

Cosmic sirens: discovery of gravitational waves and their impact on astrophysics and fundamental physics

Sanjeev Dhurandhar^{1,*} and Bangalore S. Sathyaprakash²

¹Inter University Centre for Astronomy and Astrophysics, Pune, India

²Institute for Gravitation and Cosmos, Department of Physics, Penn State University, USA and School of Physics and Astronomy, Cardiff University, Cardiff, United Kingdom

On 14 September 2015, the twin detectors belonging to the Laser Interferometer Gravitational Wave Observatory (LIGO) made a triple discovery: the first direct detection of gravitational waves (GWs), first observation of formation of a black hole and first observation of a binary black hole. Since then LIGO has reported two other events and a marginal candidate. These discoveries have heralded a new era in observational astronomy. They will help us in exploring extremes of astrophysics and gravity. GWs are our best chance of getting an idea of what went on a small fraction of a second after the big bang, even if that takes many more decades. With LIGO's discoveries we hope to solve many puzzles in astronomy and fundamental physics, but GWs are guaranteed to show up objects and phenomena never imagined before.

Keywords:

Introduction

Two remarkable predictions of Albert Einstein's general relativity are black holes and gravitational waves (GWs). Karl Schwarzschild derived the first black hole metric¹ and Einstein showed the existence of wave-like solutions in the linear approximation^{2,3}, both within a few months after he presented his general relativity (GR) to the Prussian Academy on 25 November 1915 (ref. 4).

Almost exactly after 100 years, on 14 September 2015, the twin instruments of the Laser Interferometer Gravitational Wave Observatory (LIGO) located at Livingston, Louisiana and Hanford, Washington, USA, confirmed both these predictions with the observation of GWs from a pair of colliding black holes⁵. This transient event called GW150914 has been hailed as the greatest physics discovery of the century and the biggest progress in astronomy since Galileo Galilei's observation of the sky.

Two other high-confidence detections made on 26 December 2015 (ref. 6) and 4 January 2017 (ref. 7), and a

marginal detection on 12 October 2015, are also from coalescing black hole binaries⁸. These discoveries have ushered in a new era in understanding the cosmos. The nascent field of GW astronomy and astrophysics will not only open a new window on the Universe⁹, but more importantly, it will allow precision tests of GR and alternative theories of gravity⁸. This article will focus on the salient points of the theory of GWs and LIGO's discovery, with particular emphasis on the properties of GWs and LIGO's black holes.

Gravitational waves – their generation, propagation and properties

In GR, Einstein's equations admit wave solutions. This is readily seen if we make a weak field approximation². A weak GW is described by a metric perturbation $h_{\mu\nu}$ in GR. To get an idea of the weakness of GW, typically, for detectable astrophysical GW sources, $h_{\mu\nu} \lesssim 10^{-22}$. To the linear order in $h_{\mu\nu}$, it can be easily shown in the transverse and traceless (TT) gauge¹⁰, that Einstein's field equations reduce to the wave equations

$$\left(\frac{\partial^2}{c^2 \partial t^2} - \nabla^2 \right) h_{\mu\nu} = \frac{16\pi G}{c^4} T_{\mu\nu}, \quad (1)$$

where $c \sim 3 \times 10^5$ km/s is the speed of light, G Newton's gravitation constant and $T_{\mu\nu}$ the stress–energy tensor. It is clear that GR predicts GWs and GW travel with the speed of light c . Thus GWs are waves in the metric field $g_{\mu\nu}$. Since the curvature or Riemann tensor is essentially the second derivative of the metric, GWs can also be described as ripples in the curvature of space–time. One may use either the metric or the curvature to describe GWs.

Further, the following properties of GWs can be deduced from GR: they are transverse, and they have two polarizations denoted by h_+ and h_\times . The two polarization states are easily understood, if we examine the effect of the waves on free test particles. We need at least two spatially separated particles to observe the effect of GWs (a single free mass particle cannot detect a wave (or any

*For correspondence. (e-mail: sanjeev@iucaa.in)

SPECIAL SECTION:

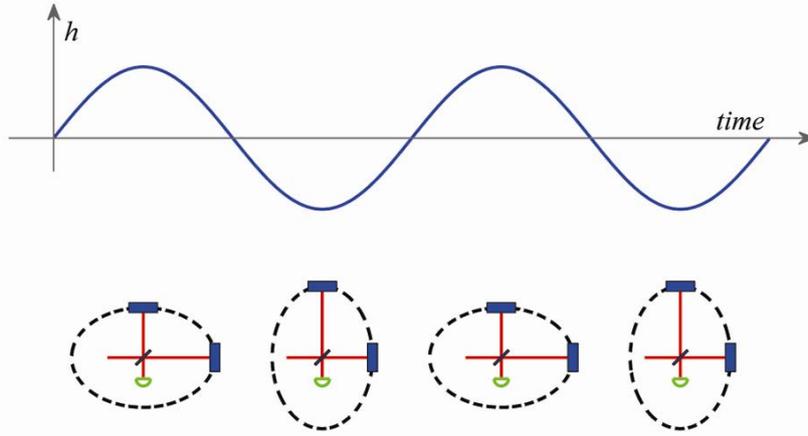


Figure 1. A circular ring of test particles deformed into an ellipse by an incident gravitational wave (image taken from ref. 44). Phases, half a cycle apart, are shown for the + polarization. The length change in the interferometric arms is also schematically shown.

gravity) because of the equivalence principle of GR); one tracks the variation in the distance between the particles as a function of time. Since $h_{\mu\nu}$ is a symmetric trace-free second rank tensor (in a certain gauge), a circular ring of test particles best demonstrates the effect of a wave which shears the circle into an ellipse. If a weak monochromatic GW of + polarization is incident on a ring of test particles, the ring is deformed into an ellipse as shown in Figure 1. Phases, half a cycle apart, of the GWs are shown in the figure. For the \times polarization the ellipses are rotated by an angle of 45° . A general wave is a linear combination of the two polarizations.

