as a function of an abstract vector $\boldsymbol{\theta}$, which collectively represents the source parameters, such as R, t_c, $\mathcal{M}$, η, and so on. The signal-to-noise ratio S/N is then computed (see below). The actual values of these parameters $\tilde{\boldsymbol{\theta}}$ are unknown prior to the measurement. When $\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}$, the linear filter becomes the Wiener optimum

filter which is well known to yield the largest possible signal-to-noise ratio (Wainstein and Zubakov, 1962). A detection can be confirmed and the source parameters determined by maximizing S/N over a broad collection of expected signals $S(t;\boldsymbol{\theta})$, loosely referred to as "templates." This method is called "matched filtering"; see Finn and Chernoff (1993), Cutler and Flanagan (1994), and Poisson and Will (1995) for the foundations of this approach for binary inspiral.

It has been established that it is the *phasing* of the signal that plays the largest role in parameter estimation. This is because a slight variation in the parameters can quickly cause $h(t;\boldsymbol{\theta})$ to get out of phase with respect to the true signal $h(t;\tilde{\boldsymbol{\theta}})$, thus seriously reducing S/N from its maximum possible value. Therefore a good match between the phases of the template and the measured signal throughout the observed cycles singles out, to a large extent, the value of the source parameters.

The leading term in $\Phi(F)$ gives the chirp mass. If there is sufficient sensitivity to measure the 1PN correction term, then because of the dependence on η and m, one can measure the individual masses m_1 and m_2. Measuring higher-order terms can give information about the spins, and can also yield tests of general relativity.

The key ingredient in matched filtering is the Fourier transform of $S(t)$, given by $\tilde{S}(f) = (2\pi)^{-1}\int_{-\infty}^{\infty} S(t)e^{2\pi i f t}dt$. Using the stationary phase approximation, and confining attention to positive frequencies, it is straightforward to show that

$$\tilde{S}(f) = \mathcal{A}f^{-7/6}e^{i\psi(f)}\,, \tag{12.43}$$

where $\mathcal{A} \propto \mathcal{M}^{5/6}Q(\text{angles})/R$, and

$$\psi(f) = 2\pi f t_c - \frac{\pi}{4} + f\int^{f} f'^{-2}\Phi(f')df'\,, \tag{12.44}$$

where t_c is the time corresponding to the phase Φ_c, and the gravitational-wave frequency F is to be replaced by the Fourier frequency f in Eq. (12.42). In general relativity, the phase $\psi(f)$ is given by

$$\begin{aligned}\psi(f) = 2\pi f t_c - \Phi_c - \frac{\pi}{4} + \frac{3}{128}(\pi\mathcal{M}f)^{-5/3}\Bigg[1 + \frac{20}{9}\left(\frac{743}{336} + \frac{11}{4}\eta\right)(\pi m f)^{2/3} \\ -4(4\pi - \beta)(\pi m f) + 10\left(\frac{3058673}{1016064} + \frac{5429}{1008}\eta + \frac{617}{144}\eta^2 - \sigma\right)(\pi m f)^{4/3}\Bigg]\,.\end{aligned} \tag{12.45}$$

Then, for a detector characterized by a Gaussian noise spectral density $S_n(f)$, the signal-to-noise ratio associated with a measurement of a signal $S(t)$ is given by

$$S/N \equiv 4\int_0^{\infty} \frac{|\tilde{S}(f)|^2}{S_n(f)}df. \tag{12.46}$$

Defining the Fisher information matrix by

$$\Gamma_{ab} \equiv 2\int_0^{\infty} \frac{\tilde{S}^*_{,a}\tilde{S}_{,b} + \tilde{S}_{,a}\tilde{S}^*_{,b}}{S_n(f)}df, \tag{12.47}$$

where subscripts $,a$ and $,b$ denote partial derivatives with respect to one of the parameters θ^a characterizing the signal, it can be shown that the error σ_a in measuring the parameter

θ^a, and the correlation coefficients c^{ab} between parameters θ^a and θ^b are obtained from the inverse of the Fisher matrix, according to the prescription

$$\sigma_a^2 = \langle (\Delta\theta^a)^2 \rangle = (\Gamma^{-1})^{aa}\,,$$
$$c^{ab} = \frac{\langle \Delta\theta^a \Delta\theta^b \rangle}{\sigma_a \sigma_b} = \frac{(\Gamma^{-1})^{ab}}{\sigma_a \sigma_b}\,. \tag{12.48}$$

This sensitivity to phase is one of the keys to testing general relativity with gravitational waves. In an alternative theory of gravity, the coefficients that appear in Eq. (12.45) may differ from those in general relativity, and if those differences depend on a parameter, such as the scalar-tensor coupling parameter ω_0, one can use matched filtering to place a bound on such parameters.

