IOP Concise Physics

Detecting the Stochastic Gravitational-Wave Background

Carlo Nicola Colacino

Chapter 1

A brief history of gravitational radiation

1.1 Introduction

More than 100 years ago, in 1915, Einstein published his general theory of relativity (GR). This theory changed dramatically our understanding of the gravitational force, or, to say it in modern parlance, interaction. Gravitation was no longer seen as a force, but rather as the curvature of spacetime itself. Newton's theory of gravitation, that had withstood the test of time for almost 300 years, became just an approximation of a radically different theory. It must be stressed that, although many physicists struggled to understand the novelty of Einstein's theory and Einstein himself did not win the Nobel Prize in Physics for GR—he was awarded the prize in 1921 for his pioneering work on quantum mechanics and his explanation of the photoelectric effect and, precisely because of the hostile reception by the physics community to his GR, did not collect the prize himself—100 years thereafter, Nature has always said a sound and clear yes to all predictions of GR. No experiment so far has even planted the smallest seed of doubt concerning the validity of GR. GR—the most beautiful of the physical theories because of its elegant and consistent mathematical formulation and richness—ranks amongst the most experimentally confirmed scientific theories, a milestone in our understanding of the Universe, its evolution and its structure. GR predicts the existence of ripples of spacetime that propagate with the speed of light, the so-called gravitational waves (GWs). This specific prediction was verified for the first time at the end of the last century in an indirect way: two radioastronomers, Hulse and Taylor, observed for more than 20 years the rotation period of the binary system PSR1913+16, a system made up of two neutron stars rotating around their centre of mass, and noticed that the period of rotation changed with time. They attributed this change to the energy loss due to the emission of GWs and showed that their data were in perfect agreement, within the experimental error, with the predictions of GR, actually one of the best agreements ever found in science between theory and data. Hulse and Taylor were awarded the Nobel Prize in 1993. On 11 February 2016 the Laser

doi:10.1088/978-1-6817-4082-9ch1

Interferometric Gravitational Observatory (LIGO) in the US and the European project Virgo announced the first direct detection of GWs, GW150914 [1]. This was an historical moment. Two massive black holes had merged and the event's gravitational-wave signal detected. The event that was announced on 11 February but had been recorded on the instruments on 14 September 2015, therefore 14 September has internationally become Gravitational-Wave Day. Three subsequent detections of the same kind of event have followed since: GW151226 [2], GW170104 [3] and GW170814 [4]. As a result of these detections the Physics Nobel Prize 2017 was awarded to three LIGO scientists: Kip S Thorne (Caltech), Rainer Weiss (MIT) and Barry C Barish (Caltech). Another breakthrough was about to come: on 16 October 2017 it was announced that thanks to triangulation made by the three detectors, a GW signal, GW170817, had been recorded and its source localised in the sky, which made it possible to observe the event with electromagnetic detectors as well. The conclusion was reached that the radiation had been produced by the merging not of black holes but of neutron stars. Multimessenger astronomy, i.e. combining GW and electromagnetic observations, was officially born on 17 August 2017, the day the signal was detected by the interferometers [5].

1.2 Sources of gravitational radiation

The gravitational radiation that was detected on 14 September and announced on 11 February was produced by the coalescence of a binary system made up of two supermassive black holes of roughly 36 and 29 solar masses, respectively. The signal was named **GW150914**, from the date it was observed. Such coalescences are described by scientists as a three-step process and gravitational radiation is produced in all three different phases: first, as the two objects rotate around one another there is the inward spiralling of the two compact objects; as they get closer and closer they reach the merging phase—mergers of compact objects are the most violent events in the history of the Universe after the Big Bang itself—and finally there is the ringdown of the single resulting black hole. The coalescence of a compact binary was the most likely candidate for the first detection and remains the main source of gravitational radiation in the eyes of scientists, both theoreticians and data-analysts alike. GR describes rather well the first and third stage, i.e. the spiralling and the ringdown. Theoretical waveforms are known for these two phases and can give information on a number of physical parameters such as the masses of the compact objects. The strong gravitational fields produced during the merging phase cannot be described too accurately by GR but can be tackled by using numerical relativity. Data analysis is greatly enhanced by the use of the matched filtering technique.