We will confine ourselves to interferometric detection¹¹. If we select two masses on this ring of test masses at right angles and monitor their distance with respect to the centre of the ring, which we take to be the reference point, we will find that during an half cycle of the wave one arm shortens while the other arm elongates. In the next half cycle of the wave the opposite happens. Using a laser interferometric arrangement, a passing GW will produce a time-varying path difference which can be detected on a photodiode.

However, there is a catch. The changes in distances are exceedingly small in astrophysical situations. For example, a neutron star binary at a distance of 100 Mpc (1 Mpc or mega parsec is $\sim 10^{19}$ km) – a typical distance to a GW source – will produce a differential length change of $\sim 10^{-16}$ cm for test masses kept a few kilometres apart, which is the typical length of the arm of a large-scale ground-based interferometric detector. For a GW source, the well-known Landau–Lifschitz quadrupole formula¹² relates the metric perturbation $h_{\mu\nu}$ to the source described by $T_{\mu\nu}$. This formula can be obtained by integrating the inhomogeneous wave equation, i.e. eq. (1) under certain assumptions. The formula relates the GW amplitude h_{ik} (spatial part) in the TT gauge to the second

time derivative of the quadrupole moment of the source. It reads

$$h_{ik}^{\text{TT}} = \frac{2G}{Rc^4} \left[\ddot{I}_{ik} \left(t - \frac{R}{c} \right) \right]^{\text{TT}}, \quad (2)$$

where R is the distance to the source, I_{ik} the quadrupole moment tensor of the source at the retarded time $t - R/c$, the double dot over I_{ik} represents its second time derivative and the superscript TT denotes the transverse traceless part of the tensor. The luminosity of a GW source is obtained by integrating the energy flux of gravitational radiation over all directions in the sky and is given by the formula

$$\mathcal{L} = \frac{G}{5c^5} \langle \ddot{I}_{ik}^{\text{TT}} \ddot{I}^{ik\text{TT}} \rangle, \quad (3)$$

where the angular brackets denote the time average over a period. If M is the mass of the system, r the size of the system and Ω its frequency, then $\mathcal{L} \sim (G/c^5) M^2 r^4 \Omega^6$, where $\alpha \sim 1$ is a measure of the nonsphericity of the source.

Note that \ddot{I}_{ik} has dimensions of energy and the operation of taking its transverse traceless part means that it is the nonspherical motion of the source, the kinetic energy in the nonspherical motion, which we denote by $E_{\text{nonspherical}}^{\text{kinetic}}$, is responsible for producing GWs. If we consider $E_{\text{nonspherical}}^{\text{kinetic}}/c^2$ of the order of a solar mass and the distance to the source ranging from galactic scale of tens of kpc to cosmological distances of Gpc, then h (a typical component of h_{ik}) ranges from 10^{-17} to 10^{-22} . These numbers then set the scale for the sensitivities at which the detectors must operate.

The quantity h relates to the change in distance between the test particles. For ground-based detectors if L is the distance separating the test masses, then the change in distance δL due to a GW with metric perturbation h is easily obtained from the geodesic deviation equation

$$\delta L \sim hL. \quad (4)$$

Detection involves impossibly small measurements, and therefore, the noise in the detector needs to be suppressed by several orders of magnitude in order that there is a chance of extracting the signal from the noise by statistical signal detection methods. A host of noise sources contaminate the data. Some of the main sources of noise are seismic noise at low frequencies, thermal noise at mid-frequencies up to a few hundred hertz and the photon shot noise at high frequencies. Sophisticated techniques have been evolved and devised over the decades to control such noises. These efforts have shown that long arm lengths, high laser power, and extremely well-controlled laser stability are essential to reach the requisite sensitivity.

The response in the detector is a linear combination of the two time-dependent polarization amplitudes h_+ and h_\times , and is given by

$$h(t) = h_+(t)F_+(\phi, \theta, \psi) + h_\times(t)F_\times(\phi, \theta, \psi), \quad (5)$$

where F_+ and F_\times are the antenna pattern functions, and ϕ , θ , ψ carry information about the direction to the source in the sky and polarization relative to the orientation of the detector¹³. For the binary inspiral, the polarization amplitudes at the leading order take the form

$$h_+(t) = A(t)(1 + \cos^2 \iota) \cos \Phi(t), \quad (6)$$

$$h_\times(t) = -2A(t) \cos \iota \sin \Phi(t). \quad (7)$$

The phase evolution $\Phi(t)$ determines the masses which in turn determine the absolute amplitude of the wave. The inclination angle ι of the orbital plane to the line of sight may be deduced from a network of detectors with different orientations from polarization information. The observed amplitude depends on the absolute amplitude, inclination angle and distance. If the first two can be measured, then one can solve for the distance to the binary.