12.2.2 Compact binary inspiral in scalar-tensor theories

At present, the evolution of the gravitational-wave frequency in scalar-tensor theory cannot be written down to the same PN order as that displayed in Eq. (12.39) for general relativity. This is because dipole gravitational radiation contributes to the waveform at -0.5PN order and to the energy flux at -1PN order, and therefore to obtain terms in dF/dt at nPN order beyond the quadrupole term, one needs to evaluate the scalar wave field and the equations of motion to $(n+1)$PN order. In addition, the clean expansion in powers of $(\pi m F)^{2/3} \sim v^2$ shown in Eq. (12.39) is no longer so clean when dipole radiation is present. Even though the dipole term is formally of leading order, its contribution to the waveform and energy flux is proportional to ζ. Since solar-system experiments already place the constraint $\zeta < 10^{-5}$, and $\mathcal{S}$ is typically less than 0.5, the dipole flux is likely to be small compared to the quadrupole flux in the late inspiral phase, where $v > 10^{-2}$, except for situations where spontaneous scalarization yields extremely large values of $\mathcal{S}$. Thus a double expansion of quantities such as dF/dt and $\Phi(F)$ must be carried out, with one parameter being $(\pi m F)^{2/3}$, the other being the ratio of the two dominant types of flux (see Sennett et al. (2016) for discussion). The expansions will look very different depending on whether the evolution is dipole-radiation dominated or quadrupole-radiation dominated.

Here we will keep things simple, and work only to the equivalent of quadrupole order, mainly to demonstrate how dipole radiation alters the frequency evolution of the inspiral (Will, 1994). From the binary equation of motion (10.69) without spins, we follow the method described in Section 12.2.1 to obtain the quasicircular orbit expressions (compare Eqs. (12.36) and (12.37))

$$\omega^2 = \frac{\mathcal{G}m}{r^3}\left[1 - \frac{\mathcal{G}m}{r}\left(3 - \eta + \bar{\gamma} + 2\bar{\beta}_+ + 2\Delta\bar{\beta}_-\right)\right],$$
$$E = -\eta\frac{\mathcal{G}m^2}{2r}\left[1 - \frac{1}{4}\frac{\mathcal{G}m}{r}\left(7 - \eta + 4\bar{\gamma}\right)\right]. \tag{12.49}$$

From Eqs. (11.59), the energy loss rate from a quasicircular orbit is given by

$$\frac{dE}{dt} = -\frac{8}{15}\eta^2\kappa_1\frac{m}{r}\left(\frac{\mathcal{G}m}{r}\right)^4 - \frac{4}{3}\zeta\eta^2\frac{m}{r}\left(\frac{\mathcal{G}m}{r}\right)^3\mathcal{S}_-^2\,, \tag{12.50}$$

where κ_1 is given by Eq. (11.121). The final result for the evolution of the gravitational-wave frequency is

$$\frac{dF}{dt} = \frac{96\pi}{5} F^2 (\pi \mathcal{G}\mathcal{M}F)^{5/3} \left[\xi + \frac{5}{24} (\pi \mathcal{G} m F)^{-2/3} \zeta \mathcal{S}_-^2 \right], \tag{12.51}$$

where

$$\begin{aligned} \xi \equiv 1 + \frac{5}{12}\bar{\gamma} - \frac{5}{144}\zeta \mathcal{S}_-^2 \left(3 + 7\eta - 4\bar{\gamma} + 16\bar{\beta}_+ + 16\Delta\bar{\beta}_-\right) \\ + \frac{5}{6}\zeta \mathcal{S}_- \left(\frac{\mathcal{S}_-\bar{\beta}_+ + \mathcal{S}_+\bar{\beta}_-}{\bar{\gamma}} \right) + \frac{5}{6}\zeta \Delta\mathcal{S}_- \left(\frac{\mathcal{S}_+\bar{\beta}_+ + \mathcal{S}_-\bar{\beta}_-}{\bar{\gamma}} \right). \end{aligned} \tag{12.52}$$

Whereas in general relativity, the post-Newtonian corrections to the leading quadrupole frequency evolution are given by a power series in increasing powers of $(\pi \mathcal{G} m F)^{2/3}$, the dipole correction term depends on a negative power that parameter, reflecting its -1PN nature. But by measuring the evolution of F in an inspiral, one can place a bound on the scalar-tensor parameter ζ. Unfortunately, for the sensitivities of the advanced LIGO-Virgo instruments, the bounds are not likely to be competitive with the solar-system bound (Will, 1994), except possibly for scalar-tensor theories that admit spontaneous scalarization, leading to anomalously large values of $\mathcal{S}_\pm$. On the other hand, such bounds would be derived from tests of the strong-field dynamical regime, not the weak field, slow motion regime of the solar system. Note, however, that for binary black hole inspirals, with $s_1 = s_2 = 1/2$, the gravitational-wave signal is identical to that of general relativity, after rescaling the masses by the factor $1 - \zeta$, so no test of basic scalar-tensor theory is possible for such sources. One needs a neutron star inspiral, such as the event GW170817, to test these theories.

Sennett et al. (2016) have analyzed the next-order contributions to the frequency and phase evolution.

12.2.3 Compact binary inspiral in other theories

Compact binary inspiral in other theories of gravity has not been analyzed systematically in the same manner as have general relativity and scalar-tensor theories, so only limited conclusions about alternative theories can be drawn from the gravitational-wave detections made to date (Yunes et al., 2016). On the other hand, a number of phenomenological parametrizations have been developed that make it possible to place some preliminary bounds on general classes of theories.

Bounds on the graviton mass and the speed of gravitational waves

If gravitation is propagated by a massive field, as in massive gravity theories (Section 5.7), then the velocity of gravitational waves will depend upon their wavelength as $v_g^2 = 1 - (\lambda/\lambda_g)^2$ [Eq. (11.10)], where $\lambda_g = h/m_g$ is the graviton Compton wavelength. In the case of inspiralling compact binaries, gravitational waves emitted at low frequency early in the inspiral will thus travel slightly more slowly than those emitted at high frequency later, resulting in an offset in the relative arrival times at a detector. This modifies the observed

frequency evolution of the observed inspiral gravitational waveform, similar to that caused by post-Newtonian corrections to the quadrupole evolution (Will, 1998).

