The basics of gravitational wave theory

New Journal of Physics
The open-access journal for physics

The basics of gravitational wave theory

Éanna É Flanagan[1] **and Scott A Hughes**[2]

[1] Center for Radiophysics and Space Research, Cornell University, Ithaca, NY 14853, USA
[2] Department of Physics and Center for Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
E-mail: flanagan@astro.cornell.edu and sahughes@mit.edu

New Journal of Physics **7** (2005) 204
Received 11 January 2005
Published 29 September 2005
Online at http://www.njp.org/
doi:10.1088/1367-2630/7/1/204

Abstract. Einstein's special theory of relativity revolutionized physics by teaching us that space and time are not separate entities, but join as 'spacetime'. His general theory of relativity further taught us that spacetime is not just a stage on which dynamics takes place, but is a participant: the field equation of general relativity connects matter dynamics to the curvature of spacetime. Curvature is responsible for gravity, carrying us beyond the Newtonian conception of gravity that had been in place for the previous two and a half centuries. Much research in gravitation since then has explored and clarified the consequences of this revolution; the notion of dynamical spacetime is now firmly established in the toolkit of modern physics. Indeed, this notion is so well established that we may now contemplate using spacetime as a *tool* for other sciences. One aspect of dynamical spacetime—its radiative character, 'gravitational radiation'—will inaugurate entirely new techniques for observing violent astrophysical processes. Over the next 100 years, much of this subject's excitement will come from learning how to exploit spacetime as a tool for astronomy. This paper is intended as a tutorial in the basics of gravitational radiation physics.

PII: S1367-2630(05)92710-9

Contents

1. Introduction: spacetime and gravitational waves

Einstein's special relativity [1] taught us that space and time are not simply abstract, external concepts, but must in fact be considered measured observables, like any other quantity in physics. This reformulation enforced the philosophy that Newton sought to introduce in laying out his laws of mechanics [2]:

> . . . I frame no hypotheses; for whatever is not reduced from the phenomena is to be called an hypothesis; and hypotheses . . . have no place in experimental philosophy . . .

Despite his intention to stick only with that which can be observed, Newton described space and time using exactly the abstract notions that he otherwise deplored [3]:

> Absolute space, in its own nature, without relation to anything external, remains always similar and immovable
>
> Absolute, true, and mathematical time, of itself, and from its own nature, flows equably without relation to anything external.

Special relativity put an end to these abstractions: time is nothing more than that which is measured by clocks, and space is that which is measured by rulers. The properties of space and time thus depend on the properties of clocks and rulers. The constancy of the speed of light as measured by observers in different reference frames, as observed in the Michelson–Morley experiment, forces us inevitably to the fact that space and time are mixed into spacetime. Ten years after his paper on special relativity, Einstein endowed spacetime with curvature and made it dynamical [5]. This provided a *covariant* theory of gravity [6], in which all predictions for physical measurements are invariant under changes in coordinates. In this theory, general relativity, the notion of 'gravitational force' is reinterpreted in terms of the behaviour of geodesics in the curved manifold of spacetime.

To be compatible with special relativity, gravity must be causal: any change to a gravitating source must be communicated to distant observers no faster than the speed of light, c. This leads immediately to the idea that there must exist some notion of 'gravitational radiation'. As demonstrated by Schutz [7], one can actually calculate with surprising accuracy many of the properties of gravitational radiation simply by combining a time-dependent Newtonian potential with special relativity.

The first calculation of gravitational radiation in general relativity is due to Einstein. His initial calculation [8] was 'marred by an error in calculation' (Einstein's words), and was corrected in 1918 [9] (albeit with an overall factor of two error). Modulo a somewhat convoluted history (discussed in great detail by Kennefick [10]) owing (largely) to the difficulties of analysing radiation in a nonlinear theory, Einstein's final result stands today as the leading-order 'quadrupole formula' for gravitational wave emission. This formula plays a role in gravity theory analogous to the dipole formula for electromagnetic radiation, showing that gravitational waves (hereafter abbreviated GWs) arise from accelerated masses exactly as electromagnetic waves arise from accelerated charges.

The quadrupole formula tells us that GWs are difficult to produce—very large masses moving at relativistic speeds are needed. This follows from the weakness of the gravitational interaction. A consequence of this is that it is *extremely* unlikely there will ever be an interesting laboratory source of GWs. The only objects massive and relativistic enough to generate detectable GWs are astrophysical. Indeed, experimental confirmation of the existence of GWs has come from the study of binary neutron star systems—the variation of the mass quadrupole in such systems is large enough that GW emission changes the system's characteristics on a timescale short enough to be observed. The most celebrated example is the 'Hulse–Taylor' pulsar, B1913+16, reported by Hulse and Taylor in 1975 [11]. Thirty years of observation have shown that the orbit is decaying; the results match with extraordinary precision general relativity's prediction for such a decay due to the loss of orbital energy and angular momentum by GWs. For a summary of the most recent data, see figure 1 of [12]. Hulse and Taylor were awarded the Nobel Prize for this

discovery in 1993. Since this pioneering system was discovered, several other double neutron star systems 'tight' enough to exhibit strong GW emission have been discovered [13]–[16].

Studies of these systems prove beyond a reasonable doubt that GWs exist. What remains is to detect the waves directly and exploit them—to *use* GWs as a way to study astrophysical objects. The contribution to this Focus Issue by Aufmuth and Danzmann [21] discusses the challenges and the method of directly measuring these waves. Intuitively, it is clear that measuring these waves must be difficult—the weakness of the gravitational interaction ensures that the response of any detector to gravitational waves is very small. Nonetheless, technology has brought us to the point where detectors are now beginning to set interesting upper limits on GWs from some sources [17]–[20]. The first direct detection is now hopefully not too far in the future.

The real excitement will come when detection becomes routine. We will then have an opportunity to retool the 'physics experiment' of direct detection into the development of astronomical observatories. Some of the papers appearing in this volume will discuss likely future revolutions which, at least conceptually, should change our notions of spacetime in a manner as fundamental as Einstein's works in 1905 and 1915 (see, e.g., papers in this Focus Issue by Ashtekar and Horowitz). Such a revolution is unlikely to be driven by GW observations—most of what we expect to learn using GWs will apply to regions of spacetime that are well-described using classical general relativity; it is highly unlikely that Einstein's theory will need major revisions prompted by GW observations. Any revolution arising from GW science will instead be in astrophysics: mature GW measurements have the potential to study regions of the Universe that are currently inaccessible to our instruments. During the next century of spacetime study, spacetime will itself be exploited to study our Universe.

1.1. Why this paper?

As GW detectors have improved and approached maturity, many papers have been written reviewing this field and its promise. One might ask: do we really need another one? As a partial answer to this question, we note that R. Price requested this paper very nicely. More seriously, our goal is to provide a brief tutorial on the basics of GW science, rather than a comprehensive survey of the field. The reader who is interested in such a survey can find them in [22]–[30]. Other reviews on the basics of GW science can be found in [31, 32]; we also recommend the dedicated conference procedings [33]–[35].

We assume that the reader has a basic familiarity with general relativity, at least at the level of Hartle's textbook [36]; thus, we assume the reader understands metrics and is reasonably comfortable taking covariant derivatives. We adapt what Baumgarte and Shapiro [37] call the 'Fortran convention' for indices: $a, b, c \ldots h$ denote spacetime indices which run over 0, 1, 2, 3 or t, x, y, z, while $i, j, k, \ldots, n$ denote spatial indices which run over 1, 2, 3. We use the Einstein summation convention throughout—repeated adjacent indices in superscript and subscript positions imply a sum:

$$u^a v_a \equiv \sum_a u^a v_a.$$

When we discuss linearized theory, we will sometimes be sloppy and sum over adjacent spatial indices in the same position. Hence,

$$u_i v_i \equiv u^i v^i \equiv u^i v_i \equiv \sum_i u^i v_i$$

is valid in linearized theory. (As we will discuss in section 2, this is allowable because, in linearized theory, the position of a spatial index is immaterial in Cartesian coordinates.) A quantity that is symmetrized on pairs of indices is written as

$$A_{(ab)} = \tfrac{1}{2}(A_{ab} + A_{ba}).$$

Throughout most of this paper, we use 'relativist's units', in which $G = 1 = c$; mass, space and time have the same units in this system. The following conversion factors are often useful for converting to 'normal' units:

$$\begin{aligned} 1\text{second} &= 299\,792\,458\,\text{m} \simeq 3 \times 10^8\,\text{m} \\ 1M_\odot &= 1476.63\,\text{m} \simeq 1.5\,\text{km} \\ &= 4.92549 \times 10^{-6}\,\text{seconds} \simeq 5\,\mu\text{seconds}. \end{aligned}$$

($1M_\odot$ is one solar mass.) We occasionally restore factors of G and c to write certain formulae in normal units.

Section 2 provides an introduction to linearized gravity, deriving the most basic properties of GWs. Our treatment in this section is mostly standard. One aspect of our treatment that is slightly unusual is that we introduce a gauge-invariant formalism that fully characterizes the linearized gravity's degrees of freedom. We demonstrate that the linearized Einstein equations can be written as five Poisson-type equations for certain combinations of the spacetime metric, plus a wave equation for the transverse-traceless components of the metric perturbation. This analysis helps to clarify which degrees of freedom in general relativity are radiative and which are not, a useful exercise for understanding spacetime dynamics.

Section 3 analyses the interaction of GWs with detectors whose sizes are small compared to the wavelength of the GWs. This includes ground-based interferometric and resonant-mass detectors, but excludes space-based interferometric detectors. The analysis is carried out in two different gauges; identical results are obtained from both analyses. Section 4 derives the leading-order formula for radiation from slowly moving, weakly self-gravitating sources, the quadrupole formula discussed above.

In section 5, we develop linearized theory on a curved background spacetime. Many of the results of 'basic linearized theory' (section 2) carry over with slight modification. We introduce the 'geometric optics' limit in this section, and sketch the derivation of the Isaacson stress–energy tensor, demonstrating how GWs carry energy and curve spacetime. Section 6 provides a very brief synopsis of GW astronomy, leading the reader through a quick tour of the relevant frequency bands and anticipated sources. We conclude by discussing very briefly some topics that we could not cover in this paper, with pointers to good reviews.

2. The basic basics: gravitational waves in linearized gravity

The most natural starting point for any discussion of GWs is 'linearized gravity'. Linearized gravity is an adequate approximation to general relativity when the spacetime metric, g_{ab}, may be treated as deviating only slightly from a flat metric, η_{ab}:

$$g_{ab} = \eta_{ab} + h_{ab}, \qquad \|h_{ab}\| \ll 1. \tag{2.1}$$

Here η_{ab} is defined to be diag$(-1, 1, 1, 1)$ and $\|h_{ab}\|$ means 'the magnitude of a typical non-zero component of h_{ab}'. Note that the condition $\|h_{ab}\| \ll 1$ requires both the gravitational field to

be weak, and in addition constrains the coordinate system to be approximately Cartesian. We will refer to h_{ab} as the metric perturbation; as we will see, it encapsulates GWs, but contains additional, non-radiative degrees of freedom as well. In linearized gravity, the smallness of the perturbation means that we only keep terms which are linear in h_{ab}—higher order terms are discarded. As a consequence, indices are raised and lowered using the flat metric η_{ab}. The metric perturbation h_{ab} transforms as a tensor under Lorentz transformations, but not under general coordinate transformations.