Detection and measurement of gravitational waves

Consider a data segment $[0, T]$ uniformly sampled with N time samples t_k , $k = 0, 1, \dots, N-1$ and sampling interval Δ ; we then have $T = N\Delta$ and $t_k = k\Delta$. If $x(t)$ are the data, the N samples $x(t_k) = x_k$ can be regarded as an N -dimensional column vector \mathbf{x} with components x_k (ref. 14). We

first consider the simple situation where a data vector \mathbf{x} either contains only noise, i.e. $\mathbf{x} = \mathbf{n}$, or contains a given signal \mathbf{s} plus noise, $\mathbf{x} = \mathbf{s} + \mathbf{n}$, where we have assumed additive noise. This is called a binary hypothesis and the hypotheses are denoted by H_0 and H_1 respectively. The data vector \mathbf{x} is a vector random variable with a multivariate probability distribution. The probability distributions will differ depending on whether the signal is present in the data or not. The multivariate probability distributions are denoted by $p_0(\mathbf{x})$ for H_0 and $p_1(\mathbf{x})$ for H_1 . Given \mathbf{x} we must decide between the hypotheses H_0 and H_1 . The assumption of additive noise implies $p_1(\mathbf{x}) = p_0(\mathbf{x} - \mathbf{s})$.

We will consider the case of wide sense stationary (WSS) Gaussian noise¹⁵, although in reality it is neither of these. The WSS assumption is justified, if the signals are for short durations; then the noise for that duration can be considered to satisfy the WSS conditions. The WSS assumption requires that the first two moments of the noise are invariant under time translations. That is

$$\langle n(t) \rangle = \langle n(t + \tau) \rangle, \quad \langle n(t)n(t') \rangle = \langle n(t + \tau)n(t' + \tau) \rangle, \quad (8)$$

for all time translations τ and where the angular brackets denote ensemble average. Physically, this indicates that the mean and covariance between noise samples do not depend on absolute time, but only on the time difference. We will consider zero mean noise (the DC component is removed), i.e. $\langle n(t) \rangle = 0$. The noise is then characterized by its autocorrelation function K defined as follows

$$\langle n(t)n(t') \rangle = K(t - t'), \quad (9)$$

The Fourier transform of $K(t)$ is the power spectral density (PSD) $S(f)$.

For describing probability distributions, it is convenient to go over to the sampled data vector \mathbf{x} . We define the covariance matrix of the noise vector as

$$C_{ik} = K(t_i - t_k) = K((i - k)\Delta). \quad (10)$$

Then the pdf for Gaussian noise is a multivariate Gaussian with mean zero and is given by

$$p_0(\mathbf{x}) = \frac{1}{(2\pi)^{N/2} \sqrt{\det C}} e^{-\frac{1}{2} \mathbf{x}^T C^{-1} \mathbf{x}}. \quad (11)$$

We may define the matrix μ (Fischer information matrix) as the inverse of C ; then in the exponent of the pdf we get the quantity $\mu_{ik} x_i x_k$. This motivates the definition of a scalar product. We define a scalar product of two data vectors \mathbf{x} and \mathbf{y} as

$$(\mathbf{x}, \mathbf{y}) = \mu_{ik} x_i y_k. \quad (12)$$

SPECIAL SECTION:

Then we can write

$$p_0(\mathbf{x}) = \frac{(\det \mu)^{1/2}}{(2\pi)^{N/2}} e^{-\frac{1}{2}(\mathbf{x}, \mathbf{x})}. \quad (13)$$

If we denote the set of data trains with \mathcal{D} , then with the scalar product so defined, \mathcal{D} is a Hilbert space. In the Fourier space in case of WSS noise, the metric μ_{ik} is diagonalized; $\mu_{ik} \rightarrow \delta(f-f') = S(f)$. In the literature, $S(f)$ is replaced by the one-sided PSD which we again denote by the same symbol. Then eq. (12) can be written as

$$(\mathbf{x}, \mathbf{y}) = 4\Re \int_0^\infty df \frac{\bar{x}^*(f)\bar{y}(f)}{S(f)}, \quad (14)$$

where we have taken the real part of the integral.

To decide between the hypotheses, we devise a test. A test amounts to partitioning \mathcal{D} into two disjoint regions \mathcal{R} and \mathcal{R}^c , where \mathcal{R}^c is the complement of \mathcal{R} . Thus if $\mathbf{x} \in \mathcal{R}$, the signal is present or H_1 is true, otherwise $\mathbf{x} \in \mathcal{R}^c$, H_0 is true and the signal is absent. Now the problem reduces to determining the region \mathcal{R} . In order to do this, we define the false alarm and detection probabilities P_F and P_D respectively, as follows

$$P_F(\mathcal{R}) = \int_{\mathcal{R}} p_0(\mathbf{x}) d^N x, \quad P_D(\mathcal{R}) = \int_{\mathcal{R}} p_1(\mathbf{x}) d^N x. \quad (15)$$

The solution to finding \mathcal{R} is given by the Neyman–Pearson approach as follows. We first define the likelihood ratio

$$\Lambda(\mathbf{x}) = \frac{p_1(\mathbf{x})}{p_0(\mathbf{x})}. \quad (16)$$

Given a false alarm probability $P_F = \alpha$, the detection region \mathcal{R} is determined by $P[\Lambda(\mathbf{x}) > \Lambda_0] = \alpha$ and $\mathcal{R} = [\mathbf{x} \in \mathcal{D} | \Lambda(\mathbf{x}) > \Lambda_0]$. Λ_0 is called the threshold and is decided by the false alarm probability α .