To make this quantitative, we consider the propagation of a massive "graviton" in a background Friedmann-Robertson-Walker (FRW) homogeneous and isotropic spacetime, with the line element

$$ds^2 = -dt^2 + a^2(t)[d\chi^2 + \Sigma^2(\chi)(d\theta^2 + \sin^2\theta d\phi^2)]\,, \tag{12.53}$$

where $a(t)$ is the scale factor of the universe and $\Sigma(\chi)$ is equal to χ, $\sin\chi$ or $\sinh\chi$ if the universe is spatially flat, closed or open, respectively. For a graviton moving radially from an emitter at $\chi = \chi_e$ to a receiver at $\chi = 0$, it is straightforward to show that the component of the graviton's four-momentum p_χ = constant. Using the fact that $m_g^2 = -p^\alpha p^\beta g_{\alpha\beta} = E^2 - a^{-2}p_\chi^2$, where $E = p^0$, together with $p^\chi/E = d\chi/dt$, we obtain

$$\frac{d\chi}{dt} = -\frac{1}{a}\left(1 + \frac{m_g^2 a^2}{p_\chi^2}\right)^{-1/2}\,, \tag{12.54}$$

where $p_\chi^2 = a^2(t_e)(E_e^2 - m_g^2)$. Assuming that $E_e \gg m_g$, expanding Eq. (12.54) to first order in $(m_g/E_e)^2$ and integrating, we obtain

$$\chi_e = \int_{t_e}^{t_r} \frac{dt}{a(t)} - \frac{1}{2}\frac{m_g^2}{a^2(t_e)E_e^2}\int_{t_e}^{t_r} a(t)dt\,. \tag{12.55}$$

Consider gravitons emitted at two different times t_e and t'_e, with energies E_e and E'_e, and received at corresponding arrival times (χ_e is the same for both). Assuming that $\Delta t_e \equiv t_e - t'_e \ll a/\dot{a}$, and noting that $m_g/E_e = (\lambda_g F_e)^{-1}$, where F_e is the emitted frequency, we obtain, after eliminating χ_e,

$$t_r = t'_r + (1+Z)\left[t_e - t'_e + \frac{D}{2\lambda_g^2}\left(\frac{1}{F_e^2} - \frac{1}{F_e'^2}\right)\right]\,, \tag{12.56}$$

where $Z \equiv a_0/a(t_e) - 1$ is the cosmological redshift, and

$$\begin{aligned} D &\equiv \frac{(1+Z)}{a_0}\int_{t_e}^{t_a} a(t)dt \\ &= \frac{1+Z}{H_0}\int_0^Z \frac{dZ'}{(1+Z')^2[\Omega_M(1+Z')^3 + \Omega_\Lambda]^{1/2}}\,, \end{aligned} \tag{12.57}$$

where $a_0 = a(t_r)$ is the present value of the scale factor. The expression in the second line corresponds to a spatially flat universe with matter density parameter Ω_M and dark-energy parameter Ω_Λ, with $\Omega_M + \Omega_\Lambda = 1$. Note that D is not a conventional cosmological distance measure, like the luminosity distance given by

$$\begin{aligned} D_L &\equiv (1+Z)a_0\int_{t_e}^{t_a}\frac{dt}{a(t)} \\ &= \frac{1+Z}{H_0}\int_0^Z \frac{dZ'}{[\Omega_M(1+Z')^3 + \Omega_\Lambda]^{1/2}}\,. \end{aligned} \tag{12.58}$$

For a standard ΛCDM cosmology, D/D_L ranges from unity at $Z = 0$ to 0.3 at $Z = 3$.

Because the arrival time of the signal now depends on frequency, the *observed* frequency evolution is given by

$$\frac{dF}{dt} = \frac{dF_r}{dt_r} = (1+Z)^{-1}\frac{dF_e}{dt_e}\frac{dt_e}{dt_r}\,. \tag{12.59}$$

where we have inserted the redshift factor linking F_r with F_e. We now calculate dt_r/dt_e using Eq. (12.56), and insert the general relativistic quadrupole evolution for dF_e/dt_e, Eq. (12.39), including the 1PN correction. We are assuming here that, in whatever massive gravity theory is being employed, the intrinsic frequency evolution of the system is given by the formulae from general relativity, ignoring corrections of fractional order $(r/\lambda_g)^2$, where r is the size of the binary system. Since solar-system bounds already imply that $\lambda_g > 10^{12}$ km, these are likely to be small corrections. We then obtain, for the observed frequency evolution,

$$\frac{dF}{dt} = \frac{96\pi}{5}F^2(\pi\mathcal{M}F)^{5/3}\left[1-\left(\frac{743}{336}+\frac{11}{4}\eta+\beta_g\right)(\pi mF)^{2/3}+\dots\right], \tag{12.60}$$

where F, m, and $\mathcal{M}$ are the observed frequency and masses, and where

$$\beta_g = \frac{96}{5}\frac{\pi^2 D\eta m}{\lambda_g^2(1+Z)}\,. \tag{12.61}$$

The effect of the graviton mass is to modify the coefficient of the 1PN term in the frequency evolution. Thus with sufficiently accurate measurements of the frequency evolution, it is possible to measure the individual masses and to place a bound on β_g. Early studies estimated that advanced LIGO-Virgo detectors could place a bound $\lambda_g > 10^{12}$ km, while the LISA space detector could place a bound $\lambda > 10^{16}$ km (Will, 1998; Berti et al., 2005; Stavridis and Will, 2009; Arun and Will, 2009).