Hulse and Taylor's source was of a different kind: pulsars are stars that emit pulses at radio frequencies at very precise, regular intervals as they rotate. Again, the waveform is very well known from GR. It is just a sine wave in the Solar System barycentre, a coordinate system at rest with respect to the Sun. The orbital motion of the Earth around the Sun and the rotational motion of the Earth around its own axis modulate the frequency and this makes data analysis slightly more complicated than in the previous case of the compact binary coalescences.

Supernovae, or bursts, are the third kind of source, and perhaps the most difficult to detect, for two reasons. First, we have no precise waveform of the gravitational signal such events could produce, so we cannot rely on matched filtering. Furthermore, the effects of supernovae on our detectors might look remarkably similar to instrumental noise, so it is hard to tell them apart. Therefore, bursts are so hard to detect.

The stochastic gravitational-wave background is something entirely different. With the acronym SGWB we describe all the 'random' sources, random because they arise from a large number of 'unresolved', independent and uncorrelated events. We are not talking about quantum physics here: the individual events that make up the SGWB are perfectly deterministic. It is their unresolved superposition that produces a random signal, exactly as the superposition of many broadcasting stations in the same frequency band sums up to confusion noise on our radios. Such a signal can be treated and described only statistically. This type of background can be the result of processes that took place in the very early Universe, shortly after the Big Bang, or could have arisen more recently during structure formation. It is almost impossible to detect with present-day technology, nevertheless it is perhaps the most interesting source: from a theoretical point of view, as we have already mentioned, the cosmological background, the one produced immediately after formation, when the Universe was perhaps less than a picosecond old, could—if detected—shed light on all mysteries of cosmology as well as high-energy physics. Our knowledge about the Universe comes entirely from electromagnetic observations. Our most detailed view of the Universe stems from the cosmic microwave background radiation (CMBR), an isotropic electromagnetic radiation that decoupled from matter around 5×10^5 years after the Big Bang. We have no consistent and established quantum theory of the gravitational interaction, but reliable theoretical estimates seem to suggest that if current detectors reveal an SGWB component of cosmological origin, then it will carry with it a picture of the Universe as it was about 10^{-22} s after the Big Bang. That would be a tremendous leap forward in our knowledge and therefore scientists refer to the SGWB of cosmological origin as the 'Rosetta Stone' of cosmology and of particle physics. From a more practical point of view, we know that there is a stochastic component hidden in every data set we record with our detectors. When we look for any of the previously listed sources, we are really searching for a needle in a haystack. This is sometimes very rewarding, as in the case of the first detection, but it can also be very frustrating and is in any case computationally very intensive. On the other hand, as we will see in this book, any time two GW detectors operate and take data simultaneously, there will be an SGWB component in our data. This component might be very well hidden below the noise floor, but we know it is there and this knowledge lets us gain further insight into the SGWB.

We will start by recalling how GWs arise in GR, then calculate the radiation from a rotating binary source with given parameters. We will also rederive Hulse and Taylor's result for the 1913+16 binary pulsar. We will then proceed to characterise mathematically the SGWB, discussing cosmological as well as astrophysical sources. A large section will also be devoted to the data analysis problem for the SGWB.

1.3 Gravitational waves from general relativity

Einstein's GR is a very complex theory, where gravitation is seen not as a force but rather as the curvature of spacetime itself. The essence of GR can be summarised in the famous words by John Archibald Wheeler 'matter tells space how to curve, space tells matter how to move' (the sentence would be perfect, however, if the word 'space' were to be changed into 'spacetime'). The mathematical formulation of this idea is given by the GR field equations:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu} \tag{1.1}$$

$R_{\mu\nu}$ is the so-called Ricci tensor, obtained by contraction from the Riemann, or curvature, tensor $R_{\mu\nu} \equiv R^{\alpha}_{\ \mu\alpha\nu}$. R is the the Ricci scalar $R = g^{\mu\nu}R_{\mu\nu}$. These two quantities, well known in Riemannian geometry, are purely geometrical, they measure the intrinsic curvature of the spacetime. $T_{\mu\nu}$ is the covariant form of the energy–momentum tensor, i.e. a quantity related to matter. This distinction is, however, not so clear-cut, as can be seen by taking the trace of both members from which we can recast equation (1.1) as

$$R_{\mu\nu} = 8\pi G\left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T\right) \tag{1.2}$$

The equations are highly nonlinear in the ten unknowns $g_{\mu\nu}$.