We illustrate the situation for Gaussian noise. Then the pdfs for the signal absent and signal present cases are the following

$$p_0(\mathbf{x}) = A_N e^{-\frac{1}{2}(\mathbf{x}, \mathbf{x})}, \quad p_1(\mathbf{x}) = A_N e^{-\frac{1}{2}(\mathbf{x} - \mathbf{s}, \mathbf{x} - \mathbf{s})}. \quad (17)$$

Here A_N is the normalization constant which normalizes the pdfs. It is convenient to deal with the log likelihood ratio, which is

$$\ln \Lambda(\mathbf{x}) = (\mathbf{x}, \mathbf{s}) - \frac{1}{2}(\mathbf{s}, \mathbf{s}). \quad (18)$$

The monotonicity of the functions involved allows us to translate the equation, $\Lambda(\mathbf{x}) \geq \Lambda_0$ to $\rho = (\mathbf{x}, \mathbf{s}) \geq \rho_0$. We

may write $\rho = q_i x_i$ where $q_i = \mu_{ik} s_k$. The vector \mathbf{q} is called the matched filter. The matched filter is optimal in the Neyman–Pearson sense – it maximizes the detection probability for a given false alarm probability¹⁶.

Although the above procedure is simple and easy to implement, we are not in this situation. We have a family of signals (instead of a single given signal), say $\mathbf{s}(\lambda)$ where the vector parameter λ collectively denotes the parameters λ^α , $\alpha = 1, 2, \dots, p$. For example, in the case of coalescing binary sources, the signal depends on the individual masses, spins, distance to the binary and other kinematical parameters such as time of arrival, initial phase, etc. All these constitute the λ^α . The family of vectors $\mathbf{s}(\lambda)$ traces out a p -dimensional manifold – a submanifold of \mathcal{D} . This is called the signal manifold \mathcal{S} . The signal parameters λ^α can be regarded as coordinates on the signal manifold¹⁷.

We now have to deal with composite hypotheses:

- H_0 : No signal present, i.e., $\mathbf{x} = \mathbf{n}$, noise only and the corresponding pdf is $p_0(\mathbf{x})$.
- H_1 : One among the family of signals $\mathbf{s}(\lambda)$ is present, i.e., $\mathbf{x} = \mathbf{n} + \mathbf{s}(\lambda)$ with the corresponding pdf now denoted by $p_1(\mathbf{x}; \lambda)$.

We may use the Bayesian framework to address this problem. We are provided with a prior pdf $z(\lambda)$ on the parameter space which is based on either prior astrophysical information or some other consideration. Following on the lines of the previous case, we define the likelihood ratio for a specific signal with parameters λ as

$$\Lambda(\mathbf{x}; \lambda) = \frac{p_1(\mathbf{x}; \lambda)}{p_0(\mathbf{x})}. \quad (19)$$

The Bayesian statistic is the marginalized likelihood given by

$$\Lambda_B(x) = \int d\lambda \Lambda(\mathbf{x}; \lambda) z(\lambda). \quad (20)$$

This is again optimal in the sense that the average detection probability is maximized for a given false alarm probability.

Computing the likelihood ratio is expensive – one would have to compute the matched filter at tens of millions of points in searching over the parameter space. Since our computational resources are limited, we can search the parameter space only over a finite number of grid points in \mathcal{S} . These vectors are called templates, which we take to be normalized. The template placement is achieved elegantly by defining a metric on the parameter space^{17,18}. In fact, the scalar product defined on \mathcal{D} in eq. (14) induces a metric on \mathcal{S}

$$g_{\alpha, \beta} = \left(\frac{\partial \mathbf{s}}{\partial \lambda^\alpha}, \frac{\partial \mathbf{s}}{\partial \lambda^\beta} \right). \quad (21)$$

For those points that do not coincide with the grid points, we allow a maximum mismatch of ε ; typically $\varepsilon = 0.03$ allowing for a 10% loss in detected events. The placement of templates is governed by the equation

$$g_{\alpha\beta}\Delta\lambda^\alpha\Delta\lambda^\beta = \varepsilon. \quad (22)$$

The contours traced out by $\Delta\lambda$ are approximately hyperellipsoids. The simplest arrangement is that of hypercubes, where the hypercube fits inside the hyperellipsoid. The template is at the centre of the hypercube of side Δl which is determined by requiring that the square of the distance from a vertex, namely $p(\Delta l/2)^2 = \varepsilon$. The number of templates $N_{\text{templates}}$ is the volume of the parameter space divided by the volume of the hypercube and is given by

$$N_{\text{templates}} = \frac{\int d\lambda \sqrt{g}}{\Delta l^p}. \quad (23)$$

where g is the determinant of the metric $g_{\alpha\beta}$.

However, in general, the template placement problem is difficult because one needs good coordinates on \mathcal{S} , which may not be readily available. In the spinless case, a good choice of coordinates are the chirp times, say, τ_0, τ_3 , in which the metric components are nearly constant leading to a uniform grid. However, in general, there are only suboptimal solutions. In two dimensions, an improvement over the hypercube is the hexagonal packing which reduces the number of templates by 23%. In O1 run of the LIGO detectors $\sim 250,000$ templates were used in compact binary coalescence searches. Sathyaprakash and Dhurandhar¹⁹ for the first time showed how to match filter the data with a bank of templates and further presented an optimal strategy for detecting GW signals of known shapes.

Coalescing compact binaries: their dynamics and signature in gravitational wave radiation

In eq. (1), the left-hand-side is the curvature of strain and the right-hand-side consists of the matter stress–energy tensor. Time-varying internal stresses cause ripples in the fabric of space–time that travel outwards from their sources at the speed of light. The coupling constant between the stress–energy tensor and the curvature of strain is G/c^4 and it has dimensions of 1/force. This force $G_F \equiv c^4/G \sim 10^{44}$ N, which we shall call the geometric force, is enormous. It implies that only when stresses in a system are large enough to produce forces this big can we expect the strain amplitudes to be the greatest. Let us ask when do self-gravitating systems experience this force. In a binary system of a pair of stars of total mass M , the centripetal force is $F = Mv^2/r = v^4/G$, where r is the distance between the bodies and v is the rotational speed. In the

last of the equalities we have used Kepler’s law $v^2 = GM/r$. We can rewrite the centripetal force in terms of the geometric force: $F = G_F (v/c)^4$. Thus the centripetal force is always smaller than the geometric force, coming close to it when $v \sim c$. For most astronomical binaries, $v \ll c$. However, for black hole and neutron star binaries, v can get very close to c just before they merge. It is only then that the strains produced are the largest ever possible. For this reason, black hole and neutron star binaries are considered to be the most powerful sources of GWs.