We now compute all the quantities which are needed to describe linearized gravity. The components of the affine connection (Christoffel coefficients) are given by

$$\Gamma^a{}_{bc} = \tfrac{1}{2}\eta^{ad}(\partial_c h_{db} + \partial_b h_{dc} - \partial_d h_{bc}) = \tfrac{1}{2}(\partial_c h^a{}_b + \partial_b h^a{}_c - \partial^a h_{bc}). \tag{2.2}$$

Here ∂_a means the partial derivative $\partial/\partial x^a$. Since we use η_{ab} to raise and lower indices, spatial indices can be written either in the 'up' position or the 'down' position without changing the value of a quantity: $f^x = f_x$. Raising or lowering a time index, by contrast, switches sign: $f^t = -f_t$. The Riemann tensor we construct in linearized theory is then given by

$$R^a{}_{bcd} = \partial_c \Gamma^a{}_{bd} - \partial_d \Gamma^a{}_{bc} = \tfrac{1}{2}(\partial_c \partial_b h^a{}_d + \partial_d \partial^a h_{bc} - \partial_c \partial^a h_{bd} - \partial_d \partial_b h^a{}_c). \tag{2.3}$$

From this, we construct the Ricci tensor

$$R_{ab} = R^c{}_{acb} = \tfrac{1}{2}(\partial_c \partial_b h^c{}_a + \partial^c \partial_a h_{bc} - \Box h_{ab} - \partial_a \partial_b h), \tag{2.4}$$

where $h = h^a{}_a$ is the trace of the metric perturbation and $\Box = \partial_c \partial^c = \nabla^2 - \partial_t^2$ is the wave operator. Contracting once more, we find the curvature scalar:

$$R = R^a{}_a = (\partial_c \partial^a h^c{}_a - \Box h) \tag{2.5}$$

and finally build the Einstein tensor:

$$G_{ab} = R_{ab} - \tfrac{1}{2}\eta_{ab} R = \tfrac{1}{2}(\partial_c \partial_b h^c{}_a + \partial^c \partial_a h_{bc} - \Box h_{ab} - \partial_a \partial_b h - \eta_{ab} \partial_c \partial^d h^c{}_d + \eta_{ab} \Box h). \tag{2.6}$$

This expression is a bit unwieldy. Somewhat remarkably, it can be cleaned up significantly by changing the notation: rather than working with the metric perturbation h_{ab}, we use the *trace-reversed* perturbation $\bar{h}_{ab} = h_{ab} - \frac{1}{2}\eta_{ab} h$. (Notice that $\bar{h}^a{}_a = -h$, hence the name 'trace reversed'.) Replacing h_{ab} with $\bar{h}_{ab} + \frac{1}{2}\eta_{ab} h$ in equation (2.6) and expanding, we find that all terms with the trace h are cancelled. What remains is

$$G_{ab} = \tfrac{1}{2}(\partial_c \partial_b \bar{h}^c{}_a + \partial^c \partial_a \bar{h}_{bc} - \Box \bar{h}_{ab} - \eta_{ab} \partial_c \partial^d \bar{h}^c{}_d). \tag{2.7}$$

This expression can be simplified further by choosing an appropriate coordinate system, or *gauge*. Gauge transformations in general relativity are just coordinate transformations. A general infinitesimal coordinate transformation can be written as $x^{a'} = x^a + \xi^a$, where $\xi^a(x^b)$ is an arbitrary infinitesimal vector field. This transformation changes the metric via

$$h'_{ab} = h_{ab} - 2\partial_{(a}\xi_{b)}, \tag{2.8}$$

so that the trace-reversed metric becomes

$$\bar{h}'_{ab} = h'_{ab} - \tfrac{1}{2}\eta_{ab}h' = \bar{h}_{ab} - 2\partial_{(b}\xi_{a)} + \eta_{ab}\partial^c\xi_c. \tag{2.9}$$

A class of gauges that are commonly used in studies of radiation are those satisfying the Lorentz gauge condition

$$\partial^a\bar{h}_{ab} = 0. \tag{2.10}$$

(Note the close analogy to Lorentz gauge[3] in electromagnetic theory, $\partial^a A_a = 0$, where A_a is the potential vector.)

Suppose that our metric perturbation is not in Lorentz gauge. What properties must ξ_a satisfy in order to impose Lorentz gauge? Our goal is to find a new metric h'_{ab} such that $\partial^a\bar{h}'_{ab} = 0$:

$$\partial^a\bar{h}'_{ab} = \partial^a\bar{h}_{ab} - \partial^a\partial_b\xi_a - \Box\xi_b + \partial_b\partial^c\xi_c \tag{2.11}$$

$$= \partial^a\bar{h}_{ab} - \Box\xi_b. \tag{2.12}$$

Any metric perturbation h_{ab} can therefore be put into a Lorentz gauge by making an infinitesimal coordinate transformation that satisfies

$$\Box\xi_b = \partial^a\bar{h}_{ab}. \tag{2.13}$$

One can always find solutions to the wave equation (2.13), thus achieving Lorentz gauge. The amount of gauge freedom has now been reduced from four freely specifiable functions of four variables to four functions of four variables that satisfy the homogeneous wave equation $\Box\xi^b = 0$, or, equivalently, to eight freely specifiable functions of three variables on an initial data hypersurface.

Applying the Lorentz gauge condition (2.10) to the expression (2.7) for the Einstein tensor, we find that all but one term vanishes:

$$G_{ab} = -\tfrac{1}{2}\Box\bar{h}_{ab}. \tag{2.14}$$

Thus, in Lorentz gauges, the Einstein tensor simply reduces to the wave operator acting on the trace-reversed metric perturbation (up to a factor $-1/2$). The linearized Einstein equation is therefore

$$\Box\bar{h}_{ab} = -16\pi T_{ab}; \tag{2.15}$$

in vacuum, this reduces to

$$\Box\bar{h}_{ab} = 0. \tag{2.16}$$

Just as in electromagnetism, the equation (2.15) admits a class of homogeneous solutions which are superpositions of plane waves:

$$\bar{h}_{ab}(\mathbf{x}, t) = \mathrm{Re}\int \mathrm{d}^3k A_{ab}(\mathbf{k})\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t)}. \tag{2.17}$$

Here, $\omega = |\mathbf{k}|$. The complex coefficients $A_{ab}(\mathbf{k})$ depend on the wavevector $\mathbf{k}$ but are independent of $\mathbf{x}$ and t. They are subject to the constraint $k^a A_{ab} = 0$ (which follows from the Lorentz gauge condition), with $k^a = (\omega, \mathbf{k})$, but are otherwise arbitrary. These solutions are gravitational waves.

[3] Fairly recently, it has become widely recognized that this gauge was in fact invented by Ludwig Lorenz, rather than by Hendrik Lorentz. The inclusion of the 't' seems most likely due to confusion between the similar names; see [38] for a detailed discussion. Following the practice of Griffiths ([39], p 421), we bow to the weight of historical usage in order to avoid any possible confusion.

2.1. Globally vacuum spacetimes: transverse traceless (TT) gauge

We now specialize to globally vacuum spacetimes in which $T_{ab} = 0$ everywhere, and which are asymptotically flat (for our purposes, $h_{ab} \to 0$ as $r \to \infty$). Equivalently, we specialize to the space of homogeneous, asymptotically flat solutions of the linearized Einstein equation (2.15). For such spacetimes one can, along with choosing Lorentz gauge, further specialize the gauge to make the metric perturbation be purely spatial

$$h_{tt} = h_{ti} = 0 \tag{2.18}$$

and traceless

$$h = h_i{}^i = 0. \tag{2.19}$$

The Lorentz gauge condition (2.10) then implies that the spatial metric perturbation is transverse:

$$\partial_i h_{ij} = 0. \tag{2.20}$$

This is called the transverse-traceless gauge, or TT gauge. A metric perturbation that has been put into TT gauge will be written as h^{TT}_{ab}. Since it is traceless, there is no distinction between h^{TT}_{ab} and $\bar{h}^{\mathrm{TT}}_{ab}$.

The conditions (2.18) and (2.19) comprise five constraints on the metric, while the residual gauge freedom in Lorentz gauge is parametrized by four functions that satisfy the wave equation. It is nevertheless possible to satisfy these conditions, essentially because the metric perturbation satisfies the linearized vacuum Einstein equation. When the TT gauge conditions are satisfied the gauge is completely fixed.

One might wonder *why* we would choose TT gauge. It is certainly not necessary; however, it is extremely *convenient*, since the TT gauge conditions completely fix all the local gauge freedom. The metric perturbation h^{TT}_{ab} therefore contains only physical, non-gauge information about the radiation. In TT gauge, there is a close relation between the metric perturbation and the linearized Riemann tensor R_{abcd} (which is invariant under the local gauge transformations (2.8) by equation (2.3)), namely

$$R_{itjt} = -\tfrac{1}{2}\ddot{h}^{\mathrm{TT}}_{ij}. \tag{2.21}$$

In a globally vacuum spacetime, all non-zero components of the Riemann tensor can be obtained from R_{itjt} via Riemann's symmetries and the Bianchi identity. In a more general spacetime, there will be components that are not related to radiation; this point is discussed further in section 2.2.

Transverse traceless gauge also exhibits the fact that gravitational waves have two polarization components. For example, consider a GW which propagates in the z direction: $h^{\mathrm{TT}}_{ij} = h^{\mathrm{TT}}_{ij}(t - z)$ is a valid solution to the wave equation $\Box h^{\mathrm{TT}}_{ij} = 0$. The Lorentz condition $\partial_z h^{\mathrm{TT}}_{zj} = 0$ implies that $h^{\mathrm{TT}}_{zj}(t - z) =$ constant. This constant must be zero to satisfy the condition that $h_{ab} \to 0$ as $r \to \infty$. The only non-zero components of h^{TT}_{ij} are then h^{TT}_{xx}, h^{TT}_{xy}, h^{TT}_{yx} and h^{TT}_{yy}. Symmetry and the tracefree condition (2.19) further mandate that only two of these are independent:

$$h^{\mathrm{TT}}_{xx} = -h^{\mathrm{TT}}_{yy} \equiv h_+(t - z); \tag{2.22}$$

$$h^{\mathrm{TT}}_{xy} = h^{\mathrm{TT}}_{yx} \equiv h_{\times}(t-z). \tag{2.23}$$

The quantities h_+ and $h_\times$ are the two independent waveforms of the GW (see figure 1).

For globally vacuum spacetimes, one can always satisfy the TT gauge conditions. To see this, note that the most general gauge transformation ξ^a that preserves the Lorentz gauge condition (2.10) satisfies $\Box\xi^a = 0$, from equation (2.12). A general solution to this equation can be written as

$$\xi^a = \mathrm{Re}\int \mathrm{d}^3k C^a(\mathbf{k})\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t)} \tag{2.24}$$

for some coefficients $C^a(\mathbf{k})$. Under this transformation the tensor $A_{ab}(\mathbf{k})$ in equation (2.17) transforms as

$$A_{ab} \to A'_{ab} = A_{ab} - 2\mathrm{i}k_{(a}C_{b)} + \mathrm{i}\eta_{ab}k^dC_d. \tag{2.25}$$

Achieving the TT gauge conditions (2.20) and (2.19) therefore requires finding, for each $\mathbf{k}$, a $C^a(\mathbf{k})$ that satisfies the two equations

$$0 = \eta^{ab}A'_{ab} = \eta^{ab}A_{ab} + 2\mathrm{i}k^aC_a, \tag{2.26}$$

$$0 = A'_{tb} = A_{tb} - \mathrm{i}C_bk_t - \mathrm{i}C_tk_b + \mathrm{i}\delta^t{}_b(k^aC_a); \tag{2.27}$$

$\delta^t{}_b$ is the Kronecker delta—zero for $b \neq t$, unity otherwise. An explicit solution to these equations is given by

$$C_a = \frac{3A_{bc}l^bl^c}{8\mathrm{i}\omega^4}k_a + \frac{\eta^{bc}A_{bc}}{4\mathrm{i}\omega^4}l_a + \frac{1}{2\mathrm{i}\omega^2}A_{ab}l^b, \tag{2.28}$$

where $k^a = (\omega, \mathbf{k})$ and $l^a = (\omega, -\mathbf{k})$.

2.2. *Global spacetimes with matter sources*

We now return to the more general and realistic situation in which the stress–energy tensor is non-zero. We continue to assume that the linearized Einstein equations are valid everywhere in spacetime and that we consider asymptotically flat solutions only. In this context, the metric perturbation h_{ab} contains (i) gauge degrees of freedom; (ii) physical, radiative degrees of freedom and (iii) physical, non-radiative degrees of freedom tied to the matter sources. Because of the presence of the physical, non-radiative degrees of freedom, it is not possible in general to write the metric perturbation in TT gauge. However, the metric perturbation can be split up uniquely into various pieces that correspond to the degrees of freedom (i), (ii) and (iii), and the radiative degrees of freedom correspond to a piece of the metric perturbation that satisfies the TT gauge conditions, the so-called TT piece.

This aspect of linearized theory is obscured by the standard, Lorentz gauge formulation (2.15) of the linearized Einstein equations. There, all the components of h_{ab} appear to be radiative, since all the components obey wave equations. In this subsection, we describe a formulation of linearized theory which focuses on gauge-invariant observables. In particular, we will see that

only the TT part of the metric obeys a wave equation in all gauges. We show that the non-TT parts of the metric can be gathered into a set of gauge-invariant functions; these functions are governed by Poisson equations rather than wave equations. This shows that the *non-TT pieces of the metric do not exhibit radiative degrees of freedom*. Although one can always choose gauges like Lorentz gauge in which the non-radiative parts of the metric obey wave equations and thus *appear* to be radiative, this appearance is a gauge artifact. Such gauge choices, although useful for calculations, can cause one to mistake purely gauge modes for a truly physical radiation.

Interestingly, the first analysis contrasting physical radiative degrees of freedom from purely coordinate modes appears to have been performed by Eddington in 1922 [40]. Eddington was somewhat suspicious of Einstein's analysis [9], as Einstein chose a gauge in which all metric functions propagated with the speed of light. Though the entire metric appeared to be radiative (by construction), Einstein found that *only* the 'transverse–transverse' pieces of the metric carried energy. Eddington wrote:

> Weyl has classified plane gravitational waves into three types, viz. (1) longitudinal-longitudinal; (2) longitudinal-transverse; (3) transverse-transverse. The present investigation leads to the conclusion that transverse-transverse waves are propagated with the speed of light *in all systems of co-ordinates*. Waves of the first and second types have no fixed velocity—a result which rouses suspicion as to their objective existence. Einstein had also become suspicious of these waves (in so far as they occur in his special co-ordinate system) for another reason, because he found that they convey no energy. They are not objective, and (like absolute velocity) are not detectable by any conceivable experiment. They are merely sinuosities in the co-ordinate system, and the only speed of propagation relevant to them is "the speed of thought." . . . It is evidently a great convenience in analysis to have all waves, both physical and spurious, travelling with one velocity; but it is liable to obscure physical ideas by mixing them up so completely. The chief new point in the present discussion is that when unrestricted co-ordinates are allowed the genuine waves continue to travel with the velocity of light and the spurious waves cease to have any fixed velocity.