GWs are solutions to these equations in the weak-field approximation: we need to expand Einstein's equations around the flat space Minkowski metric, therefore, we put

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{1.3}$$

with $|h_{\mu\nu}| \ll 1$ and expand equation (1.1) to linear order in $h_{\mu\nu}$, which we treat as a quantity which transforms as a tensor under Lorentz transformations. The resulting theory is called the *linearised theory*. To first order in $h_{\mu\nu}$ the Ricci tensor becomes

$$R^{(1)}_{\mu\nu} \sim \left(\partial_\lambda \Gamma^{\lambda}_{\mu\nu} - \partial_\nu \Gamma^{\lambda}_{\lambda\mu}\right) + \bigcirc(h^2) \tag{1.4}$$

where the Christoffel symbol, as it is called in the framework of GR, or the affine connection, as it is known by differential geometry mathematicians is

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2}\eta^{\lambda\alpha}(\partial_\mu h_{\alpha\nu} + \partial_\nu h_{\alpha\mu} - \partial_\alpha h_{\mu\nu}) + \bigcirc(h^2) \tag{1.5}$$

To first order in h, we must raise and lower all indices using the flat Minkowski tensor $\eta_{\mu\nu}$ and not the full tensor $g_{\mu\nu}$, that is,

$$h^{\mu}_{\nu} \equiv \eta^{\mu\varrho}h_{\varrho\nu} \qquad \partial^{\mu} \equiv \eta^{\mu\nu}\partial_\nu \qquad \Box^2 \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu \tag{1.6}$$

With this understanding, equations (1.4) and (1.5) yield the first-order Ricci tensor:

$$R^{(1)}_{\mu\nu} \equiv \frac{1}{2}(\partial_\lambda\partial_\mu h_\nu^\lambda + \partial_\lambda\partial_\nu h_\mu^\lambda - \Box^2 h_{\mu\nu} - \partial_\mu\partial_\nu h_\lambda^\lambda) \tag{1.7}$$

and equation (1.2) becomes

$$\partial_\lambda\partial_\mu h_\nu^\lambda + \partial_\lambda\partial_\nu h_\mu^\lambda - \Box^2 h_{\mu\nu} - \partial_\mu\partial_\nu h_\lambda^\lambda = 16\pi G S_{\mu\nu} \tag{1.8}$$

where

$$S_{\mu\nu} = T_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}T \tag{1.9}$$

As in electromagnetism, we can exploit a gauge symmetry of the linearised theory to get rid of spurious degrees of freedom. In fact, if we change coordinates

$$x^\mu \to x'^\mu = x^\mu + \xi^\mu \tag{1.10}$$

$h_{\mu\nu}$ transforms to lowest order as

$$h_{\mu\nu}(x) \to h'_{\mu\nu}(x') = h_{\mu\nu}(x) - (\partial_\mu\xi_\nu + \partial_\nu\xi_\mu) \tag{1.11}$$

so, if $|\partial_\mu\xi_\nu|$ is of the same order of magnitude as $|h_{\mu\nu}|$ then the condition $|h_{\mu\nu}| \ll 1$ is preserved and therefore this is a diffeomorphism of the theory.