A binary system is the perfect example of a negative feedback system; the emission of radiation causes the system to become more unstable and emit an even greater amount of radiation by bringing the component bodies together. After evolving for millions of years, the two bodies would approach each other at relativistic speeds, collide and merge, emitting in the process a burst of radiation. Ground-based interferometric detectors are designed to observe the final moments of and minutes of their lifetime.

From the quadrupole formula given in eq. (2), we can estimate the strain amplitude from an inspiralling binary of total mass M and reduced mass μ as follows. The non-spherical kinetic energy in the system is $E_{\text{non-spherical}}^{\text{kinetic}} \simeq \mu v^2/2$: From Kepler’s law, $v^2 = GM/r$, where r is the size of the orbit. Hence the amplitude is

$$h \sim 2\eta \frac{GM}{c^2 R} \frac{GM}{c^2 r}, \quad (24)$$

where $0 \leq \eta = \mu/M \leq 1/4$ is the symmetric mass ratio and R is the luminosity distance to the source. The magnitude of the strain is $h \sim 2.5 \times 10^{-23}$ for a binary of total mass $M = 1 M_\odot$, when the bodies are separated by a distance of $r = 10 GM/c^2$.

The dynamics of an inspiralling binary can be determined by demanding that the luminosity in GWs \mathcal{L} comes at the expense of the binding energy of the system $E \sim \eta M v^2/2$. The luminosity can be estimated from eq. (3) to be

$$\mathcal{L} = \frac{32c^5\eta^2}{5G} \left(\frac{GM}{rc^2} \right)^5. \quad (25)$$

The formula holds good for a binary on a quasi-circular orbit, quasi because the orbit is in reality shrinking due to the loss of energy. A binary on an eccentric orbit with the same semi-major axis as a circular orbit has a greater luminosity.

The emission of the radiation has a back reaction on the binary, causing the two stars to spiral in. Starting from a separation r , the rate \dot{r} at which the orbit shrinks can be estimated by equating the luminosity to the rate of change of the orbital energy, i.e. $dE/dt = -\mathcal{L}$. From the

SPECIAL SECTION:

latter equation, the rate at which the separation r changes as a function time is

$$\frac{dr}{dt} = -\frac{64G^3\eta}{5c^5} \left(\frac{M^3}{r^3} \right). \quad (26)$$

From this equation we can work out the lifetime t_C of the system starting from a separation $r = r_0$ to be

$$t_C = \frac{3}{256} \frac{c^5}{G^3} \frac{r_0^4}{\eta M^3}. \quad (27)$$

t_C is called the chirp-time. It is often convenient to write the chirp-time in terms of the GW frequency f which, due to the symmetry of the system, is equal to twice the orbital frequency f_{orb} . Using Kepler's law we find $\pi^2 f^2 = GM/r^3$. Upon substituting this expression in the equation for t_C we get

$$t_C = \frac{5GM}{256c^3\eta(\pi GMf_0/c^3)^{8/3}}. \quad (28)$$

Starting from a frequency of $f_0 = 30$ Hz, an equal mass system of total mass $M = 65 M_\odot$ will last for $t_C = 25$ ms. This is the in-band duration of the first GW signal detected by LIGO in the inspiral regime and increases rapidly as the detector sensitivity improves at lower frequencies.

Imprint in the amplitude and phase of the signal are the parameters of the source which include the luminosity distance R , the component masses (m_1, m_2) and spins (\vec{S}_1, \vec{S}_2) of the binary, the sky position of the source (θ, ϕ), inclination of the orbit relative to the line of sight ι (i.e. angle between the radial vector from the detector to the source and instantaneous angular momentum of the binary) and the orientation of the ellipse formed by the orbit of the binary projected on the sky (which essentially determines the combination of the polarizations measured by a detector). For long-lived signals, the amplitude and phase could be modulated by the Earth's motion relative to the source, but this is not important for signals that last for less than half an hour in the sensitive band of the detector.

The source parameters could be extracted by finding the waveform parameters that best describe the signal. This is done by Bayesian inference. Given the data x , one computes the posterior probability of the parameters θ_k

$$P(\theta_k | x) = \frac{P(x | \theta_k)P(\theta_k)}{P(x)}, \quad (29)$$

where $P(x|\theta_k)$ is the likelihood, $P(\theta_k)$ the prior probability of the parameters and $P(x) = \int P(x|\theta_k)d^n\theta_k$ is the evidence.

As the source parameters are varied, they do not necessarily produce differences in waveforms that can be discerned from the background noise. This degeneracy of the waveforms indicates that not all parameters can be measured equally well. Table 1 shows the parameters of the LIGO sources detected so far.

LIGO's black holes are the heaviest stellar mass black holes known to us. The observed luminosity in GWs for the three confirmed events reached $\mathcal{L} \sim \text{few} \times 10^{56}$ erg/s, which is only three orders of magnitude smaller than the theoretical maximum luminosity c^5/G that a source could have. At that moment the luminosity in GWs exceeded the electromagnetic luminosity of the entire Universe.