Unfortunately, Eddington's wry dismissal of unphysical modes as propagating with 'the speed of thought' is often taken by skeptics (and crackpots) as applying to *all* gravitational perturbations. Eddington in fact showed quite the opposite. We do so now using a somewhat more modern notation; our presentation is essentially the flat-spacetime limit of Bardeen's [41] gauge-invariant cosmological perturbation formalism. A similar treatment can be found in the lecture notes by Bertschinger [42].

We begin by defining the decomposition of the metric perturbation h_{ab}, in any gauge, into a number of irreducible pieces. Assuming that $h_{ab} \to 0$ as $r \to \infty$, we define the quantities ϕ, β_i, γ, H, ε_i, λ and h_{ij}^{TT} via the equations

$$h_{tt} = 2\phi, \tag{2.29}$$

$$h_{ti} = \beta_i + \partial_i \gamma, \tag{2.30}$$

$$h_{ij} = h_{ij}^{\mathrm{TT}} + \tfrac{1}{3} H \delta_{ij} + \partial_{(i}\varepsilon_{j)} + \left(\partial_i \partial_j - \tfrac{1}{3}\delta_{ij}\nabla^2\right)\lambda, \tag{2.31}$$

together with the constraints

$$\partial_i \beta_i = 0 \quad \text{(one constraint)}, \tag{2.32}$$

$$\partial_i \varepsilon_i = 0 \quad \text{(one constraint)}, \tag{2.33}$$

$$\partial_i h_{ij}^{\mathrm{TT}} = 0 \quad \text{(three constraints)}, \tag{2.34}$$

$$\delta^{ij} h_{ij}^{\mathrm{TT}} = 0 \quad \text{(one constraint)}. \tag{2.35}$$

and boundary conditions

$$\gamma \to 0, \quad \varepsilon_i \to 0, \quad \lambda \to 0, \quad \nabla^2 \lambda \to 0 \tag{2.36}$$

as $r \to \infty$. Here $H \equiv \delta^{ij} h_{ij}$ is the trace of the *spatial* portion of the metric perturbation, not to be confused with the spacetime trace $h = \eta^{ab} h_{ab}$ that we used earlier. The spatial tensor h_{ij}^{TT} is transverse and traceless, and is the TT piece of the metric discussed above which contains the physical radiative degrees of freedom. The quantities β_i and $\partial_i \gamma$ are the transverse and longitudinal pieces of h_{ti}. The uniqueness of this decomposition follows from taking a divergence of equation (2.30) giving $\nabla^2 \gamma = \partial_i h_{ti}$, which has a unique solution by the boundary condition (2.36). Similarly, taking two derivatives of equation (2.31) yields the equation $2\nabla^2 \nabla^2 \lambda = 3\partial_i \partial_j h_{ij} - \nabla^2 H$, which has a unique solution by equation (2.36). Having solved for λ, one can obtain a unique ε_i by solving $3\nabla^2 \varepsilon_i = 6\partial_j h_{ij} - 2\partial_i H - 4\partial_i \nabla^2 \lambda$.

The total number of free functions in the parametrization (2.29)–(2.31) of the metric is 16: four scalars (ϕ, γ, H and λ), six vector components (β_i and ε_i) and six symmetric tensor components (h_{ij}^{TT}). The number of constraints (2.32)–(2.35) is six, so the number of independent variables in the parametrization is 10, consistent with a symmetric 4×4 tensor.

We next discuss how the variables ϕ, β_i, γ, H, ε_i, λ and h_{ij}^{TT} transform under gauge transformations ξ^a with $\xi^a \to 0$ as $r \to \infty$. We parametrize such gauge transformation as

$$\xi_a = (\xi_t, \xi_i) \equiv (A, B_i + \partial_i C), \tag{2.37}$$

where $\partial_i B_i = 0$ and $C \to 0$ as $r \to \infty$; thus B_i and $\partial_i C$ are the transverse and longitudinal pieces of the spatial gauge transformation. The transformed metric is $h_{ab} - 2\partial_{(a}\xi_{b)}$; decomposing this transformed metric into its irreducible pieces yields the transformation laws

$$\phi \to \phi - \dot{A}, \tag{2.38}$$

$$\beta_i \to \beta_i - \dot{B}_i, \tag{2.39}$$

$$\gamma \to \gamma - A - \dot{C}, \tag{2.40}$$

$$H \to H - 2\nabla^2 C, \tag{2.41}$$

$$\lambda \to \lambda - 2C, \tag{2.42}$$

$$\varepsilon_i \to \varepsilon_i - 2B_i, \tag{2.43}$$

$$h_{ij}^{\mathrm{TT}} \to h_{ij}^{\mathrm{TT}}. \tag{2.44}$$

Gathering terms, we see that the following combinations of these functions are gauge-invariant:

$$\Phi \equiv -\phi + \dot{\gamma} - \tfrac{1}{2}\ddot{\lambda}, \tag{2.45}$$

$$\Theta \equiv \tfrac{1}{3}(H - \nabla^2\lambda), \tag{2.46}$$

$$\Xi_i \equiv \beta_i - \tfrac{1}{2}\dot{\varepsilon}_i; \tag{2.47}$$

h_{ij}^{TT} is gauge-invariant without any further manipulation. In the Newtonian limit, Φ reduces to the Newtonian potential Φ_N, while $\Theta = -2\Phi_N$. The total number of free, gauge-invariant functions is six: one function Θ; one function Φ; three functions Ξ_i, minus one due to the constraint $\partial_i \Xi_i = 0$; and six functions h_{ij}^{TT}, minus three due to the constraints $\partial_i h_{ij}^{\mathrm{TT}} = 0$, minus one due to the constraint $\delta^{ij} h_{ij}^{\mathrm{TT}} = 0$. This is in keeping with the fact that in general the 10 metric functions contain six physical and four gauge degrees of freedom.

We would now like to enforce Einstein's equation. Before doing so, it is useful to first decompose the stress–energy tensor in a manner similar to that of our decomposition of the metric. We define the quantities ρ, S_i, S, P, σ_{ij}, σ_i and σ via the equations

$$T_{tt} = \rho, \tag{2.48}$$

$$T_{ti} = S_i + \partial_i S, \tag{2.49}$$

$$T_{ij} = P\delta_{ij} + \sigma_{ij} + \partial_{(i}\sigma_{j)} + (\partial_i\partial_j - \tfrac{1}{3}\delta_{ij}\nabla^2)\sigma, \tag{2.50}$$

together with the constraints

$$\partial_i S_i = 0, \tag{2.51}$$

$$\partial_i \sigma_i = 0, \tag{2.52}$$

$$\partial_i \sigma_{ij} = 0, \tag{2.53}$$

$$\delta^{ij}\sigma_{ij} = 0, \tag{2.54}$$

and boundary conditions

$$S \to 0, \qquad \sigma_i \to 0, \qquad \sigma \to 0, \qquad \nabla^2\sigma \to 0 \tag{2.55}$$

as $r \to \infty$. These quantities are not all independent. The variables ρ, P, S_i and σ_{ij} can be specified arbitrarily; stress–energy conservation ($\partial^a T_{ab} = 0$) then determines the remaining variables S, σ and σ_i via

$$\nabla^2 S = \dot{\rho}, \tag{2.56}$$

$$\nabla^2\sigma = -\tfrac{3}{2}P + \tfrac{3}{2}\dot{S}, \tag{2.57}$$

$$\nabla^2\sigma_i = 2\dot{S}_i. \tag{2.58}$$

We now compute the Einstein tensor from the metric (2.29)–(2.31). The result can be expressed in terms of the gauge-invariant observables:

$$G_{tt} = -\nabla^2\Theta, \tag{2.59}$$

$$G_{ti} = -\tfrac{1}{2}\nabla^2\Xi_i - \partial_i\dot{\Theta}, \tag{2.60}$$

$$G_{ij} = -\tfrac{1}{2}\Box h_{ij}^{\mathrm{TT}} - \partial_{(i}\dot{\Xi}_{j)} - \tfrac{1}{2}\partial_i\partial_j(2\Phi + \Theta) + \delta_{ij}[\tfrac{1}{2}\nabla^2(2\Phi + \Theta) - \ddot{\Theta}]. \tag{2.61}$$

We finally enforce Einstein's equation $G_{ab} = 8\pi T_{ab}$ and simplify using the conservation relations (2.56)–(2.58); this leads to the following field equations:

$$\nabla^2\Theta = -8\pi\rho, \tag{2.62}$$

$$\nabla^2\Phi = 4\pi(\rho + 3P - 3\dot{S}), \tag{2.63}$$

$$\nabla^2\Xi_i = -16\pi S_i, \tag{2.64}$$

$$\Box h_{ij}^{\mathrm{TT}} = -16\pi\sigma_{ij}. \tag{2.65}$$

Notice that **only the metric components h_{ij}^{TT} obey a wave-like equation**. The other variables Θ, Φ and Ξ_i are determined by Poisson-type equations. Indeed, in a purely vacuum spacetime, the field equations reduce to five Laplace equations and a wave equation:

$$\nabla^2\Theta^{\mathrm{vac}} = 0, \tag{2.66}$$

$$\nabla^2\Phi^{\mathrm{vac}} = 0, \tag{2.67}$$

$$\nabla^2\Xi_i^{\mathrm{vac}} = 0, \tag{2.68}$$

$$\Box h_{ij}^{\mathrm{TT,vac}} = 0. \tag{2.69}$$

This manifestly demonstrates that only the h_{ij}^{TT} metric components—the transverse, traceless degrees of freedom of the metric perturbation—characterize the radiative degrees of freedom in the spacetime. Although it is possible to pick a gauge in which other metric components *appear* to be radiative, they will not be: their radiative character is an illusion arising due to the choice of gauge or coordinates.

The field equations (2.62)–(2.65) also demonstrate that, far from a dynamic, radiating source, the time-varying portion of the physical degrees of freedom in the metric is dominated by h_{ij}^{TT}. If we expand the gauge-invariant fields Φ, Θ, Ξ_i and h_{ij}^{TT} in powers of $1/r$, then, at sufficiently large distances, the leading-order $O(1/r)$ terms will dominate. For the fields Θ, Φ and Ξ_i, the coefficients of the $1/r$ pieces are simply the conserved mass $\int \mathrm{d}^3x\,\rho$ or the conserved linear momentum $-\int \mathrm{d}^3x\,S_i$, from the conservation relations (2.56)–(2.58). Thus, the only time-varying

piece of the physical degrees of freedom in the metric perturbation at order $O(1/r)$ is the TT piece $h_{ij}^{\rm TT}$. An alternative proof of this result is given in exercise 19.1 of Misner *et al* [4].

Although the variables Φ, Θ, Ξ_i and $h_{ij}^{\rm TT}$ have the advantage of being gauge-invariant, they have the disadvantage of being non-local. Computation of these variables at a point requires knowledge of the metric perturbation h_{ab} everywhere. This non-locality obscures the fact that the physical, non-radiative degrees of freedom are causal, a fact which is explicit in Lorentz gauge [4]. On the other hand, many observations that seek to detect GWs are sensitive only to the value of the Riemann tensor at a given point in space (see section 3). For example, the Riemann tensor components R_{itjt}, which are directly observable by detectors such as LIGO, are given in terms of the gauge-invariant variables as

$$R_{itjt} = -\tfrac{1}{2}\ddot{h}_{ij}^{\rm TT} + \Phi_{,ij} + \dot{\Xi}_{(i,j)} - \tfrac{1}{2}\ddot{\Theta}\delta_{ij}. \tag{2.70}$$

Thus, at least certain combinations of the gauge-invariant variables are locally observable.

2.3. Local regions of spacetime

In the previous subsection we described a splitting of metric perturbations into radiative, non-radiative and gauge pieces. This splitting requires that the linearized Einstein equations be valid throughout the spacetime. However, this assumption is not valid in the real Universe: many sources of GWs are intrinsically strong field sources and cannot be described using linearized theory, and on cosmological scales the metric of our Universe is not close to the Minkowski metric. Furthermore, the splitting requires a knowledge of the metric throughout all of spacetime, whereas any measurements or observations can probe only finite regions of spacetime. For these reasons it is useful to consider linearized perturbation theory in finite regions of spacetime, and to try to define gravitational radiation in this more general context.