The dispersion relation $E^2 = p^2 + m_g^2$ for a massive graviton is Lorentz invariant. Mirshekari et al. (2012) considered extensions to theories of gravity that effectively violate Lorentz invariance, in which the dispersion relation could take the form

$$E^2 = p^2 + m_g^2 + Ap^\alpha\,, \tag{12.62}$$

where A and α are two parameters characterizing the Lorentz violation. For example, in some extra-dimension theories, $\alpha = 4$ and A is a parameter of the order of the square of the Planck length. Mirshekari et al. (2012) studied the bounds that could be placed on m_g and A for various values of α.

Tests of the massive graviton hypothesis have now been carried out using data from gravitational-wave detections. Data from the discovery event GW150914, led to the bound $\lambda_g > 10^{13}$ km (Abbott et al., 2016c,d). This was improved to 1.6×10^{13} km by combining data from the three events GW150914, GW151226, and GW170104. (Abbott et al., 2017b).

This limit is comparable to the solar-system limit, which comes from assuming that massive gravity also implies a modification of the Newtonian potential by the inclusion of a Yukawa factor, that is, $U = (m/r)e^{-r/\lambda_g}$. Data on Mercury's perihelion advance and on orbits of the outer planets imply that $\lambda > 10^{12}$ km (Talmadge et al., 1988), although recent improvements in knowledge of Mercury's orbit described in Section 7.3 probably push

that bound above 10^{13} km. Some have argued for a larger bound on λ_g from galactic and cluster dynamics (Hiida and Yamaguchi, 1965; Hare, 1973; Goldhaber and Nieto, 1974), noting that the evidence of bound clusters and of clear tidal interactions between galaxies argues for a range λ_g at least as large as a few megaparsecs (6×10^{19} km). However, in view of the uncertainties related to the amount of dark matter in the universe, and the fact that massive gravity theories frequently include other modifications of gravity on large scales, these bounds should be viewed with caution. Even the dynamical arguments used to establish these large-scale bounds have been questioned recently (Christodoulou and Kazanas, 2017; Mukherjee and Sounda, 2017). Bounds on the Lorentz-violating dispersion parameter A were also placed using combined data from GW150914, GW151226, and GW170104 (Abbott et al., 2017b).

A very strong bound was placed on the difference in speed between light and gravitational waves using the binary neutron-star event GW170817 (Abbott et al., 2017a). Because the gravitational-wave event could be associated with the gamma-ray burst event GRB170817A in the galaxy NGC 4993, the observed time difference between the two signals, 1.74 ± 0.05 s, combined with the estimate of the distance of 42.9 ± 3.2 Mpc, produced the bound

$$-3 \times 10^{-15} < c_g - 1 < 7 \times 10^{-16}\,. \tag{12.63}$$

Bounds on polarizations of gravitational waves

A restricted test of gravitational-wave polarizations was carried out in 2017, when the Virgo detector joined LIGO for several weeks toward the end of LIGO's second observing run. Using data from the binary black hole inspiral GW170814, seen in all three detectors, a Bayesian analysis was carried out in which a fit to the data was carried out for pure tensor response (i.e., including only F_+ and $F_\times$ in the response), pure scalar response (only $F_{\rm S}$ and $F_{\rm L}$) and pure vector response (only $F_{\rm V1}$ and $F_{\rm V2}$). The quadrupole response was strongly favored over the two alternatives by Bayes factors of 200 and 1000, respectively (Abbott et al., 2017c).

Parametrizations of binary inspiral signals

Several frameworks have been developed, modeled on the PPN formalism, in which arbitrary parameters are introduced into the predicted gravitational-wave signal from binary inspiral in order to encompass alternative theories of gravity. Arun et al. (2006a) wrote the Fourier phase $\psi(f)$ of Eq. (12.45) in the form

$$\psi(f) = 2\pi f t_c - \phi_c - \frac{\pi}{4} + \sum_k \frac{3}{128\eta}(\pi M f)^{(k-5)/3}\alpha_k\,, \tag{12.64}$$

where α_k are arbitrary parameters. The general relativistic values of the parameters can be read off from the coefficients inside the square brackets of Eq. (12.45) (see also Arun et al. (2006b) and Mishra et al. (2010)).

Yunes and Pretorius (2009) extended this post-Newtonian parametrization to include nongeneral relativistic frequency dependences, in the form

$$\psi(f) = 2\pi f t_c - \phi_c - \frac{\pi}{4} + \sum_{k=0}^{N-1} \phi_k (\pi M f)^{b_k}\,, \tag{12.65}$$

where ϕ_k and b_k are each a set of N parameters. This parametrization allows for the presence of dipole radiation, which begins at order $(\pi M f)^{-7/3}$. They also included parametrized waveform templates for both the merger and ringdown phase. This is known as the parametrized post-Einsteinian (PPE) formalism.

These parameterizations were extended and adapted for specific use within the LIGO-Virgo data analysis pipeline, based on Bayesian model selection (Li et al., 2012; Agathos et al., 2014). This "test infrastructure for general relativity" (TIGER) was used in a joint analysis of the data from GW150914 and GW151226 to place limits on the various parameters of the model (Abbott et al., 2016a). For example, the 0PN, 1PN, and 1.5PN coefficients in the model were consistent with general relativity to between 10 and 20 percent; higher PN-order coefficients were more poorly constrained. With more detections, these constraints are likely to improve.