The rate at which the frequency of GWs evolves depends on the intrinsic luminosity of a binary, while the amplitude of the radiation contains information about the extrinsic luminosity of the source. The two together can be used to infer the luminosity distance to a source. GW150914 and GW151226 are both at a distance of 1.3 billion light years, while GW170104 and LVT151012 are both measured to be at a distance of almost 3 billion light years.

Implications of the discovery for astrophysics and fundamental physics

In this section we will briefly discuss the implications of LIGO's discoveries to fundamental physics and astrophysics.

Astrophysical implications

From Table 1 we see that the masses of four companions out of eight black holes are larger than $\sim 20 M_\odot$. Indeed, LIGO's black holes are among the heaviest stellar mass

Table 1. Parameters of the source corresponding to LIGO's three detections GW150914, GW151226 and GW170104 and one candidate LVT151012 as of June 2017: matched filter signal-to-noise ratio, signal arrival time difference between the Hanford and Livingston detectors Δt_{HL} , peak luminosity in gravitational waves (GWs), radiated energy in the process of merger, luminosity distance to the source, masses of the companions and total mass of binary, and mass and dimensionless spin of the remnant black hole

Source parameter	150914	151226	170104	151012
Signal-to-noise ratio	24	13	13	9.7
Δt_{HL} /ms	7	1.1	-2.9	-0.6
Peak GW luminosity/(10^{56} erg/s)	3.6	3.3	3.1	3.1
Radiated energy/ $M_\odot c^2$	3.0	1.0	2.0	1.5
Luminosity distance/Mpc	420	440	880	1020
Primary mass/ M_\odot	36	15	31	23
Secondary mass/ M_\odot	29	7	19	13
Total mass/ M_\odot	65	22	51	37
Remnant mass M/M_\odot	62	21	49	35
Remnant spin $c \vec{S} /(GM^2)$	0.68	0.74	0.64	0.66

black holes known in astronomy. X-ray observations have identified many black hole candidates, but these are typically in the mass range $5\text{--}29 M_{\odot}$ (ref. 20). In fact, according to current understanding of stellar evolution, it is not possible to form black holes of mass greater than $\sim 20 M_{\odot}$ from stellar processes unless they have sub-solar metallicities^{9,21}, which were considered to be unlikely. Thus LIGO's discoveries are already impacting astrophysical models of the evolution of massive stars and their environments.

Binary black holes can form in isolated binaries or in dense stellar clusters by dynamical processes^{9,22,23}. Massive stars in an isolated binary could each go through the main sequence phase at the end of which they suffer gravitational collapse and undergo supernova. If the progenitor stars are more massive than about $\sim 20 M_{\odot}$, then the remnant could be a black hole²¹. If supernova kicks are large, the resulting binary may be disrupted or so wide that gravitational radiation emission may not be sufficiently large to drive the system to merger within the Hubble time. If the kicks are small enough, then the resulting binary may merge within the Hubble time and can be observed in the LIGO band.

In the second scenario massive stars in dense nuclear or globular clusters could end up as black holes that sink into the centre of the cluster by dynamical friction, where they may pair up with other black holes that have previously sunk in. Multiple interactions between such black holes could lead to hardening of the binaries that could merge within the Hubble time (ref. 9 and references therein).

LIGO's detections are currently consistent with either scenario^{7,9}. Future observations could discriminate between different models as merger rates, spin orientations, and mass and spin distributions are different for alternative models.

Tests of general relativity

LIGO's binary black holes and the remnants that form from their merger are ideal test beds of strong fields and dynamical space-times. The dimensionless gravitational potential at the Earth due the gravitational field of the Sun is $\phi/c^2 = GM_{\odot}/rc^2 \sim 10^{-8}$, where $r = 1$ AU is the distance between the Sun and the Earth. Even in radio binary pulsars like the PSR1913+16 and J0737-3039, it is $\phi/c^2 \sim 10^{-4}$. The potential is of the order unity for two black holes at the time of merger $\phi/c^2 \sim 0.5$ (refs 24, 25). Moreover, the strongly nonlinear nature of the problem implies that it is necessary to compute the solution to the two-body dynamics in the perturbative post-Newtonian approximation, which would ensure that the waveforms built from such approximations are accurate enough for the purpose of detection, measurement and tests of GR.

In the post-Newtonian approximation²⁶, one describes the phase evolution of the orbit by an expansion in the

small parameter v , the characteristic speed of the companion stars in the binary. (For the sake of simplicity, we shall neglect logarithmic terms that are present in such an expansion. Tests of the post-Newtonian approximation are actually performed on the coefficients of the phase in the Fourier phase; the time-domain description given here provides the gist of the tests.)

$$\Phi(v(t)) = \sum_n \Phi_k(v/c)^k. \quad (30)$$

The post-Newtonian coefficients Φ_k depend on the masses and spins of the companion black holes. One can use LIGO's detections to test if these parameters are as predicted by post-Newtonian approximation or differ from their values in GR²⁷. The same approach can also be used to test the waveforms for the full inspiral, merger and ringdown that are obtained analytically and matched to the exact solutions obtained by numerical solutions at a finite number of points in the parameter space^{28,29}. These tests have so far not revealed any departure of the data from GR, once again proving Einstein right^{7,8}.

The late time dynamics of the system requires its full nonlinear description and hence it is during the merger phase that we might expect GR to fail. We have looked for any departure of the post-inspiral signal with that signal from the inspiral regime³⁰, but found no deviations from the predictions of GR^{7,8}.