Consider therefore a finite volume $\mathcal{V}$ in space. Can we split up the metric perturbation h_{ab} in $\mathcal{V}$ into radiative and non-radiative pieces? In general, the answer is no: within any finite region, GWs cannot be distinguished from time-varying near-zone fields generated by sources outside that region. One way to see this is to note that in finite regions of space, the decomposition of the metric into various pieces becomes non-unique, as does the decomposition of vectors into transverse and longitudinal pieces. (For example the vector $(x^2 - y^2)\partial_z$ is both transverse and longitudinal.) Alternatively, we note that within any finite vacuum region $\mathcal{V}$, one can *always* find a gauge which is locally TT, that is, a gauge which satisfies the conditions (2.18)–(2.20) within the region. (This fact does not seem to be widely known, so we give a proof in appendix A.) In particular, this applies to the static Coulomb-type field of a point source, as long as the source itself is outside of $\mathcal{V}$. Consequently, isolating the TT piece of the metric perturbation does not yield just the radiative degrees of freedom within a local region—a TT metric perturbation may also contain, for example, Coulomb-type fields.

Within finite regions of space, therefore, GWs cannot be defined at a fundamental level—one simply has time-varying gravitational fields. However, there is a certain limit in which GWs can be approximately defined in local regions, namely the limit in which the wavelength of the waves is much smaller than length and timescales characterizing the background metric.

[4] One way to see that the guage-invariant degrees of freedom are causal is to combine the vacuum wave equation (2.16) for the metric perturbation with the expression (2.33) for the gauge-invariant Riemann tensor. This gives the wave equation $\Box R_{\alpha\beta\gamma\delta} = 0$.

This definition of gravitational radiation is discussed in detail and in a more general context in section 5. As discussed in that section, this limit will always be valid when one is sufficiently far from all radiating sources.

3. Interaction of gravitational waves with a detector

The usual notion of 'gravitational force' disappears in general relativity, replaced instead by the idea that freely falling bodies follow geodesics in spacetime. Given a spacetime metric g_{ab} and a set of spacetime coordinates x^a, geodesic trajectories are given by the equation

$$\frac{\mathrm{d}^2x^a}{\mathrm{d}\tau^2} + \Gamma^a{}_{bc}\frac{\mathrm{d}x^b}{\mathrm{d}\tau}\frac{\mathrm{d}x^c}{\mathrm{d}\tau} = 0, \tag{3.1}$$

where τ is a proper time as measured by an observer travelling along the geodesic. By writing the derivatives in the geodesic equation (3.1) in terms of coordinate time t rather than proper time τ, and by combining the $a = t$ equation with the spatial, $a = j$ equations, we obtain an equation for the coordinate acceleration:

$$\frac{\mathrm{d}^2x^i}{\mathrm{d}t^2} = -(\Gamma^i_{tt} + 2\Gamma^i_{tj}v^j + \Gamma^i_{jk}v^jv^k) + v^i(\Gamma^t_{tt} + 2\Gamma^t_{tj}v^j + \Gamma^t_{jk}v^jv^k), \tag{3.2}$$

where $v^i = \mathrm{d}x^i/\mathrm{d}t$ is the coordinate velocity.

Let us now specialize to linearized theory, with the non-flat part of our metric dominated by a GW in TT gauge. Further, let us specialize to non-relativistic motion for our test body. This implies that $v^i \ll 1$, and to a good approximation we can neglect the velocity-dependent terms in equation (3.2):

$$\frac{\mathrm{d}^2x^i}{\mathrm{d}t^2} + \Gamma^i{}_{tt} = 0. \tag{3.3}$$

In linearized theory and TT gauge,

$$\Gamma^i{}_{tt} = \Gamma_{itt} = \tfrac{1}{2}\left(2\partial_t h^{\mathrm{TT}}_{jt} - \partial_j h^{\mathrm{TT}}_{tt}\right) = 0 \tag{3.4}$$

since $h^{\mathrm{TT}}_{at} = 0$. Hence, we find that $\mathrm{d}^2x^i/\mathrm{d}t^2 = 0$.

Does this result mean that the GW has no effect? Certainly not! It just tells us that, in TT gauge the *coordinate location* of a slowly moving, freely falling body is unaffected by the GW. In essence, the coordinates move with the waves.

This result illustrates why, in general relativity, it is important to focus upon coordinate-invariant observables—a naive interpretation of the above result would be that freely falling bodies are not influenced by GWs. In fact, the GWs cause the *proper separation* between two freely falling particles to oscillate, even if the *coordinate separation* is constant. Consider two spatial freely falling particles, located at $z = 0$, and separated on the x-axis by a coordinate distance L_c. Consider a GW in TT gauge that propagates down the z-axis, $h^{\mathrm{TT}}_{ab}(t, z)$. The proper distance L between the two particles in the presence of the GW is given by

$$\begin{aligned} L &= \int_0^{L_c} \mathrm{d}x\sqrt{g_{xx}} = \int_0^{L_c} \mathrm{d}x\sqrt{1 + h^{\mathrm{TT}}_{xx}(t, z = 0)} \\ &\simeq \int_0^{L_c} \mathrm{d}x[1 + \tfrac{1}{2}h^{\mathrm{TT}}_{xx}(t, z = 0)] = L_c[1 + \tfrac{1}{2}h^{\mathrm{TT}}_{xx}(t, z = 0)]. \end{aligned} \tag{3.5}$$

Notice that we use the fact that the coordinate location of each particle is fixed in TT gauge! In a gauge in which the particles move with respect to the coordinates, the limits of integration would have to vary. Equation (3.5) tells us that the proper separation of the two particles oscillates with a fractional length change $\delta L/L$ given by

$$\frac{\delta L}{L} \simeq \tfrac{1}{2} h^{\mathrm{TT}}_{xx}(t, z = 0). \tag{3.6}$$

Although we used TT gauge to perform this calculation, the result is gauge-independent; we will derive it in a different gauge momentarily. Notice that h^{TT}_{xx} acts as a strain—a fractional length change. The magnitude h of a wave is often referred to as the 'wave strain'. The proper distance we have calculated here is a particularly important quantity since it directly relates to the accumulated phase which is measured by laser interferometric GW observatories (cf the contribution by Danzmann in this volume). The 'extra' phase $\delta\phi$ accumulated by a photon that travels down and back the arm of a laser interferometer in the presence of a GW is $\delta\phi = 4\pi\delta L/\lambda$, where λ is the photon's wavelength and δL is the distance the end mirror moves relative to the beam splitter[5]. We now give a different derivation of the fractional length change (3.6) based on the concept of *geodesic deviation*. Consider a geodesic in spacetime given by $x^a = z^a(\tau)$, where τ is the proper time, with four velocity $u^a(\tau) = \mathrm{d}z^a/\mathrm{d}\tau$. Suppose we have a nearby geodesic $x^a(\tau) = z^a(\tau) + L^a(\tau)$, where $L^a(\tau)$ is small. We can regard the coordinate displacement L^a as a vector $\vec{L} = L^a\partial_a$ on the first geodesic; this is valid to first order in $\vec{L}$. Without loss of generality, we can make the connecting vector be purely spatial: $L^a u_a = 0$. Spacetime curvature causes the separation vector to change with time—the geodesics will move further apart or closer together, with an acceleration given by the geodesic deviation equation

$$u^b\nabla_b(u^c\nabla_c L^a) = -R^a{}_{bcd}[\vec{z}(\tau)]u^b L^c u^d; \tag{3.7}$$

see, e.g., [36], chapter 21. This equation is valid to linear order in L^a; fractional corrections to this equation will scale as $L/\mathcal{L}$, where $\mathcal{L}$ is the lengthscale over which the curvature varies.

For application to GW detectors, the shortest such lengthscale $\mathcal{L}$ is the wavelength λ of the GWs. Thus, the geodesic deviation equation will have fractional corrections of order L/λ. For ground-based detectors L is $\lesssim$ a few km, while $\lambda \gtrsim 3000$ km (see section 6.1); thus the approximation will be valid. For detectors with $L \gtrsim \lambda$ (e.g. the space-based detector LISA), the analysis here is not valid and other techniques must be used to analyse the detector.

A convenient coordinate system for analysing the geodesic deviation equation (3.7) is the *local proper reference frame* of the observer who travels along the first geodesic. This coordinate system is defined by the requirements

$$z^i(\tau) = 0, \qquad g_{ab}(t, \mathbf{0}) = \eta_{ab}, \qquad \Gamma^a_{bc}(t, \mathbf{0}) = 0, \tag{3.8}$$

which imply that the metric has the form

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}\mathbf{x}^2 + O\left(\frac{\mathbf{x}^2}{\mathcal{R}^2}\right). \tag{3.9}$$

[5] This description of the phase shift holds only if $L \ll \lambda$, so that the metric perturbation does not change value very much during a light travel time. This condition will be violated in the high-frequency regime for space-based GW detectors; a more careful analysis of the phase shift is needed in this case [43].

Here $\mathcal{R}$ is the radius of curvature of spacetime, given by $\mathcal{R}^{-2} \sim \|R_{abcd}\|$. It also follows from the gauge conditions (3.8) that proper time τ and coordinate time t coincide along the first geodesic, that $\vec{u} = \partial_t$ and that $L^a = (0, L^i)$.

Consider now the proper distance between the two geodesics, which are located at $x^i = 0$ and $x^i = L^i(t)$. From the metric (3.9), we see that this proper distance is just $|\mathbf{L}| = \sqrt{L_i L_i}$, up to fractional corrections of order $L^2/\mathcal{R}^2$. For a GW of amplitude h and wavelength λ, we have $\mathcal{R}^{-2} \sim h/\lambda^2$, so the fractional errors are $\sim hL^2/\lambda^2$. (Notice that $\mathcal{R} \sim \mathcal{L}/\sqrt{h}$—the wave's curvature scale $\mathcal{R}$ is much larger than the lengthscale $\mathcal{L}$ characterizing its variations.) Since we are restricting attention to detectors with $L \ll \lambda$, these fractional errors are much smaller than the fractional distance changes $\sim h$ caused by the GW (equation (3.6)). Therefore, we can simply identify $|\mathbf{L}|$ as the proper separation.

We now evaluate the geodesic deviation equation (3.7) in the local proper reference frame coordinates. From the conditions (3.8), it follows that we can replace the covariant time-derivative operator $u^a \nabla_a$ with $\partial/(\partial t)$. Using $\vec{u} = \partial_t$ and $L^a = (0, L^i)$, we get

$$\frac{\mathrm{d}^2 L^i(t)}{\mathrm{d}t^2} = -R_{itjt}(t, \mathbf{0}) L^j(t). \tag{3.10}$$

Note that the key quantity entering into the equation, R_{itjt}, is gauge-invariant in linearized theory, so we can use any convenient coordinate system to evaluate it. Using the expression (2.21) for the Riemann tensor in terms of the TT gauge metric perturbation h_{ij}^{TT}, we find that

$$\frac{\mathrm{d}^2 L^i}{\mathrm{d}t^2} = \frac{1}{2} \frac{\mathrm{d}^2 h_{ij}^{\mathrm{TT}}}{\mathrm{d}t^2} L^j. \tag{3.11}$$

Integrating this equation using $L^i(t) = L_0^i + \delta L^i(t)$ with $|\delta \mathbf{L}| \ll |\mathbf{L}_0|$ gives

$$\delta L^i(t) = \tfrac{1}{2} h_{ij}^{\mathrm{TT}}(t) L_0^j. \tag{3.12}$$

This equation is ideal for analysing an interferometric GW detector. We choose Cartesian coordinates such that the interferometer's two arms lie along the x- and y-axes, with the beam splitter at the origin. For concreteness, let us imagine that the GW propagates along the z-axis. Then, as discussed in section 2.1, the only non-zero components of the metric perturbation are $h_{xx}^{\mathrm{TT}} = -h_{yy}^{\mathrm{TT}} = h_+$ and $h_{xy}^{\mathrm{TT}} = h_{yx}^{\mathrm{TT}} = h_\times$, where $h_+(t - z)$ and $h_\times(t - z)$ are the two polarization components. We take the ends of one of the interferometer's two arms as defining the two nearby geodesics; the first geodesic is defined by the beam splitter at $\mathbf{x} = 0$, the second by the end-mirror. From equation (3.12) we then find that the distances $L = |\mathbf{L}|$ of the arms' ends from the beam splitter vary with time as

$$\frac{\delta L_x}{L} = \tfrac{1}{2} h_+, \qquad \frac{\delta L_y}{L} = -\tfrac{1}{2} h_+. \tag{3.13}$$

(Here the subscripts x and y denote the two different arms, not the components of a vector.) These distance changes are then measured via laser interferometry. Notice that the GW (which is typically a sinusoidally varying function) acts tidally, squeezing along one axis and stretching along the other. In this configuration, the detector is sensitive only to the + polarization of the GW. The $\times$ polarization acts similarly, except that it squeezes and stretches along a set of axes

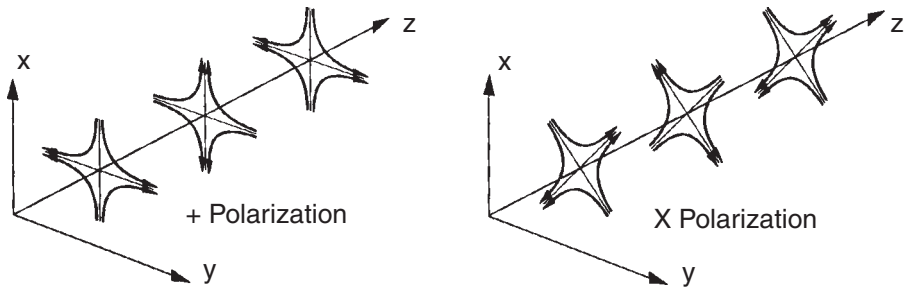

Figure 1. Lines of force for a purely + GW (left), and for a purely × GW (right). Figure kindly provided by Kip Thorne; originally published in [44].

that are rotated with respect to the x and y axes by 45°. The force lines corresponding to the two different polarizations are illustrated in figure 1.