12.3 Exploring Spacetime near Compact Objects

One of the difficulties of testing GR in the strong-field regime is the possibility of contamination by uncertain or complex physics. In the solar system, weak-field gravitational effects can in most cases be measured cleanly and separately from nongravitational effects. The remarkable cleanliness of many binary pulsars permits precise measurements of gravitational phenomena in a strong-field context. The gravitational waves from inspiraling compact binaries detected by LIGO and Virgo appear to be likewise amazingly clean, not entirely surprising, since four of the first five detections involved binary black holes.

Unfortunately, nature is rarely so kind. Still, under suitable conditions, qualitative and even quantitative strong-field tests of GR are possible. The combination of improved astronomical observations in all wavelength bands, from radio to gamma rays, and better theoretical and computational modelling of complex physical processes have offered the promise of performing striking tests of strong-field predictions of GR. This is a rapidly evolving field, and therefore we will be unable to give a thorough account here. Instead we will give a few examples of interesting arenas where possible astrophysical tests of general relativity might occur, and will otherwise refer readers to a number of recent reviews (Psaltis, 2008; Berti et al., 2015; Johannsen, 2016).

12.3.1 SgrA*: A black hole in the galactic center

Lynden-Bell and Rees (1971) suggested that there might be a supermassive black hole in the center of the Milky Way, and Balick and Brown (1974) observed a bright, unresolved synchroton radio source at the precise galactic center, naming it Sagittarius A*, or SgrA* for short. While evidence mounted for the existence of massive black holes in the centers of quasars and active galactic nuclei, a black hole in the Milky Way remained only an

intriguing possibility until advances in infrared interferometry and adaptive optics made it possible to detect stars very near the location of SgrA* and to observe their orbital motions (Eckart and Genzel, 1996; Ghez et al., 1998). A major breakthrough came with the unambiguous determination of the orbit of the star denoted S2 with an orbital period of about 16 years, and projected semimajor axis of 0.1 arcseconds (Schödel et al., 2002; Ghez et al., 2003). This made it possible, using Kepler's third law, to determine the mass of the black hole directly. Combined with more recent data on S2 and on other stars in the central cluster, this has led to an estimate for the black-hole mass of $4.4 \pm 0.4 \times 10^6\, M_\odot$ and an improved value for the distance to the galactic center of 8.3 ± 0.4 kiloparsecs (see Genzel et al. (2010) for a thorough review of SgrA*). In addition to opening a window on the innermost region of the galactic center, the discovery of these stars has made it possible to contemplate using orbital dynamics to probe the curved spacetime of a rotating black hole, with the potential to test general relativity in the strong-field regime.

Soon after the detection of stars orbiting SgrA*, numerous authors pointed out the relativistic effects that are potentially observable, including the gravitational redshift and time-dilation effects (Zucker et al., 2006) and the pericenter advance (Jaroszynski, 1998; Fragile and Mathews, 2000; Rubilar and Eckart, 2001; Weinberg et al., 2005; Kraniotis, 2007). Although these effects were not detectable with the instrumentation at the time, notably the infrared interferometer at the Keck Observatory in Hawaii and the Very Large Telescope infrared Interferometer (VLTI) at the European Southern Observatory, the next generation of the instrumentation may bring relativistic tests within reach. The upgrade of the VLTI, known as GRAVITY, made its first observations of the galactic center in June 2016 (GRAVITY Collaboration et al., 2017). The Thirty Meter Telescope (TMT), currently under development, will also have the capability to make such measurements.

For example, the gravitational redshift of spectra of the stars S2 and S102 will be detectable during their next pericenter passages in 2018 and 2021, respectively. The pericenter precession of S2 is another target for detection in 2018. In general relativity, the leading contribution to the pericenter advance rate is given by $\dot{\omega} = 6\pi m/P_{\rm b} a(1-e^2)$. However, if the precession of the orbit is being measured astrometrically from Earth, then we must project the orbit onto the plane of the sky and determine the rate of change of the relevant point on the orbit as seen from Earth. Thus the projected pericenter shift is given by $\dot{\omega}_{\rm proj} = [a(1 \pm e)/R]\dot{\omega}\cos\iota$ where the plus (minus) sign corresponds to measuring the shift at apocenter (pericenter), ι is the inclination of the orbit, and R is the distance to the galactic center. Using the mass and distance of SgrA* quoted earlier, the result is

$$\dot{\omega}_{\rm proj} = 98.3\,\mu{\rm as\,yr}^{-1}\left(\frac{1\,{\rm yr}}{P_{\rm b}}\right)\frac{\cos\iota}{1\mp e}\,, \tag{12.66}$$

where the minus sign now corresponds to a measurement at apocenter. For S2 and S102, the rates are 37 and 23 microarcseconds per year. Current observations of S2 and S102 are within a factor of 10 of measuring the precession (Hees et al., 2017).

Among the goals of second-generation projects such as GRAVITY and the TMT are astrometric precisions of tens of microarcseconds per year, and sensitivities enabling the detection of stars orbiting much closer to the black hole than the S-stars (if such stars exist). This makes it possible to consider doing more than merely detect relativistic effects, but

Table 12.3 Orbital parameters of selected stars orbiting the galactic center black hole SgrA*. Data taken from Gillessen et al. (2009) and Meyer et al. (2012).