We can choose a coordinate transformation such that

$$g^{\mu\nu}\Gamma^\lambda_{\mu\nu} = 0 \tag{1.12}$$

This is called the harmonic, or De Donder gauge. Some texts call it also Lorenz gauge in analogy to electromagnetism. To first order it implies that:

$$\partial_\mu h_\nu^\mu = \frac{1}{2}\partial_\nu h_\mu^\mu \tag{1.13}$$

In this gauge, the linearised field equations (1.8) become

$$\Box^2 h_{\mu\nu} = -16\pi G S_{\mu\nu} \tag{1.14}$$

The solution to this equation is given by the retarded potential:

$$h_{\mu\nu}(\bar{x}, t) = 4G\int d^4x' \frac{S_{\mu\nu}(\bar{x}', t)}{|x - x'|} \tag{1.15}$$

In order to understand the physical content of equation (1.14) let us solve it in the region far away from the source, i.e.

$$\Box^2 h_{\mu\nu} = 0 \tag{1.16}$$

The most general solutions to equation (1.16) can be written as plane-wave solutions:

$$h_{\mu\nu}(\vec{x}, t) = e_{\mu\nu} \exp\left[\mathrm{i}(\omega t - \vec{k} \cdot \vec{x})\right] + e^{*}_{\mu\nu}\exp\left[-\mathrm{i}(\omega t - \vec{k} \cdot \vec{x})\right] \tag{1.17}$$

This satisfies equation (1.16) if

$$k_\mu k^\mu = 0 \tag{1.18}$$

and equation (1.13) if

$$k_\mu e^\mu_\nu = \frac{1}{2} k_\mu e^\nu_\nu \tag{1.19}$$

The matrix $e_{\mu\nu}$ is obviously symmetric and is called the *polarisation tensor.*

This tensor is symmetric in its indices, so it has $n(n + 1)/2 = 4 \cdot 5/2 = 10$ independent components. The four relations (19) reduce this number to six, but of these six only two represent physically significant degrees of freedom. As we have seen, under a coordinate transformation $x^\mu \to x^\mu + \xi^\mu$ the metric $\eta_{\mu\nu} + h_{\mu\nu}$ transforms into a new metric $\eta_{\mu\nu} + h'_{\mu\nu}$, where $h'_{\mu\nu}$ is given by equation (1.11). If we choose:

$$\xi^\mu(x) = \mathrm{i}\xi^\mu \exp\left(\mathrm{i}k_\mu x^\mu\right) - \mathrm{i}\xi^{\mu *}\exp\left(-\mathrm{i}k_\mu x^\mu\right) \tag{1.20}$$

then equation (1.11) gives

$$h'_{\mu\nu} = e'_{\mu\nu} \exp(\mathrm{i}k_\lambda x^\lambda) + e'^{*}_{\mu\nu}\exp(-\mathrm{i}k_\lambda x^\lambda) \tag{1.21}$$

with

$$e'_{\mu\nu} = e_{\mu\nu} + k_\mu \xi_\nu + k_\nu \xi_\mu \tag{1.22}$$

It is easy to see that $e_{\mu\nu}$ and $e'_{\mu\nu}$ represent the very same physical situation for arbitrary values of the four parameters ξ^μ, so, of the six independent components of $e_{\mu\nu}$, only $6 - 4 = 2$ represent physical degrees of freedom.

To illustrate this point and to make our notation uniform to that of many books and articles on GWs, let us consider, as an example, a wave travelling in the z-direction, i.e. with wave vector

$$k^1 = k^2 = 0 \qquad k^3 = k^0 = k > 0 \tag{1.23}$$

In this case equation (1.19) gives

$$e_{31} + e_{01} = e_{32} + e_{02} = 0$$

$$e_{33} + e_{03} = -e_{03} - e_{00} = \frac{1}{2}(e_{11} + e_{22} + e_{33} - e_{00})$$

These equations allow us to express e_{i0} and e_{22} in terms of the other six $e_{\mu\nu}$:

$$e_{31} = -e_{01}; \qquad e_{02} = -e_{32}; \qquad e_{03} = \frac{1}{2}(e_{33} + e_{00}); \qquad e_{22} = -e_{11} \tag{1.24}$$