Of course, we do not expect nature to provide GWs that so perfectly align with our detectors. In general, we will need to account for the detector's antenna pattern, meaning that we will be sensitive to some weighted combination of the two polarizations, with the weights depending upon the location of a source on the sky, and the relative orientation of the source and the detector. See [45], equations (104a,b) and associated text for further discussion.

Finally, in our analysis so far of detection, we have assumed that the only contribution to the metric perturbation is the GW contribution. However, in reality time-varying near-zone gravitational fields produced by sources in the vicinity of the detector will also be present. From equation (3.10) we see that the quantity that is actually measured by interferometric detectors is the spacetime–spacetime or electric-type piece R_{itjt} of the Riemann tensor (or more precisely the time-varying piece of this within the frequency band of the detector). From the general expression (2.70) for this quantity, we see that R_{itjt} contains contributions from both $h_{ij}^{\rm TT}$ describing GWs, and also additional terms describing the time-varying near-zone gravitational fields. There is no way for the detector to separate these two contributions, and the time-varying near-zone gravitational fields produced by motions of bedrock, air, human bodies, and tumbleweeds can all contribute to the output of the detector and act as sources of noise [46]–[48].

4. The generation of gravitational waves: putting in the source

4.1. Slow-motion sources in linearized gravity

Gravitational waves are generated by the matter source term on the right-hand side of the linearized Einstein equation

$$\Box\bar{h}_{ab} = -16\pi T_{ab}, \tag{4.1}$$

cf equation (2.15) (presented here in Lorentz gauge). In this section we will compute the leading-order contribution to the spatial components of the metric perturbation for a source whose internal motions are slow compared to the speed of light ('slow-motion sources'). We will then compute the TT piece of the metric perturbation to obtain the standard quadrupole formula for the emitted radiation.

Equation (4.1) can be solved by using a Green's function. A wave equation with source generically takes the form

$$\Box f(t, \mathbf{x}) = s(t, \mathbf{x}), \tag{4.2}$$

where $f(t, \mathbf{x})$ is the radiative field, depending on time t and position $\mathbf{x}$, and $s(t, \mathbf{x})$ is a source function. The Green's function $G(t, \mathbf{x}; t', \mathbf{x}')$ is the field which arises due to a delta function source; it tells how much field is generated at the 'field point' $(t, \mathbf{x})$ per unit source at the 'source point' $(t', \mathbf{x}')$:

$$\Box G(t, \mathbf{x}; t', \mathbf{x}') = \delta(t - t')\delta(\mathbf{x} - \mathbf{x}'). \tag{4.3}$$

The field which arises from our actual source is then given by integrating the Green's function against $s(t, \mathbf{x})$:

$$f(t, \mathbf{x}) = \int \mathrm{d}t' \, \mathrm{d}^3x' \, G(t, \mathbf{x}; t', \mathbf{x}')s(t', \mathbf{x}'). \tag{4.4}$$

The Green's function associated with the wave operator $\Box$ is very well known (see, e.g. [49]):

$$G(t, \mathbf{x}; t', \mathbf{x}') = -\frac{\delta(t' - [t - |\mathbf{x} - \mathbf{x}'|/c])}{4\pi|\mathbf{x} - \mathbf{x}'|}. \tag{4.5}$$

The quantity $t - |\mathbf{x} - \mathbf{x}'|/c$ is the *retarded time*; it takes into account the lag associated with the propagation of information from events at $\mathbf{x}$ to position $\mathbf{x}'$. The speed of light c has been restored here to emphasize the causal nature of this Green's function; we set it back to unity in what follows.

Applying this result to equation (4.1), we find

$$\bar{h}_{ab}(t, \mathbf{x}) = 4 \int \mathrm{d}^3x' \, \frac{T_{ab}(t - |\mathbf{x} - \mathbf{x}'|, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}. \tag{4.6}$$

As already mentioned, the radiative degrees of freedom are contained entirely in the spatial part of the metric, projected transverse and traceless. Firstly, consider the spatial part of the metric:

$$\bar{h}_{ij}(t, \mathbf{x}) = 4 \int \mathrm{d}^3x' \, \frac{T^{ij}(t - |\mathbf{x} - \mathbf{x}'|, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}. \tag{4.7}$$

We have raised indices on the right-hand side, using the rule that the position of spatial indices in linearized theory is irrelevant.

We now evaluate this quantity at large distances from the source. This allows us to replace the factor $|\mathbf{x} - \mathbf{x}'|$ in the denominator with $r = |\mathbf{x}|$. The corresponding fractional errors scale as $\sim L/r$, where L is the size of the source; these errors can be neglected. We also make the same replacement in the time argument of T_{ij}:

$$T_{ij}(t - |\mathbf{x} - \mathbf{x}'|, \mathbf{x}') \approx T_{ij}(t - r, \mathbf{x}'). \tag{4.8}$$

Using the formula $|\mathbf{x}-\mathbf{x}'| = r - n^i x'^i + O(1/r)$, where $n^i = x^i/r$, we see that the fractional errors generated by the replacement (4.8) scale as L/τ, where τ is the timescale over which the system is changing. This quantity is just the velocity of internal motions of the source (in units with $c = 1$), and is therefore small compared to one by our assumption. These replacements give

$$\bar{h}_{ij}(t,\mathbf{x}) = \frac{4}{r}\int \mathrm{d}^3x'\, T^{ij}(t-r,\mathbf{x}'), \tag{4.9}$$

which is the first term in a multipolar expansion of the radiation field.

Equation (4.9) almost gives us the quadrupole formula that describes GW emission (at leading order). To get the rest of the way there, we need to massage this equation a bit. The stress–energy tensor must be conserved, which means $\partial_a T^{ab} = 0$ in linearized theory. Breaking this up into time and space components, we have

$$\partial_t T^{tt} + \partial_i T^{ti} = 0, \tag{4.10}$$

$$\partial_t T^{ti} + \partial_j T^{ij} = 0. \tag{4.11}$$

From this, it follows rather simply that

$$\partial_t^2 T^{tt} = \partial_k \partial_l T^{kl}. \tag{4.12}$$

Multiply both sides of this equation by $x^i x^j$. We first manipulate the left-hand side:

$$\partial_t^2 T^{tt} x^i x^j = \partial_t^2 (T^{tt} x^i x^j). \tag{4.13}$$

Next, manipulate the right-hand side of equation (4.12), multiplied by $x^i x^j$:

$$\partial_k \partial_l T^{kl} x^i x^j = \partial_k \partial_l (T^{kl} x^i x^j) - 2\partial_k \left(T^{ik} x^j + T^{kj} x^i\right) + 2T^{ij}. \tag{4.14}$$

This identity is easily verified[6] by expanding the derivatives and applying the identity $\partial_i x^j = \delta_i{}^j$. We thus have

$$\partial_t^2 (T^{tt} x^i x^j) = \partial_k \partial_l (T^{kl} x^i x^j) - 2\partial_k (T^{ik} x^j + T^{kj} x^i) + 2T^{ij}. \tag{4.15}$$

This yields

$$\begin{aligned}
\frac{4}{r}\int \mathrm{d}^3x'\, T_{ij} &= \frac{4}{r}\int \mathrm{d}^3x' \left[\tfrac{1}{2}\partial_t^2 (T^{tt} x'^i x'^j) + \partial_k (T^{ik} x'^j + T^{kj} x'^i) - \tfrac{1}{2}\partial_k \partial_l (T^{kl} x'^i x'^j)\right] \\
&= \frac{2}{r}\int \mathrm{d}^3x'\, \partial_t^2 (T^{tt} x'^i x'^j) \\
&= \frac{2}{r}\frac{\partial^2}{\partial t^2}\int \mathrm{d}^3x'\, T^{tt} x'^i x'^j \\
&= \frac{2}{r}\frac{\partial^2}{\partial t^2}\int \mathrm{d}^3x'\, \rho\, x'^i x'^j.
\end{aligned} \tag{4.16}$$

[6] Although one of us (SAH) was unable to do this simple calculation while delivering lectures at a summer school in Brownsville, TX. Never attempt to derive the quadrupole formula while medicated.

In going from the first line to the second, we used the fact that the second and third terms under the integral are divergences. Using Gauss's theorem, they can thus be recast as surface integrals; taking the surface outside the source, their contribution is zero. In going from the second line to the third, we used the fact that the integration domain is not time-dependent, so we can take the derivatives out of the integral. Finally, we used the fact that T^{tt} is the mass density ρ. Defining the second moment I_{ij} of the mass distribution via

$$I_{ij}(t) = \int \mathrm{d}^3x'\, \rho(t, \mathbf{x}')x'^i x'^j, \tag{4.17}$$

and combining equations (4.9) and (4.16) now gives

$$\bar{h}_{ij}(t, \mathbf{x}) = \frac{2}{r}\frac{\mathrm{d}^2 I_{ij}\,(t-r)}{\mathrm{d}t^2}. \tag{4.18}$$

When we subtract the trace from I_{ij}, we obtain the quadrupole moment tensor:

$$\mathcal{I}_{ij} = I_{ij} - \tfrac{1}{3}\delta_{ij} I, \qquad I = I_{ii}. \tag{4.19}$$

This tensor will prove handy shortly.

To complete the derivation, we must project out the non-TT pieces of the right-hand side of equation (4.18). Since we are working to leading order in $1/r$, at each field point $\mathbf{x}$ this operation reduces to algebraically projecting the tensor perpendicularly to the local direction of propagation $\mathbf{n} = \mathbf{x}/r$, and subtracting off the trace. It is useful to introduce the projection tensor,

$$P_{ij} = \delta_{ij} - n_i n_j. \tag{4.20}$$

This tensor eliminates vector components parallel to $\mathbf{n}$, leaving only transverse components. Thus,

$$\bar{h}^T_{ij} = \bar{h}_{kl} P_{ik} P_{jl} \tag{4.21}$$

is a transverse tensor. Finally, we remove the trace; what remains is

$$h^{\mathrm{TT}}_{ij} = \bar{h}_{kl} P_{ik} P_{jl} - \tfrac{1}{2} P_{ij} P_{kl} \bar{h}_{kl}. \tag{4.22}$$

Substituting equation (4.18) into (4.22), we obtain our final quadrupole formula:

$$h^{\mathrm{TT}}_{ij}(t, \mathbf{x}) = \frac{2}{r}\frac{\mathrm{d}^2 \mathcal{I}_{kl}\,(t-r)}{\mathrm{d}t^2} P_{ik}(\mathbf{n}) P_{jl}(\mathbf{n}). \tag{4.23}$$

4.2. Extension to sources with non-negligible self-gravity

Our derivation of the quadrupole formula (4.23) assumed the validity of the linearized Einstein equations. In particular, the derivation is not applicable to systems with weak (Newtonian) gravity whose dynamics are dominated by self-gravity, such as binary star systems[7]. This shortcoming

[7] Stress–energy conservation in linearized gravity, $\partial^a T_{ab} = 0$, forces all bodies to move on geodesics of the Minkowski metric.

of the above linearized-gravity derivation of the quadrupole formula was first pointed out by Eddington. However, it is very straightforward to extend the derivation to encompass systems with non-negligible self-gravity.

In full general relativity, we define the quantity $\bar{h}^{ab}$ via

$$\sqrt{-g}g^{ab} = \eta^{ab} - \bar{h}^{ab}, \tag{4.24}$$

where $\eta^{ab} \equiv \mathrm{diag}(-1, 1, 1, 1)$. When gravity is weak this definition coincides with our previous definition of $\bar{h}^{ab}$ as a trace-reversed metric perturbation. We impose the harmonic gauge condition

$$\partial_a(\sqrt{-g}g^{ab}) = \partial_a\bar{h}^{ab} = 0. \tag{4.25}$$

In this gauge, the Einstein equation can be written as

$$\Box_{\mathrm{flat}}\bar{h}^{ab} = -16\pi(T^{ab} + t^{ab}), \tag{4.26}$$

where $\Box_{\mathrm{flat}} \equiv \eta^{ab}\partial_a\partial_b$ is the flat-spacetime wave operator and t^{ab} is a pseudo-tensor that is constructed from $\bar{h}^{ab}$. Taking a coordinate divergence of this equation and using the gauge condition (4.25), shows that stress–energy conservation can be written as

$$\partial_a(T^{ab} + t^{ab}) = 0. \tag{4.27}$$

Equations (4.25)–(4.27) are precisely the same equations as are used in the linearized-gravity derivation of the quadrupole formula, except for the fact that the stress–energy tensor T^{ab} is replaced by $T^{ab} + t^{ab}$. Therefore, the derivation of the last subsection carries over, with the modification that the formula (4.17) for I_{ij} is replaced by

$$I_{ij}(t) = \int \mathrm{d}^3x' \, [T^{tt}(t, \mathbf{x}') + t^{tt}(t, \mathbf{x}')]x'^i x'^j. \tag{4.28}$$

In this equation the term t^{tt} describes gravitational-binding energy, roughly speaking. For systems with weak gravity, this term is negligible in comparison with the term T^{tt} describing the rest-masses of the bodies. Therefore, the quadrupole formula (4.23) and the original definition (4.17) of I_{ij} continue to apply to the more general situation considered here.