Star	a (a.u)	e	$\iota(^\circ)$	P_b (yr)
S2	1020 ± 8	0.880 ± 0.003	135.25 ± 0.47	15.8 ± 0.1
S4	2470 ± 160	0.406 ± 0.022	77.83 ± 0.32	59.5 ± 2.6
S5	2080 ± 350	0.842 ± 0.017	143.7 ± 4.7	45.7 ± 6.9
S9	2430 ± 430	0.825 ± 0.020	81.0 ± 0.7	58.0 ± 9.5
S14	2125 ± 83	0.963 ± 0.006	99.4 ± 1.0	47.3 ± 2.9
S38	1160 ± 340	0.802 ± 0.041	166 ± 22	18.9 ± 5.8
S102	812 ± 32	0.68 ± 0.02	151 ± 3	11.5 ± 0.3

rather to provide the first test of the black hole "no-hair" or uniqueness theorems of general relativity (Will, 2008). According to those theorems, an electrically neutral black hole is completely characterized by its mass m and angular momentum J. As a consequence, all the multipole moments of its external spacetime are functions of m and J. Specifically, the quadrupole moment $Q_2 = -J^2/m$.

To see how such a test might be carried out, we work in the post-Newtonian limit, and write down the equation of motion for a test body in the field of a body with mass m, angular momentum $\boldsymbol{J}$, and quadrupole moment Q_2. This can be obtained from Eq. (6.70), using general relativistic values of the PPN parameters, setting $\eta = 0$, and including the fact that the dimensionless quadrupole moment J_2 and Q_2 are related by $mR^2 J_2 = -Q_2$. We also include the spin-orbit terms from Eq. (9.4a), with $\gamma = 1$, $\alpha_1 = 0$, and $\boldsymbol{S_1} = m_1 = 0$, hence $\boldsymbol{\sigma} = 0$. The result is

$$\begin{aligned}\frac{d\boldsymbol{v}}{dt} = &-\frac{m\boldsymbol{n}}{r^2}\left[1 + v^2 - \frac{4m}{r}\right] + \frac{4m\boldsymbol{v}\dot{r}}{r^2} \\ &- \frac{2J}{r^3}\left[2\boldsymbol{v}\times\boldsymbol{e} - 3\dot{r}\boldsymbol{n}\times\boldsymbol{e} - 3r^{-1}\boldsymbol{n}(\boldsymbol{h}\cdot\boldsymbol{e})\right] \\ &- \frac{3}{2}\frac{Q_2}{r^4}\left[5\boldsymbol{n}(\boldsymbol{n}\cdot\boldsymbol{e})^2 - 2\boldsymbol{e}(\boldsymbol{n}\cdot\boldsymbol{e}) - \boldsymbol{n}\right], \end{aligned} \tag{12.67}$$

where $\boldsymbol{v}$ is the velocity of the body, $\boldsymbol{n} = \boldsymbol{x}/r$, $\boldsymbol{h} = \boldsymbol{x}\times\boldsymbol{v}$, and $\boldsymbol{e} = \boldsymbol{J}/J$ is a unit vector along the symmetry axis of the black hole. If we define a dimensionless spin parameter χ by $\chi \equiv J/m^2$, then, for a Kerr black hole, $0 \le \chi \le 1$ and $Q_2 = -M^3\chi^2$.

We consider a body with orbital elements a, e, ι, Ω and ω defined using the plane of the sky as the X–Y plane of the reference system (see Figure 6.1). The axis of the black hole has an unknown orientation relative to the reference system. Substituting the perturbing acceleration of Eq. (12.67) into the Lagrange planetary equations (6.74) and integrating over an orbit, we obtain the changes $\Delta e = \Delta a = 0$, and

$$\begin{aligned}\Delta\varpi &= A_S - 2A_J\cos\alpha - \frac{1}{2}A_Q(1 - 3\cos^2\alpha), \\ \Delta\Omega &= \frac{\sin\alpha\sin\beta}{\sin\iota}(A_J - A_Q\cos\alpha), \\ \Delta i &= \sin\alpha\cos\beta(A_J - A_Q\cos\alpha), \end{aligned} \tag{12.68}$$

where

$$A_S = 6\pi m/p\,,$$
$$A_J = 4\pi J/(mp^3)^{1/2}\,,$$
$$A_Q = 3\pi Q_2/mp^2\,, \tag{12.69}$$

where $\Delta\varpi = \Delta\omega + \cos\iota\,\Delta\Omega$, and $p = a(1-e^2)$. The angles α and β are the polar angles of the black hole's symmetry axis $\boldsymbol{e}$ with respect to the star's orbital plane defined by the line of nodes $\boldsymbol{e}_\Omega$, and the vector in the orbital plane $\boldsymbol{e}_\perp$ orthogonal to $\boldsymbol{e}_\Omega$ and $\mathbf{h}$, according to $\cos\alpha \equiv \boldsymbol{e}\cdot\hat{\boldsymbol{h}}$, $\sin\alpha\cos\beta \equiv \boldsymbol{e}\cdot\boldsymbol{e}_\Omega$ and $\sin\alpha\sin\beta \equiv \boldsymbol{e}\cdot\boldsymbol{e}_\perp$.