In Einstein's theory, the waves travel at the speed of light and suffer practically no dispersion as they traverse over vast distances of billions of light years from their sources to the Earth. In a massive graviton theory the waves will be dispersed, causing lower frequencies to be delayed by a larger magnitude than higher frequencies, hence inducing change in the phase evolution of the signals³¹. The very first observations have already begun to set interesting upper limits on the possible dispersion and also a limit on the mass of the graviton of $m_g < 7 \times 10^{-23}/c^2$ eV (ref. 7).

Outlook

In 1909, Geiger and Marsden smashed alpha particles at gold foil³², which in 1911 led Rutherford to discover the structure of the atom³³. A 100 years hence we are at the verge of exploring the very structure of space-time with a similar experiment by observing black holes – pure geometric objects – smashing against each other. LIGO's first discoveries are just the beginning: GW astronomy will herald a new era in fundamental physics, cosmology and astrophysics³⁴, giving us access to processes with phenomenal energies, inconceivable in accelerators, and luminosities far exceeding all processes but the big bang itself.

Neutron stars are the densest objects that we know; the density at their cores can be two times that of atomic

SPECIAL SECTION:

nuclei at $\sim 4 \times 10^{17} \text{ g/cm}^3$. The structure of matter at such densities has not been understood, although there are many different models for the state of matter at such densities which could consist (mostly) of just neutrons, quark-gluon plasma or hyperons. Close to merger, the phase evolution of GWs from binary neutron stars will be different from binary black holes due to tidal deformation of neutron stars. The extent of that deformation, and the induced quadrupole moment, depend on the equation-of-state. Hence a careful measurement of the waveform buried in detector noise could potentially reveal the structure of neutron star cores³⁵.

Formation and evolution of compact binaries consisting of neutron stars and black holes is an important question in astrophysics²². For example, black hole binaries can form from massive stellar binaries when the companion stars each end their lives in a supernova resulting in black holes. Alternatively, they may form in dense stellar clusters where black holes formed from isolated massive stars sink to the centre of the cluster, where they pair up with black holes that have sunk before. Distribution of masses and spins of compact objects deduced from GW observations could be used to distinguish between different formation scenarios of binary compact objects. A related question is when did massive black holes found in the nuclei of most galaxies form and what were their masses at birth and how did they grow? By observing black holes throughout the Universe, which is possible with more sensitive detectors, it should be possible to answer this question in the coming years.

LIGO's detections so far have not been able to test many predictions of GR so far. We cannot really tell if the signals detected by LIGO are really from black holes. The signal is consistent with other compact objects such as boson stars³⁶ or gravastars³⁷ that are also expected to result in an exponentially damped ringdown of the final remnant. The true signature of a black hole³⁸ is in the quasi-normal modes it emits³⁹, which consist of an infinite spectrum of damped sinusoids whose frequency and time constant depend only on two parameters – the black hole mass and spin. Ringdowns of other compact objects will depend on the detailed inner structure of the stars. Thus, detecting quasi-normal modes that depend only on two parameters would be the tell-tale signature of the formation of a black hole^{40–42}.

Other nonlinear effects expected to be detected in future observations include precession of the orbital plane, and higher modes in the radiation such as the octupole that contains GW frequencies that are all multiples of the orbital frequency.

In GR, GWs have only two independent polarizations, although as many as six polarizations are possible in a generic metric theory of gravity⁴³. Measuring the wave polarization would require a network of three or more non-colocated interferometers.

Understanding matter and gravity at extreme densities will be the science goal for future observations by LIGO, but to observe the full spectrum of GR predictions and constrain astrophysical models would require a new generation of detectors that are 10–30 times more sensitive than advanced detectors. Indeed, this nascent field is also poised to make new discoveries of objects and phenomena unknown to us today, but observation of stochastic GWs that should have been produced at the birth of the Universe is our only window to understand what happened in the earliest moments of creation and the quantum nature of space–time.

1. Schwarzschild, K., *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften* (Berlin), 1916, **1**, 189.
2. Einstein, A., *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften* (Berlin), Seite, 1916, 688–696.
3. Einstein, A., *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math Phys.)* 1918, **154**, 1918.
4. Einstein, A., *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1915 **1**, 778 [Addendum: *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1915, **799**, 1915].
5. Abbott, B. P. *et al.* (LIGO Scientific Collaboration, Virgo Collaboration). *Phys. Rev. Lett.*, 2016, **116**, 061102, arXiv:1602.03837 [gr-qc].
6. Abbott, B. P. *et al.* (LIGO Scientific Collaboration, Virgo Collaboration). *Phys. Rev. Lett.*, 2016, **116**, 241103, arXiv:1606.04855 [gr-qc].
7. Abbott, B. P. *et al.* (VIRGO, LIGO Scientific). *Phys. Rev. Lett.*, 2017, **118**, 221101, arXiv:1706.01812 [gr-qc].
8. Abbott, B. P. *et al.* (LIGO Scientific Collaboration, Virgo Collaboration). *Phys. Rev.*, 2016, **X6**, 041015, arXiv:1606.04856 [gr-qc].
9. Abbott, B. P. *et al.* (LIGO Scientific Collaboration, Virgo Collaboration). *Astrophys. J. Lett.*, 2016, **818**, L22, arXiv:1602.03846 [astro-ph.HE].
10. Schutz, B., *A First Course in General Relativity*, Series in Physics, Cambridge University Press, 1985.
11. Abbott, B. P. *et al.* (LIGO Scientific Collaboration and Virgo Collaboration). *Phys. Rev. Lett.*, 2016, **116**, 131103, arXiv:1602.03838 [gr-qc].
12. Landau, L. D. and Lifshitz, E. M., *The Classical Theory of Fields: Course of Theoretical Physics*, Pergamon, Oxford, UK, 1971, 3rd edn (translated from Russian).
13. Dhurandhar, S. V. and Tinto, M., *MNRAS*, 1988, **234**, 663.
14. Dhurandhar, S. V. and Schutz, B. F., *Phys. Rev. D*, 1994, **50**, 2390.
15. Papoulis, A. and Pillai, S. U., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill Higher Education (a division of McGraw-Hill Companies), New York, USA, 2002, 4th edn, ISBN 978-0-07-366011-0.
16. Helstrom, C. W., *Statistical Theory of Signal Detection*, Pergamon Press, London, England, 1968, 2nd edn.
17. Balasubramanian, R., Sathyaprakash, B. S. and Dhurandhar, S. V., *Phys. Rev. D*, 1996, **53**, 3033, gr-qc/9508011.
18. Owen, B. J., *Phys. Rev. D*, 1996, **53**, 6749, gr-qc/9511032.
19. Sathyaprakash, B. S. and Dhurandhar, S. V., *Phys. Rev. D*, 1991, **44**, 3819.
20. Miller, M. C. and Miller, J. M., *Phys. Rep.*, 2014, **548**, 1, arXiv:1408.4145 [astro-ph.HE].
21. Belczynski, K., Bulik, T., Fryer, C. L., Ruiter, A., Vink, J. S. and Hurley, J. R., *Astrophys. J.*, 2010, **714**, 1217, arXiv:0904.2784 [astro-ph.SR].