4.3. Dimensional analysis

The rough form of the leading GW field that we just derived, equation (4.23), can be deduced using simple physical arguments. First, we define some moments of the mass distribution. The zeroth moment is just the mass itself:

$$M_0 \equiv \int \rho \, \mathrm{d}^3x = M. \tag{4.29}$$

(More accurately, this is the total mass-energy of the source.) Next, we define the dipole moment:

$$M_1 \equiv \int \rho \, x_i \, \mathrm{d}^3x = ML_i. \tag{4.30}$$

L_i is a vector with the dimension of length; it describes the displacement of the centre of mass from our chosen origin. (As such, M_1 is clearly not a very meaningful quantity—we can change its value simply by choosing a different origin.)

If our mass distribution exhibits internal motion, then moments of the *mass current*, $j_i = \rho v_i$, are also important. The first moment is the spin angular momentum:

$$S_1 \equiv \int \rho v_j \, x_k \, \epsilon_{ijk} \, \mathrm{d}^3x = S_i. \tag{4.31}$$

Finally, we look at the second moment of the mass distribution:

$$M_2 \equiv \int \rho x_i x_j \, \mathrm{d}^3x = ML_{ij}, \tag{4.32}$$

where L_{ij} is a tensor with the dimension length squared.

Using dimensional analysis and simple physical arguments, it is simple to see that the first moment that can contribute to GW emission is M_2. Consider first M_0. We want to combine M_0 with the distance to our source, r, in such a way as to produce a dimensionless wavestrain h. The only way to do this (bearing in mind that the strain should fall off as $1/r$, and restoring factors of G and c) is to put

$$h \sim \frac{G}{c^2}\frac{M_0}{r}. \tag{4.33}$$

Does this formula make sense for radiation? Not at all! Conservation of mass-energy tells us that M_0 for an isolated source cannot vary dynamically. This h cannot be radiative; it corresponds to a Newtonian potential, rather than a GW.

How about the moment M_1? In order to get the dimensions right, we must take one time derivative:

$$h \sim \frac{G}{c^3}\frac{\mathrm{d}}{\mathrm{d}t}\frac{M_1}{r}. \tag{4.34}$$

(The extra factor of c converts the dimension of the time derivative to space, so that the whole expression is dimensionless.) Think carefully about the derivative of M_1:

$$\frac{\mathrm{d}M_1}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\int \rho x_i \, \mathrm{d}^3x = \int \rho v_i \, \mathrm{d}^3x = P_i. \tag{4.35}$$

This is the total momentum of our source. Our guess for the form of a wave corresponding to M_1 becomes

$$h \sim \frac{G}{c^3}\frac{P}{r}. \tag{4.36}$$

Can this describe a GW? Again, not a chance: the momentum of an isolated source must be conserved. By boosting into a different Lorentz frame, we can always set $P = 0$. Terms like this can only be gauge artifacts; they do not correspond to radiation. (Indeed, terms like (4.36) appear in the metric of a moving black hole and correspond to the relative velocity of the hole and the observer (see [50], chapter 5).)

How about S_1? Dimensional analysis tells us that radiation from S_1 must take the form

$$h \sim \frac{G}{c^4}\frac{\mathrm{d}}{\mathrm{d}t}\frac{S_1}{r}. \tag{4.37}$$

Conservation of angular momentum tells us that the total spin of an isolated system cannot change, so we reject this term for the same reason that we rejected (4.33)—it cannot correspond to radiation.

Finally, we examine M_2:

$$h \sim \frac{G}{c^4}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{M_2}{r}. \tag{4.38}$$

There is *no* conservation principle that allows us to reject this term. Comparing to equation (4.23), we see that this is the quadrupole formula we derived earlier, up to numerical factors.

In 'normal' units, the prefactor of this formula turns out to be G/c^4—a small number divided by a *very* big number. In order to generate interesting amounts of GWs, the quadrupole moment's variation must be enormous. The only interesting sources of GWs will be those which have very large masses undergoing extremely rapid variation; even in this case, the strain we expect from typical sources is tiny. The smallness of GWs reflects the fact that gravity is the weakest of the fundamental interactions.

4.4. Numerical estimates

Consider a binary star system, with stars of mass m_1 and m_2 in a circular orbit with separation R. The quadrupole moment is given by

$$\mathcal{I}_{ij} = \mu(x_i x_j - \tfrac{1}{3}R^2\delta_{ij}), \tag{4.39}$$

where $\mu = m_1 m_2/(m_1 + m_2)$ is the binary's reduced mass and $\mathbf{x}$ is the relative displacement, with $|\mathbf{x}| = R$. We use the centre-of-mass reference frame and choose the coordinate axes so that the binary lies in the xy plane, so $x = x_1 = R\cos\Omega t$, $y = x_2 = R\sin\Omega t$ and $z = x_3 = 0$. Let us further choose to evaluate the field on the z-axis, so that $\mathbf{n}$ points in the z-direction. The projection operators in equation (4.23) then simply serve to remove the zj components of the tensor. Bearing this in mind, the quadrupole formula (4.23) yields

$$h_{ij}^{\mathrm{TT}} = \frac{2\ddot{\mathcal{I}}_{ij}}{r}. \tag{4.40}$$

The quadrupole moment tensor is

$$\mathcal{I}_{ij} = \mu R^2 \begin{bmatrix} \cos^2\Omega t - \frac{1}{3} & \cos\Omega t\sin\Omega t & 0 \\ \cos\Omega t\sin\Omega t & \cos^2\Omega t - \frac{1}{3} & 0 \\ 0 & 0 & -\frac{1}{3} \end{bmatrix}; \tag{4.41}$$

its second derivative is

$$\ddot{\mathcal{I}}_{ij} = -2\Omega^2\mu R^2 \begin{bmatrix} \cos 2\Omega t & \sin 2\Omega t & 0 \\ -\sin 2\Omega t & -\cos 2\Omega t & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{4.42}$$

The magnitude h of a typical non-zero component of h_{ij}^{TT} is

$$h = \frac{4\mu\Omega^2 R^2}{r} = \frac{4\mu M^{2/3}\Omega^{2/3}}{r}. \tag{4.43}$$

We used Kepler's third law[8] to replace R with powers of the orbital frequency Ω and the total mass $M = m_1 + m_2$. For the purpose of our numerical estimate, we will take the members of the binary to have equal masses, so that $\mu = M/4$:

$$h = \frac{M^{5/3}\Omega^{2/3}}{r}. \tag{4.44}$$

Finally, we insert numbers corresponding to plausible sources:

$$\begin{aligned} h &\simeq 10^{-21}\left(\frac{M}{2M_\odot}\right)^{5/3}\left(\frac{1\,\mathrm{h}}{P}\right)^{2/3}\left(\frac{1\,\text{kiloparsec}}{r}\right) \\ &\simeq 10^{-22}\left(\frac{M}{2.8M_\odot}\right)^{5/3}\left(\frac{0.01\,\text{second}}{P}\right)^{2/3}\left(\frac{100\text{megaparsecs}}{r}\right). \end{aligned} \tag{4.45}$$

The first line corresponds roughly to the mass, distance and orbital period ($P = 2\pi/\Omega$) expected for the many close binary white dwarf systems in our galaxy. Such binaries are so common that they are likely to be a confusion-limited source of GWs for space-based detectors, acting in some cases as an effective source of noise. The second line contains typical parameter values for binary neutron stars that are on the verge of spiralling together and merging. Such waves are targets for the ground-based detectors that have recently begun operations. The tiny magnitude of these waves illustrates why detecting GWs is so difficult.

5. Linearized theory of gravitational waves in a curved background

At the most fundamental level, GWs can only be defined within the context of an approximation in which the wavelength of the waves is much smaller than lengthscales characterizing the background spacetime in which the waves propagate. In this section, we discuss perturbation theory of curved spacetimes, describe the approximation in which GWs can be defined, and derive the effective stress tensor which describes the energy content of GWs. The material in this section draws on the treatments given in chapter 35 of Misner *et al* [4], section 7.5 of Wald [51], and the review papers [31, 32].

5.1. Perturbation theory of curved vacuum spacetimes

Throughout this section we will for simplicity restrict attention to vacuum spacetime regions. We consider a one-parameter family of solutions of the vacuum Einstein equation, parametrized by ε, of the form

$$g_{ab} = g^{\mathrm{B}}_{ab} + \varepsilon h_{ab} + \varepsilon^2 j_{ab} + O(\varepsilon^3). \tag{5.1}$$

[8] In units with $G = 1$ and for circular orbits of radius R, $R^3\Omega^2 = M$.

Here $g^{\rm B}_{ab}$ is the background metric; it was taken to be the Minkowski metric in sections 2, 4 and 2.2. Here we allow $g^{\rm B}_{ab}$ to be any vacuum solution of the Einstein equations. The quantity h_{ab} is the linear-order metric perturbation, as in the previous sections; j_{ab} is a second-order metric perturbation which will be used in section 5.3. We can regard ε as a formal expansion parameter; we set its value to unity at the end of our calculations.

The derivation of the linearized Einstein equation proceeds as before. Most of the formulae for linearized perturbations of Minkowski spacetime continue to apply, with η_{ab} replaced by $g^{\rm B}_{ab}$, and with partial derivatives ∂_a replaced by covariant derivatives with respect to the background, $\nabla^{\rm B}_a$. Some of the formulae acquire extra terms involving coupling to the background Riemann tensor.

Inserting equation (5.1) into the formula for connection coefficients gives

$$\Gamma^a_{\ bc} = \tfrac{1}{2} g^{ad}(\partial_c g_{db} + \partial_b g_{dc} - \partial_d g_{bc}) \tag{5.2}$$

$$= \tfrac{1}{2}(g^{{\rm B}\ ad} - \varepsilon h^{ad})(\partial_c g^{\rm B}_{db} + \varepsilon\partial_c h_{db} + \partial_b g^{\rm B}_{dc} + \varepsilon\partial_b h_{dc} - \partial_d g^{\rm B}_{bc} - \varepsilon\partial_d h_{bc}) + O(\varepsilon^2)$$

$$= \Gamma^{{\rm B}\,a}_{\ bc} + \varepsilon\delta\Gamma^a_{\ bc} + O(\varepsilon^2). \tag{5.3}$$

Here $\Gamma^{{\rm B}\,a}_{\ bc}$ are the connection coefficients of the background metric $g^{\rm B}_{ab}$, and the first-order corrections to the connection coefficients are given by

$$\delta\Gamma^a_{\ bc} = -\tfrac{1}{2} h^{ad} g^{\rm B}_{de}\Gamma^{{\rm B}\,e}_{\ bc} + \tfrac{1}{2} g^{{\rm B}ad}(\partial_c h_{db} + \partial_b h_{dc} - \partial_d h_{bc})$$

$$= \tfrac{1}{2} g^{{\rm B}ad}(\nabla^{\rm B}_c h_{db} + \nabla^{\rm B}_b h_{dc} - \nabla^{\rm B}_d h_{bc}), \tag{5.4}$$

where $\nabla^{\rm B}_a$ is the covariant-derivative operator associated with the background metric. Equation (5.4) can be derived more easily, at any given point in spacetime, by evaluating the expression (5.2) in a coordinate system in which the background connection coefficients vanish at that point, so that $\partial_a = \nabla^{\rm B}_a$. The result (5.4) for general coordinate systems then follows from general covariance.