The structure of the expressions for $\Delta\Omega$ and $\Delta\iota$ can be understood as follows: Eq. (12.67) implies that the orbital angular momentum $\boldsymbol{h}$ precesses according to $d\boldsymbol{h}/dt = \boldsymbol{\omega}_{\rm P}\times\mathbf{h}$, where the orbit-averaged precession vector is given by $\boldsymbol{\omega}_{\rm P} = \boldsymbol{e}(A_J - A_{Q_2}\cos\alpha)$. The orbit element variations are given by $d\iota/dt = \boldsymbol{\omega}_{\rm P}\cdot\boldsymbol{e}_\Omega$ and $\sin\iota\, d\Omega/dt = \boldsymbol{\omega}_{\rm P}\cdot\boldsymbol{e}_\perp$. As a consequence, we have the purely geometric relationship,

$$\frac{\sin\iota\, d\Omega/dt}{di/dt} = \tan\beta\,. \tag{12.70}$$

To get rough idea of the astrometric size of these precessions, we define an angular precession rate amplitude $\dot\Theta_n = (a/R)A_n/P_{\rm b}$, where R is the distance to the galactic center and $P_{\rm b}$ is the orbital period. Using $m = 4.4\times10^6\,M_\odot$, $D = 8.3$ kpc, we obtain the rates, in microarcseconds per year,

$$\dot\Theta_S \approx 98.3\,P_{\rm b}^{-1}(1-e^2)^{-1}\,,$$
$$\dot\Theta_J \approx 1.07\,\chi P_{\rm b}^{-4/3}(1-e^2)^{-3/2}\,,$$
$$\dot\Theta_Q \approx 1.3\times10^{-2}\,\chi^2 P_{\rm b}^{-5/3}(1-e^2)^{-2}\,, \tag{12.71}$$

where $P_{\rm b}$ is measured in years. The observable precessions will be reduced somewhat from these raw rates because the orbit itself must be projected onto the plane of the sky. As we noted earlier, the contribution to $\Delta\varpi$ is reduced by $\cos\iota$, and the contribution to $\Delta\iota$ is reduced by $\sin\iota$. For an orbit in the plane of the sky, changes in the inclination are unmeasurable, and changes in the nodal angle become degenerate with changes in the pericenter.

For the quadrupole precessions to be observable, it is clear that the black hole must have a decent angular momentum ($\chi > 0.5$) and that the star must be in a short-period high-eccentricity orbit. For example, for $\chi = 0.7$, $P_{\rm b} = 0.1$ yr, and $e = 0.9$, the three amplitudes listed in Eq. (12.71) have the values 5200, 195 and 8 μ as per year, respectively.

Although the pericenter advance is the largest relativistic orbital effect, it is *not* the most suitable effect for testing the no-hair theorems. The frame-dragging and quadrupole effects are small corrections of the leading Schwarzschild pericenter precession, and thus one would need to know m, a and e to sufficient accuracy to be able to subtract that dominant term to reveal the smaller effects of interest. Furthermore, the pericenter advance is affected by a number of complicating phenomena, including 2PN effects, the disturbing effects of surrounding mass in the form of gas, stars or dark matter, and the effects of tidal distortion of the star near its pericenter.

On the other hand, the precessions of the node and inclination are relatively immune from such effects. Any spherically symmetric distribution of mass has no effect on these orbit elements. As long as any tidal distortion of the star is quasi-equilibrium with negligible tidal lag, the resulting perturbing forces are purely radial, and thus have no effect on the node or inclination. On the other hand, even if a surrounding cluster of stars is spherical on average, the "graininess" of the perturbing forces from a finite number of such stars will have an effect on the orbital plane of a chosen target star. These perturbing effects were studied for a range of hypothetical distributions of cluster stars by Merritt et al. (2010) and Sadeghian and Will (2011)

In order to test the no-hair theorem using the precessions $\Delta\Omega$ and $\Delta\iota$, we need to measure four quantities, the magnitudes J and Q_2, and the angles α and β. Thus it will be necessary to measure $\Delta\Omega$ and $\Delta\iota$ for at least two stars whose orbital planes are appropriately nondegenerate. Even if the precision is not sufficient to be sensitive to the small quadrupole effect, it would still be possible to measure the spin magnitude and direction of the black hole, which would shed light on how it evolved and grew during the lifetime of the galaxy.

Observing stars is not the only way to explore the spacetime near SgrA*. If a pulsar were observed orbiting sufficiently close to the black hole, observations of its orbital precession using pulsar timing rather than astrometry could also contribute to a test of the no-hair theorem (Wex and Kopeikin, 1999; Liu et al., 2012). In addition to stars, there are a number of gaseous disks orbiting the central black hole (Genzel et al., 2010). Working at submillimeter wavelengths and linking a worldwide set of radio telescopes to form a giant VLBI array, a collaboration known as the Event Horizon Telescope (EHT) is approaching the capability of imaging SgrA* with event-horizon-scale angular resolution (Doeleman et al., 2009). Observation of accretion phenomena in the innermost disk at these angular resolutions could provide tests of the spacetime geometry very close to the black hole. (Johannsen et al., 2016a,b). And combining data from stellar astrometry, pulsar timing and EHT could have significant advantages for testing the no-hair theorem (Psaltis et al., 2016).

The observations needed to explore the strong-field region around SgrA* are very challenging, but steady progress is being made, notably with GRAVITY and the EHT. In addition, some luck will be called for: it is not known whether a population of sufficiently bright young stars exists close enough to the black hole to make no-hair tests feasible. Nor is it known if suitable pulsars exist sufficiently close to SgrA*, although the discovery of a magnetar (a pulsar with an extremely large magnetic field) in the neigborhood of the black hole gives reason for hope.