When the coordinate system transforms according to equation (1.20) these six independent components transform according to equation (1.22):

$$e'_{11} = e_{11} \qquad e'_{12} = e_{12}$$

$$e'_{13} = e_{13} + k\xi_1 \qquad e'_{23} = e_{23} + k\xi_2$$

$$e'_{33} = e_{33} + 2k\xi_1 \qquad e'_{00} = e_{00} - 2k\xi_0 \tag{1.25}$$

We see therefore that it is only e_{11} and e_{12} that have an absolute physical significance. We can arrange that all components of $e'_{\mu\nu}$ vanish except for e'_{11} and e'_{12} and $e'_{22} = -e'_{11}$ by performing a coordinate transformation with

$$\xi_1 = -\frac{e_{13}}{k}; \qquad \xi_2 = -\frac{e_{23}}{k}; \qquad \xi_3 = -\frac{e_{33}}{2k}; \qquad \xi_0 = \frac{e_{00}}{2k} \tag{1.26}$$

The distinction between the different components of the polarisation tensor becomes clear if we study how $e_{\mu\nu}$ changes when we subject the coordinate system to a rotation about the z-axis, a Lorentz transformation of the form:

$$R_1^{\ 1} = \cos\theta; \qquad R_1^{\ 2} = \sin\theta$$

$$R_2^{\ 1} = -\sin\theta; \qquad R_2^{\ 2} = \cos\theta$$

$$R_3^{\ 3} = R_0^{\ 0} = 1 \qquad \text{other } R_\nu^{\ \mu} = 0 \tag{1.27}$$

Since this transformation leaves k_μ invariant, the only effect is to transform $e_{\mu\nu}$ into

$$e'_{\mu\nu} = R_\mu^{\ \alpha} R_\nu^{\ \beta} e_{\alpha\beta} \tag{1.28}$$

Using the relations (24) we find that:

$$e'_\pm = \exp(\pm 2\mathrm{i}\theta) e_\pm \tag{1.29}$$

$$f'_\pm = \exp(\pm \mathrm{i}\theta) f_\pm \tag{1.30}$$

$$e'_{33} = e_{33}; \qquad e'_{00} = e_{00} \tag{1.31}$$

where

$$e_\pm \equiv e_{11} \mp \mathrm{i}e_{12} = -e_{22} \mp \mathrm{i}e_{12} \tag{1.32}$$

$$f_\pm \equiv e_{31} \mp \mathrm{i}e_{32} = -e_{01} \pm \mathrm{i}e_{12} \tag{1.33}$$

In general any plane wave ψ which is transformed by a rotation of any angle θ about the direction of propagation into

$$\psi' = \exp(\mathrm{i}h\theta)\,\psi \tag{1.34}$$

is said to have **helicity** h. We have shown that a GW can be decomposed into parts $e_\pm$ with helicity ± 2, parts $f_\pm$ with helicity ± 1 and parts e_{00} and e_{33} with helicity zero. However, it is *only the parts with helicity ± 2 that have physical significance.*

We can write the tensor $h_{\mu\nu}$ as

$$h_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_{11} & h_{12} & 0 \\ 0 & h_{12} & -h_{11} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.35}$$

This is for a wave travelling along the z-direction. $h_{11} = -h_{22}$ is also referred to in literature as h_+, whereas $h_{12} = -h_{21}$ is called $h_\times$.

We can understand this argument in greater depth by exploring the analogy with electromagnetism. Maxwell's equations in empty space, together with the harmonic gauge condition, are:

$$\Box^2 A_\mu = 0; \qquad \partial_\mu A^\mu = 0 \tag{1.36}$$

A plane-wave solution is of the form:

$$A_\alpha = e_\alpha \exp(\mathrm{i}k_\beta x^\beta) + e_\alpha^* \exp(-\mathrm{i}k_\beta x^\beta) \tag{1.37}$$

where

$$k_\alpha k^\alpha = 0 \tag{1.38}$$

$$k_\alpha e^\alpha = 0 \tag{1.39}$$

In general e_α would have four independent components but the condition (1.39) reduces this number to three, just as equation (1.19) would reduce the number of independent $e_{\mu\nu}$ components from ten to six. Furthermore, without changing the physical fields $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$ and without leaving the Lorenz gauge, we can change A_α by a gauge transformation:

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \Phi$$

$$\Phi(x) = \mathrm{i}\xi \exp(\mathrm{i}k_\lambda x^\lambda) - \mathrm{i}\xi^* \exp(-\mathrm{i}k_\lambda x^\lambda)$$

in analogy with equations (1.20) and (1.21). The new potential can be written:

$$A'_\alpha = e'_\alpha \exp\big(\mathrm{i}k_\beta x^\beta\big) + e'^*_\alpha \exp(-\mathrm{i}k_\beta x^\beta)$$

$$e'_\alpha = e_\alpha - \xi k_\alpha$$

The parameter ξ is arbitrary, so of the algebraically three independent components only $3 - 1 = 2$ are physically significant. To identify the two significant components of e_α we may consider a wave travelling in the z-direction, with k^α given by equation (1.23). The condition that $k_\mu e^\mu = 0$ allows us to determine e^0:

$$e_0 = -e_3$$

The preceding gauge transformation leaves e_1 and e_2 invariant, but changes e_3 into

$$e_3' = e_3 - \xi k$$

hence e_3' can be set equal to zero by choosing $\xi = e_3/k$ and so it is only e_1 and e_2 that carry physical significance. To work out the meaning of these two components we can subject the electromagnetic wave to the rotation defined by equation (1.27). The polarisation vector is then changed into:

$$e_\alpha' = R_\alpha^{\ \beta} e_\beta$$

and therefore

$$e_\pm' = \exp(\pm i\theta) e_\pm$$

$$e_3' = e_3$$

where

$$e_\pm \equiv e_1 \mp ie_2$$

Thus an electromagnetic wave can be decomposed into parts with helicity ± 1 and parts with helicity zero. However, the only physically significant ones are those with helicity ± 1, just as for GWs they are ± 2. This is what we mean when we say that electromagnetism and gravitation are carried by waves of spin 1 and spin 2, respectively.

It is a general fact that *waves that propagate with the speed of light have only two helicity states.* The hypotetical carrier of the gravitational interaction, the graviton, carries spin 2.

1.4 Background reading

My mathematical treatment of GR follows closely that of Steven Weinberg's *Gravitation and Cosmology* (Wiley 1972): although written 44 years ago, and with no exercises to solve, it remains a masterpiece, where the physical reasoning goes perfectly hand-in-hand with the mathematical formalism. Another excellent, and more modern book, about GR is *Introduction to General Relativity* by Lewis Ryder (Cambridge University Press 2009). As it concerns the SGWB, I am very grateful to Bruce Allen, whose lectures on the SGWB during a summer school back in 1999 in the beautiful location of Lake Como inspired me to devote my energies to the SGWB. Allen's lectures were published in the *Proceedings of the Les Houches on Astrophysical Sources of Gravitational Waves,* eds Jean-Alain Marck and Jean-Pierre Lesota (Cambridge University Press 1996).

References

[1] Abbott B P *et al* 2016 (LIGO Scientific Collaboration and Virgo Collaboration) Observation of gravitational waves from a binary black hole merger *Phys. Rev. Lett.* **116** 061102

[2] Abbott B P *et al* 2016 (LIGO Scientific Collaboration and Virgo Collaboration) GW151226: Observation of gravitational waves from a 22-solar-mass binary black hole coalescence *Phys. Rev. Lett.* **116** 241103

[3] Abbott B P *et al* 2017 (LIGO Scientific and Virgo Collaboration) GW170104: Observation of a 50-solar-mass binary black hole coalescence at redshift 0.2 *Phys. Rev. Lett.* **118** 221101

[4] Abbott B P *et al* 2017 (LIGO Scientific Collaboration and Virgo Collaboration) GW170814: A three-detector observation of gravitational waves from a binary black hole coalescence *Phys. Rev. Lett.* **119** 141101

[5] Abbott B P 2017 (LIGO Scientific Collaboration and Virgo Collaboration) *et al* GW170817: Observation of gravitational waves from a binary neutron star inspiral *Phys. Rev. Lett.* **119** 161101