Next, insert the expansion (5.3) of the connection coefficients into the formula

$$R^a_{\ bcd} = \partial_c\Gamma^a_{\ bd} - \partial_d\Gamma^a_{\ bc} + \Gamma^a_{\ ce}\Gamma^e_{\ bd} - \Gamma^a_{\ de}\Gamma^e_{\ bc} \tag{5.5}$$

for the Riemann tensor. Evaluating the result in a coordinate system in which $\Gamma^{{\rm B}\,a}_{\ bc} = 0$ at the point of evaluation gives

$$R^a_{\ bcd} = \partial_c\Gamma^{{\rm B}\,a}_{\ bd} - \partial_d\Gamma^{{\rm B}\,a}_{\ bc} + \varepsilon(\partial_c\delta\Gamma^a_{\ bd} - \partial_d\delta\Gamma^a_{\ bc}) + O(\varepsilon^2)$$

$$= R^{{\rm B}a}_{\ \ bcd} + \varepsilon\delta R^a_{\ bcd} + O(\varepsilon^2). \tag{5.6}$$

Here $R^{{\rm B}\,a}_{\ \ bcd}$ is the Riemann tensor of the background metric and $\delta R^a_{\ bcd} = \partial_c\delta\Gamma^a_{\ bd} - \partial_d\delta\Gamma^a_{\ bc}$ is the linear perturbation to the Riemann tensor. It follows from general covariance that the expression for $\delta R^a_{\ bcd}$ in a general coordinate system is

$$\delta R^a_{\ bcd} = \nabla^{\rm B}_{\ c}\delta\Gamma^a_{\ bd} - \nabla^{\rm B}_d\delta\Gamma^a_{\ bc}. \tag{5.7}$$

Using the expression (5.4) now gives

$$\delta R^a{}_{bcd} = \tfrac{1}{2}(\nabla^{\rm B}_c\nabla^{\rm B}_b h^a{}_d + \nabla^{\rm B}_c\nabla^{\rm B}_d h^a{}_b - \nabla^{\rm B}_c\nabla^{\rm B\,a} h_{bd} - \nabla^{\rm B}_d\nabla^{\rm B}_b h^a_c - \nabla^{\rm B}_d\nabla^{\rm B}_c h^a{}_b + \nabla^{\rm B}_d\nabla^{\rm B\,a} h_{bc}). \quad (5.8)$$

Contracting on the indices a and c yields the linearized Ricci tensor δR_{bd}:

$$\delta R_{bd} = -\tfrac{1}{2}\Box^{\rm B} h_{bd} - \tfrac{1}{2}\nabla^{\rm B}_d\nabla^{\rm B}_b h + \nabla^{\rm B}_a\nabla^{\rm B}{}_{(b}h^a{}_{d)}, \quad (5.9)$$

where $\Box^{\rm B} \equiv \nabla^{\rm B}_a\nabla^{\rm B\,a}$, indices are raised and lowered with the background metric and $h = h^a{}_a$. Reversing the trace to obtain the linearized Einstein tensor δG_{bd}, and writing the result in terms of the trace-reversed metric perturbation

$$\bar{h}_{ab} = h_{ab} - \tfrac{1}{2}g^{\rm B}_{ab}g^{\rm B\,cd}h_{cd} \quad (5.10)$$

yields the linearized vacuum Einstein equation

$$0 = \delta G_{bd} = -\tfrac{1}{2}\Box^{\rm B}\bar{h}_{bd} + R^{\rm B}_{adbc}\bar{h}^{ac} - \tfrac{1}{2}g^{\rm B}_{bd}\nabla^{\rm B}_a\nabla^{\rm B}_c\bar{h}^{ac} + \tfrac{1}{2}\nabla^{\rm B}_b\nabla^{\rm B}_a\bar{h}^a{}_d + \tfrac{1}{2}\nabla^{\rm B}_d\nabla^{\rm B}_a\bar{h}^a{}_b. \quad (5.11)$$

As in section 2, the linearized Einstein equation can be simplified considerably by a suitable choice of gauge. Under a gauge transformation parametrized by the vector field ξ^a, the metric transforms as

$$h_{ab} \to h'_{ab} = h_{ab} - 2\nabla^{\rm B}{}_{(a}\xi_{b)}; \quad (5.12)$$

the divergence of the trace-reversed metric perturbation thus transforms as

$$\nabla^{\rm B\,a}\bar{h}'_{ab} = \nabla^{\rm B\,a}\bar{h}_{ab} - \Box^{\rm B}\xi_b. \quad (5.13)$$

We can enforce in the new gauge the transverse condition

$$\nabla^{\rm B\,a}\bar{h}'_{ab} = 0 \quad (5.14)$$

by requiring that ξ_b satisfies the wave equation $\Box^{\rm B}\xi_b = \nabla^{\rm B\,a}\bar{h}_{ab}$. We can further specialize the gauge to satisfy $h' = 0$. Dropping the primes, the metric perturbation is thus traceless and transverse:

$$\nabla^{\rm B\,a}h_{ab} = h = 0. \quad (5.15)$$

In this gauge, the linearized Einstein equation (5.11) simplifies to

$$0 = \delta G_{bd} = -\tfrac{1}{2}\Box^{\rm B}h_{bd} + R^{\rm B}_{adbc}h^{ac}. \quad (5.16)$$

(Note, however, that one cannot in this context impose the additional gauge conditions $h_{0a} = 0$ used in the definition of TT gauge for perturbations of flat spacetime.)

To see that the traceless condition $h = 0$ can be achieved, note that the trace transforms as

$$h \to h' = h - 2\nabla^{\rm B\,a}\xi_a. \quad (5.17)$$

Therefore, it is sufficient to find a vector field ξ^a that satisfies $\Box^{\rm B}\xi_a = 0$ and

$$\nabla^{B\ a}\xi_a - h/2 = 0. \tag{5.18}$$

We can choose initial data for ξ_a on any Cauchy hypersurface for which the quantity (5.18) and also its normal derivative vanish. Since the quantity (5.18) satisfies the homogeneous wave equation by equations (5.11) and (5.14), it will vanish everywhere.

The wave equation (5.16) differs from its flat spacetime counterpart (2.16) in two respects: firstly, there is an explicit coupling to the background Riemann tensor; and secondly, there is a coupling to the background curvature through the connection coefficients that appear in the covariant wave operator $\Box^{\rm B}$. In the limit (discussed below) where the wavelength of the waves is much smaller than the lengthscales characterizing the background metric, these couplings have the effect of causing gradual evolution in the properties of the wave. These gradual changes can be described using the formalism of geometric optics, which shows that GWs travel along null geodesics with slowly evolving amplitudes and polarizations. See [31] for a detailed description of this formalism. Outside the geometric optics limit, the curvature couplings in equation (5.16) can cause the dynamics of the metric perturbation to be strongly coupled to the dynamics of the background spacetime. An example of such coupling is the parametric amplification of metric perturbations during inflation in the early Universe [52].

5.2. General definition of gravitational waves: the geometric optics regime

The linear perturbation formalism described in the last section can be applied to any perturbation of any vacuum background spacetime. Its starting point is the separation of the spacetime metric into a background piece plus a perturbation. In most circumstances, this separation is merely a mathematical device and can be chosen arbitrarily; no unique separation is determined by local physical measurements. [Although $g^{\rm B}_{ab}$ and h_{ab} are uniquely determined once one specifies the one parameter family of metrics $g_{ab}(\varepsilon)$, a given physical situation will be described by a single metric $g_{ab}(\varepsilon_0)$ for some fixed value of ε_0 of ε, not by the one parameter family of metrics.] However, in special circumstances, a unique separation into background plus perturbation is determined by the local physical measurements, and it is only in this context that GWs can be defined. Such circumstances arise when the wavelength λ of the waves is very much smaller than the characteristic lengthscales $\mathcal{L}$, characterizing the background curvature. In this case, one can define the background metric and perturbation, to linear order, via

$$g^{\rm B}_{ab} \equiv \langle g_{ab}\rangle, \tag{5.19}$$

$$\varepsilon h_{ab} \equiv g_{ab} - g^{\rm B}_{ab}. \tag{5.20}$$

Here the angular brackets $\langle\cdots\rangle$ denote an average over lengthscales large compared to λ but small compared to $\mathcal{L}$; a suitable covariant definition of such averaging has been given by Brill and Hartle [53]. A useful analogy to consider is the surface of an orange, which contains curvatures on two different lengthscales: An overall, roughly spherical background curvature (analogous to the background metric), and a dimpled texture on small scales (analogous to the GW). The regime $\lambda \ll \mathcal{L}$ is called the *geometric optics regime*.

We will argue below that the short-wavelength perturbation εh_{ab} gives rise to an effective stress tensor of order $\varepsilon^2 h^2/\lambda^2$, where h is a typical size of h_{ab}. This effective stress tensor

contributes to the curvature of the background metric g^{B}_{ab}. This contribution to the curvature is $\lesssim 1/\mathcal{L}^2$. It follows that $\varepsilon^2 h^2/\lambda^2 \lesssim 1/\mathcal{L}^2$, or

$$\varepsilon h \lesssim \frac{\lambda}{\mathcal{L}}. \tag{5.21}$$

Since we are assuming that $\lambda \ll \mathcal{L}$, it follows that the short-wavelength piece εh_{ab} of the metric is small compared to the background metric, and so we can use the perturbation formalism of section 5.1. Consider now the splitting of the Riemann tensor into a background piece plus a perturbation given by equation (5.6):

$$R_{abcd} = R^{\mathrm{B}}_{abcd} + \varepsilon \delta R_{abcd} + O(\varepsilon^2). \tag{5.22}$$

By the definition (5.19) of the background metric, it follows that g^{B}_{ab} and R^{B}_{abcd} vary only over lengthscales $\gtrsim \mathcal{L}$, and therefore it follows that to a good approximation

$$\langle R^{\mathrm{B}}_{abcd} \rangle = R^{\mathrm{B}}_{abcd}. \tag{5.23}$$

Hence the perturbation to the Riemann tensor can be obtained via

$$\varepsilon \delta R_{abcd} = R_{abcd} - \langle R_{abcd} \rangle, \tag{5.24}$$

the same unique and local procedure as for the metric perturbation (5.20). This Riemann tensor perturbation is often called the *GW Riemann tensor*; it is a tensor characterizing the GWs that propagate in the background metric g^{B}_{ab}.

The operational meaning of the GW fields εh_{ab} and $\varepsilon \delta R_{abcd}$ follows directly from the equivalence principle and from their meaning in the context of flat spacetime (section 2). Specifically, suppose that $\mathcal{P}$ is a point in spacetime and pick a coordinate system in which $g^{\mathrm{B}}_{ab} = \eta_{ab}$ and $\Gamma^{\mathrm{B}}{}^{a}_{bc} = 0$ at $\mathcal{P}$. Then we have

$$g_{ab} = \eta_{ab} + O\left(\frac{x^2}{\mathcal{L}^2}\right) + \varepsilon h_{ab} + O(\varepsilon^2), \tag{5.25}$$

where x is the distance from $\mathcal{P}$. Therefore, within a spacetime region around $\mathcal{P}$ in which $x \ll \mathcal{L}$, the flat-spacetime perturbation theory and measurement analysis of section 2 can be applied. Thus, the gravitational waveforms seen by observers performing local experiments will just be given by components of the GW Riemann tensor in the observer's local proper reference frames.

We remark that the splitting of the metric into a background plus a linear perturbation can sometimes be uniquely defined even in the regime $\lambda \sim \mathcal{L}$. Some examples are when the background spacetime is static (e.g. perturbations of a static star), or homogeneous (e.g. Friedman–Robertson–Walker cosmological models). In these cases the dynamic metric perturbation are not actually GWs, although their evolution can be computed using the linearized Einstein equation. For example, consider the evolution of a metric perturbation mode which is parametrically amplified during inflation in the early Universe. At early times during inflation, the mode's wavelength λ is smaller than the Hubble scale ($\mathcal{L}$); the mode is said to be 'inside the horizon'. Any excitation of the mode is locally measurable (although such modes are usually assumed to start in their vacuum state). As inflation proceeds, the mode's wavelength redshifts

and becomes larger due to the rapid expansion of the Universe, and eventually becomes larger than the Hubble scale $\mathcal{L}$; the mode is then 'outside the horizon'. At this point, excitations in the mode are not locally measurable and are thus not GWs. Finally, after inflation ends, the mode 're-enters the horizon' and excitations of the mode are locally measurable. The mode is now a true GW once again.

Finally, we note that for perturbations of flat spacetime, the definition of GWs given here does not always coincide with the definition in terms of the TT component of the metric given in section 2.2. However, far from sources of GWs (the regime relevant to observations), the two definitions do coincide. This is because the TT piece of the metric will vary on scales of a wavelength λ which is short compared to the lengthscale $\sim r$ over which other pieces of the metric vary (except for other dynamic pieces of the metric such as the time-varying quadrupole term in the gauge-invariant field Φ; those pieces vary on short lengthscales but are unimportant since they are smaller than the TT piece by a factor $\sim\lambda^2/r^2$ or smaller).

5.3. Effective stress–energy tensor of gravitational waves

Two major conceptual building blocks are needed for the derivation of the energy and momentum carried by GWs [54]: the perturbation theory of section 5.1, generalized to second order in ε, and the separation of lengthscales $\lambda \ll \mathcal{L}$ discussed in the previous subsection.