-
22. Postnov, K. A. and Yungelson, L. R., *Living Rev. Relativ.*, 2014, **17**, 3, arXiv:1403.4754 [astro-ph.HE].
23. Rodriguez, C. L., Morscher, M., Pattabiraman, B., Chatterjee, S., Haster, C.-J. and Rasio, F. A., *Phys. Rev. Lett.*, 2015, **115**, 051101, arXiv:1505.00792 [astro-ph.HE].
24. Arun, K. G., Iyer, B. R., Qusailah, M. S. S. and Sathyaprakash, B. S., *Phys. Rev. D*, 2006, **74**, 024006, arXiv:gr-qc/0604067 [gr-qc].
25. Mishra, C. K., Arun, K. G., Iyer, B. R. and Sathyaprakash, B. S., *Phys. Rev. D*, 2010, **82**, 064010, arXiv:1005.0304 [gr-qc].
26. Blanchet, L., *Living Rev. Relativ.*, 2014, **17**, 2, arXiv:1310.1528 [gr-qc].
27. Li, T. G. F. *et al.*, *Phys. Rev. D*, 2012, **85**, 082003, arXiv:1110.0530 [gr-qc].
28. Khan, S., Husa, S., Hannam, M., Ohme, F., Pürrer, M., Jiménez Forteza, X. and Bohé, A., *Phys. Rev. D*, 2016, **93**, 044007, arXiv:1508.07253 [gr-qc].
29. Husa, S., Khan, S., Hannam, M., Pürrer, M., Ohme, F., Forteza, X. J. and Bohé, A., *Phys. Rev. D*, 2016, **93**, 044006, arXiv:1508.07250 [gr-qc].
30. Ghosh, A. *et al.*, *Phys. Rev. D*, 2016, **94**, 021101, arXiv:1602.02453 [gr-qc].
31. Mirshekari, S., Yunes, N. and Will, C. M., *Phys. Rev. D*, 2012, **85**, 024041, arXiv:1110.2720 [gr-qc].
32. Geiger, H. and Marsden, E., *Proc. R. Soc. London Ser.*, 1909, **82**, 495.
33. Rutherford, E., *Philos. Mag.*, 1991, **21**, 669.
34. Sathyaprakash, B. S. and Schutz, B. F., *Living Rev. Relativ.*, 2009, **12**, 10.12942/lrr-2009-2, arXiv:0903.0338 [gr-qc].
35. Flanagan, E. E. and Hinderer, T., *Phys. Rev. D*, 2008, **77**, 021502, arXiv:0709.1915 [astro-ph].
36. Kolb, E. W. and Tkachev, I. I., *Phys. Rev. Lett.*, 1993, **71**, 3051, arXiv:hep-ph/9303313 [hep-ph].
37. Mazur, P. O. and Mottola, E., *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 9545, arXiv:gr-qc/0407075 [gr-qc].
38. Dreyer, O., Kelly, B., Krishnan, B., Finn, L. S., Garrison, D. and Lopez-Aleman, R., *Classical Quant. Grav.*, 2004, **21**, 787, arXiv:gr-qc/0309007.
39. Vishveshwara, C. V., *Nature*, 1970, **227**, 936.
40. Berti, E., Cardoso, J., Cardoso, V. and Cavaglia, M., *Phys. Rev. D*, 2007, **76**, 104044, arXiv:0707.1202 [gr-qc].
41. Kamaretsos, I., Hannam, M., Husa, S. and Sathyaprakash, B., *Phys. Rev. D*, 2012, **85**, 024018, arXiv:1107.0854 [gr-qc].
42. Gossan, S., Veitch, J. and Sathyaprakash, B., *Phys. Rev. D*, 2012, **85**, 124056, arXiv:1111.5819 [gr-qc].
43. Will, C. M., *Living Rev. Relativ.*, 2014, **17**, 4, arXiv:1403.7377 [gr-qc].
44. Abbott, B. P. *et al.*, *Reports on Progress in Physics*, 2009, **72**, 076901.
45. Abbott, B. P. *et al.* (LIGO Scientific Collaboration, Virgo Collaboration). *Phys. Rev. Lett.*, 2016, **116**, 221101, arXiv:1602.03841 [gr-qc].
- doi:
-