12.3.2 Neutron stars and black holes

Neutron stars and stellar-mass black holes may also be important arenas for testing strong-field gravity.

Studies of certain kinds of accretion known as advection-dominated accretion flow (ADAF) in low-luminosity binary X-ray sources have yielded hints of the signature of the black hole event horizon (Narayan and McClintock, 2008). The spectrum of frequencies of quasi-periodic oscillations (QPO) from accretion onto black holes and neutron stars in

binaries may permit measurement of the spins of the compact objects. Aspects of strong-field gravity and frame-dragging may be revealed in spectral shapes of iron fluorescence lines and continuum emission from the inner regions of accretion disks (Reynolds, 2013; Miller and Miller, 2015). See Abramowicz and Fragile (2013) for a review of accretion onto black holes and neutron stars.

The structure of neutron stars depends strongly on both the equation of state of nuclear matter and the theory of gravity, and there is considerable degeneracy between these two ingredients in such parameters as the mass and radius of the neutron star. However, the discovery of the "I-Love-Q" phenomenon, a relation between the moment of inertia, the Love number and the quadrupole moment of rotating neutron star models that is remarkably insensitive to the equation of state (Yagi and Yunes, 2013), may break the degeneracy and open up ways to test alternative theories of gravity.

For detailed reviews of strong-field tests of GR involving neutron stars and black holes using electromagnetic observations, see Psaltis (2008) and Johannsen (2016).

12.4 Cosmological Tests

From a few seconds after the Big Bang until the present, the underlying physics of the universe is well understood, in terms of a standard model of a nearly spatially flat universe, 13.6 billion years old, dominated by cold dark matter and dark energy, called the ΛCDM model. Notwithstanding some observational tensions, such as slightly differing values of the Hubble parameter coming from different observational techniques, difficulties accounting for the distribution of structures over all galactic scales, and the failure to date to detect the fundamental particle that is presumed to constitute dark matter, the general relativistic ΛCDM model agrees remarkably well with a wide range of observations.

Other theories, such as Brans-Dicke theory, are sufficiently close to general relativity (for large enough ω_0) that they conform to all cosmological observations, within the uncertainties. Certain generalized scalar-tensor theories, however, could have small values of ω at early times making aspects of early universe evolution highly non general relativistic, while evolving to large ω today, thereby agreeing with all solar system and astrophysical experiments, and hewing closely to late-time ΛCDM cosmology (Damour and Nordtvedt, 1993a,b).

One way to test such theories is through Big-Bang nucleosynthesis (BBN), since the abundances of the light elements produced when the temperature of the universe was about 1 MeV are sensitive to the rate of expansion at that epoch, which in turn depends on the field equations of the theory. Because the universe is radiation-dominated at that epoch, uncertainties in the amount of cold dark matter or dark energy (Λ) are unimportant. The nuclear reaction rates are reasonably well understood from laboratory experiments and theory, and the number of light neutrino families (3) conforms with evidence from particle accelerators. Thus, within modest uncertainties, one can assess the quantitative difference between the BBN predictions of general relativity and other theories of gravity under strong-field conditions and compare with observations. For recent analyses, primarily

within scalar-tensor theories see Santiago et al. (1997), Damour and Pichon (1999), Clifton et al. (2005), and Coc et al. (2006).

A different class of theories has been developed in part to provide an alternative to the dark energy of the standard ΛCDM model, by modifying gravity on large, cosmological scales, while preserving the conventional solar and stellar-system phenomenology of general relativity. The $f(R)$ theories are examples. Since we are now in a period of what may be called "precision cosmology," one can begin to test alternative theories in this class using the accumulation of data on many fronts, including the growth of large-scale structure, cosmic background radiation fluctuations, galactic rotation curves, BBN, weak lensing, baryon acoustic oscillations, and so on.

Apart from direct measurements of the expansion rate of the universe through the observations of standard candles such as Type II supernovae, most information about the universe comes from studying deviations from a homogeneous, isotropic Friedmann-Robertson-Walker background spacetime, expressed roughly in the form

$$ds^2 = -dt^2 + a(t)^2 \left[\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right] + h_{\mu\nu} dx^\mu dx^\nu \,, \tag{12.72}$$

where t is proper time as measured by a clock at rest, $a(t)$ is the scale factor, here with units of distance, r, θ, ϕ, and x^μ are dimensionless coordinates, and $k = 0$ or ± 1 determines the curvature of the spatial sections. Unlike the PPN formalism, where the background spacetime was flat and it was relatively simple to characterize the deviations represented by $h_{\mu\nu}$, in the cosmological case, there are many different approaches, with a range of possible gauge choices. In addition, the evolution of the background will itself depend on the theory of gravity being studied. Finally the split between "background" and "perturbation" may not be as clean as it was in the post-Newtonian case.

This has led to a rich variety of approaches to treating cosmological tests of gravitational theories, including Amin et al. (2008), Daniel et al. (2010), Dossett et al. (2011), Dossett and Ishak (2012), Zuntz et al. (2012), Hojjati et al. (2012), Baker et al. (2013), and Sanghai and Clifton (2017). For a comprehensive review, see Ishak (2018). This is another rapidly evolving field, and further details are beyond the scope of this book.