We start by discussing the second-order perturbation theory. By inserting the expansion (5.1) into the vacuum Einstein equation, we obtain

$$0 = G_{ab} = G_{ab}[g^{\mathrm{B}}_{cd}] + \varepsilon G^{(1)}_{ab}[h_{cd}; g^{\mathrm{B}}_{ef}] + \varepsilon^2 G^{(1)}_{ab}[j_{cd}; g^{\mathrm{B}}_{ef}] + \varepsilon^2 G^{(2)}_{ab}[h_{cd}; g^{\mathrm{B}}_{ef}] + O(\varepsilon^3). \tag{5.26}$$

Here $G_{ab}[g^{\mathrm{B}}_{cd}]$ is the Einstein tensor of the background metric, and $G^{(1)}_{ab}[\ldots; g^{\mathrm{B}}_{ef}]$ is the linear differential operator on metric perturbations giving the linear perturbation to the Einstein tensor generated by a metric perturbation. The explicit expression for $G^{(1)}_{ab}[h_{cd}, g^{\mathrm{B}}_{ef}]$ is given by equation (5.11). The term $G^{(2)}_{ab}[h_{cd}; g^{\mathrm{B}}_{ef}]$ is the piece of the Einstein tensor that is quadratic in h_{ab}; it can be computed by extending the computation of section 5.1 to one higher order, and is a sum of terms of the form $h_{ab}\nabla^{\mathrm{B}}_c\nabla^{\mathrm{B}}_d h_{ef}$ and $(\nabla^{\mathrm{B}}_a h_{bc})(\nabla^{\mathrm{B}}_d h_{ef})$ with various index contractions; see equation (35.58b) of MTW [4]. It is worth recalling that j_{ab} is a second-order metric perturbation. We must take the calculation to second order to compute the effective stress–energy tensor of the waves, since an averaging is involved—the first-order contribution vanishes by the oscillatory nature of the waves.

Equating to zero the coefficients of the different powers of ε, we obtain the vacuum Einstein equation for the background spacetime

$$G_{ab}[g^{\mathrm{B}}_{cd}] = 0, \tag{5.27}$$

the linearized Einstein equation

$$G^{(1)}_{ab}[h_{cd}; g^{\mathrm{B}}_{ef}] = 0, \tag{5.28}$$

together with the equation for the second-order metric perturbation j_{ab}

$$G^{(1)}_{ab}[j_{cd}; g^{\mathrm{B}}_{ef}] = -G^{(2)}_{ab}[h_{cd}; g^{\mathrm{B}}_{ef}]. \tag{5.29}$$

We now specialize to the geometric optics regime $\lambda \ll \mathcal{L}$. We split the second-order metric perturbation into a piece $\langle j_{ab}\rangle$ that is slowly varying, and a piece

$$\Delta j_{ab} = j_{ab} - \langle j_{ab}\rangle \tag{5.30}$$

that is rapidly varying. The full metric can be now be written as

$$g_{ab} = (g^{\rm B}_{ab} + \varepsilon^2\langle j_{ab}\rangle) + (\varepsilon h_{ab} + \varepsilon^2 \Delta j_{ab}) + O(\varepsilon^3), \tag{5.31}$$

where the first term varies slowly on lengthscales $\sim\mathcal{L}$ and the second term varies rapidly on lengthscales $\sim\lambda$. Consider next the average of the second-order Einstein equation (5.29). Using the fact that the averaging operation $\langle\cdots\rangle$ commutes with derivatives, we get

$$G^{(1)}_{ab}[\langle j_{cd}\rangle; g^{\rm B}_{ef}] = -\langle G^{(2)}_{ab}[h_{cd}; g^{\rm B}_{ef}]\rangle. \tag{5.32}$$

Subtracting equation (5.32) from equation (5.27) gives an equation for Δj_{ab}:

$$G^{(1)}_{ab}[\Delta j_{cd}] = -G^{(2)}_{ab}[h_{cd}; g^{\rm B}_{ef}] + \langle G^{(2)}_{ab}[h_{cd}; g^{\rm B}_{ef}]\rangle. \tag{5.33}$$

Equation (5.32) can be rewritten using equation (5.27) as[9]

$$G_{ab}[g^{\rm B}_{cd} + \varepsilon^2\langle j_{cd}\rangle] = 8\pi T^{\rm GW,eff}_{ab} + O(\varepsilon^3), \tag{5.34}$$

where the effective GW stress–energy tensor is

$$T^{\rm GW,eff}_{ab} = -\frac{1}{8\pi}\langle G^{(2)}_{ab}[h_{cd}; g^{\rm B}_{ef}]\rangle. \tag{5.35}$$

In the effective Einstein equation (5.34), all the quantities vary slowly on lengthscales $\sim\mathcal{L}$. The left-hand side is the Einstein tensor of the slowly varying piece of the metric. The right-hand side is the effective stress–energy tensor, obtained by taking an average of the quadratic piece of the second-order Einstein tensor. It follows from equation (5.34) that $T^{\rm GW,eff}_{ab}$ is conserved with respect to the metric $g^{\rm B}_{ab} + \varepsilon^2\langle j_{ab}\rangle$. In particular, to leading order in ε, it is conserved with respect to the background metric $g^{\rm B}_{ab}$.

The effect of the GWs is thus to give rise to a correction $\langle j_{ab}\rangle$ to the background metric. This correction is locally of the same order as Δj_{ab}, the rapidly varying piece of the second-order metric perturbation. However, any measurements that probe only the long-lengthscale structure of the metric (e.g. measurements of the gravitating mass of a radiating source over timescales long compared to λ) are sensitive only to $\langle j_{ab}\rangle$. Thus, when one restricts attention to long lengthscales, GWs can thus be treated as any other form of matter source in general relativity. Typically $\langle j_{ab}\rangle$ grows secularly with time, while Δj_{ab} does not.

A fairly simple expression for the effective stress–energy tensor can be obtained as follows. Schematically, the effective stress–energy tensor has the form

$$T^{\rm GW,eff}_{ab} \sim \&\langle h_{ab}\nabla^{\rm B}_c\nabla^{\rm B}_d h_{ef}\rangle + \&\langle(\nabla^{\rm B}_a h_{bc})(\nabla^{\rm B}_d h_{ef})\rangle, \tag{5.36}$$

[9] Our derivation of the effective Einstein equation (5.34) requires the assumption $\varepsilon^2\langle j_{ab}\rangle \ll g^{\rm B}_{ab}$, since we use second-order perturbation theory. However, the final result is valid without this assumption [54]; the curvature generated by the GWs can be comparable to the background curvature.

where & means 'a sum of terms obtained by taking various contractions of'. In this expression, gradients scale as $1/\lambda$, so $\nabla^{\rm B}_c\nabla^{\rm B}_d \sim 1/\lambda^2$. However, the commutator of two derivatives scales as the background Riemann tensor, which is of order $1/\mathcal{L}^2$. Therefore, up to corrections of order $\lambda^2/\mathcal{L}^2$ which can be neglected, one can freely commute covariant derivatives in the expression (5.36). Also, the average of any total derivative will vanish in the limit $\lambda \ll \mathcal{L}$ if the averaging lengthscale is taken to be $\sqrt{\lambda\mathcal{L}}$. Therefore, one can flip derivatives from one factor to another inside the averages in equation (5.36), as in integration by parts. Using these manipulations, the expression for the effective stress–energy tensor simplifies to [4, 54]

$$T^{\rm GW,eff}_{ab} = \frac{1}{32\pi}\left\langle \nabla^{\rm B}_a \bar{h}_{cd} \nabla^{\rm B}_b \bar{h}^{cd} - \tfrac{1}{2}\nabla^{\rm B}_a \bar{h} \nabla^{\rm B}_b \bar{h} - \nabla^{\rm B}_a \bar{h}_{bc} \nabla^{\rm B}_d \bar{h}^{cd} - \nabla^{\rm B}_b \bar{h}_{ac} \nabla^{\rm B}_d \bar{h}^{cd} \right\rangle . \tag{5.37}$$

In gauges satisfying the transverse-traceless conditions (5.15), this reduces to

$$T^{\rm GW,eff}_{ab} = \frac{1}{32\pi}\langle \nabla^{\rm B}_a h_{cd} \nabla^{\rm B}_b h^{cd} \rangle . \tag{5.38}$$

For example, for the plane wave propagating in the z direction in flat spacetime, given by

$$\begin{aligned} h_{xx} &= -h_{yy} = h_0 \cos(\omega t - \omega z), \\ h_{ab} &= 0 \quad \text{(all other components)}, \end{aligned} \tag{5.39}$$

the energy density and energy flux are given by

$$T^{tt} = T^{tz} = \frac{h_0^2\omega^2}{16\pi}\langle \cos^2(\omega t - \omega z)\rangle = \frac{h_0^2\omega^2}{32\pi}. \tag{5.40}$$

If we restore factors of G and c, and insert numbers typical of bursts of waves that we hope to detect, we get the energy flux

$$T^{tz} = 1.5\,\text{mW m}^{-2} \left(\frac{h_0}{10^{-22}}\right)^2 \left(\frac{f}{1\,\text{kHz}}\right)^2 , \tag{5.41}$$

where $f = \omega/(2\pi)$. Note that this is a large energy flux by astronomical standards, despite the tiny value of h_0; it is comparable to the flux of reflected sunlight from a full moon [32].

6. A brief survey of gravitational wave astronomy

Having now reviewed the basic theory and properties of GWs, we conclude this paper by very briefly surveying the properties of important potential sources of GWs. Our goal is to give some indication of the value that GWs may provide for astronomical observations; much of this material is updated from a previous survey paper [29]. We note that since the focus of this paper is intended to be the theory of GW sources (and that this paper is *significantly* longer than was intended or requested), we are quite a bit more schematic in our treatment here than we have been

in the rest of this paper. This final section is intended to be a very brief, somewhat superficial survey, rather than a detailed review.

We begin by contrasting gravitational radiation with electromagnetic radiation, which forms the basis for almost all current astronomical observations:

Electromagnetic waves interact strongly with matter; GWs do not. The weak interaction of GWs is both blessing and curse: it means that they propagate from emission to Earth-bound observers with essentially zero absorption, making it possible to probe astrophysics that is hidden or dark to electromagnetic observations—e.g. the coalescence and merger of black holes, the collapse of a stellar core and the dynamics of the early Universe. It also means that detecting GWs is very difficult. Also, because many of the best sources are hidden or dark, they are very poorly understood today—we know very little about what are likely to be some of the most important sources of GWs.

Electromagnetic radiation typically has a wavelength smaller than the size of the emitting system, and so can be used to form an image of the source. This is because electromagnetic radiation is usually generated by moving charges in the environment of some larger source—e.g. an atomic transition in interstellar gas, or emission from hot plasma in a stellar environment. *By contrast, the wavelength of gravitational radiation is typically comparable to or larger than the size of the radiating source*. GWs are generated by the bulk dynamics of the source itself—e.g. the motion of neutron stars in a binary. As a consequence, GWs cannot be used to form an image: the radiation simply does not resolve the generating system. Instead, GWs are best thought of as analogous to sound—the two polarizations carry a stereophonic description of the source's dynamics.

Gravitons in a gravitational-wave burst are phase-coherent; photons in electromagnetic signals are usually phase-incoherent. This arises from the fact that each graviton is generated from the same bulk motion of matter or of spacetime curvature, while each photon is normally generated by different, independent events involving atoms, ions or electrons. Thus GWs are similar to laser light. We can take advantage of the phase coherence of GWs to enhance their detectability. Matched filtering techniques for detecting GW bursts with well-modelled functional form (like those generated by coalescing compact binaries) extend the distance to which sources can be seen by a factor of roughly the square root of the number of cycles in the waveform, a significant gain [45].

An extremely important consequence of this coherency is that *the direct observable of gravitational radiation is the strain h, a quantity that falls off with distance as* $1/r$. Most electromagnetic observables are some kind of energy flux, and so fall off with a $1/r^2$ law; measuring coherent GWs is analogous to measuring a coherent, $1/r$ electromagnetic radiation field. This comparatively slow fall off with radius means that relatively small improvements in the sensitivity of GW detectors can have a large impact on their science: doubling the sensitivity of a detector doubles the distance to which sources can be detected, increasing the volume of the Universe to which sources are measurable by a factor of 8. Every factor of 2 improvement in the sensitivity of a GW observatory should increase the number of observable sources by about an order of magnitude.

In most cases, electromagnetic astronomy is based on the deep imaging of small fields of view: observers obtain a large amount of information about sources on a small piece of the sky. GW astronomy will be a nearly all-sky affair: GW detectors have nearly 4π steradian sensitivity to events over the sky. A consequence of this is that their ability to localize a source on the sky is not good by usual astronomical standards; but, it means that any source on the sky will be detectable, not just sources towards which the detector is 'pointed'. The contrast between the all-sky sensitivity but poor angular resolution of GW observatories, and the pointed, high-angular resolution of telescopes is very similar to the angular resolution contrast of hearing and sight, strengthening the useful analogy of GWs with sound.

From these general considerations, we turn now to specifics. It is useful to categorize GW sources (and the methods for detecting their waves) by the frequency band in which they radiate. Broadly speaking, we may break the GW spectrum into four rather different bands: the ultra low-frequency band, $10^{-18} \lesssim f \lesssim 10^{-13}$ Hz; the very low-frequency band, $10^{-9} \lesssim f \lesssim 10^{-7}$ Hz; the low-frequency band, $10^{-5} \lesssim f \lesssim 1$ Hz; and the high frequency band, $1 \lesssim f \lesssim 10^{4}$ Hz.

For compact sources, the GW frequency band is typically related to the source's size R and mass M. Here the source size R means the lengthscale over which the source's dynamics vary; for example, it could be the actual size of a particular body or the separation of members of a binary. The 'natural' GW frequency of such a source is $f_{\rm GW} \sim (1/2\pi)\sqrt{GM/R^3}$. Because $R \gtrsim 2GM/c^2$ (the Schwarzschild radius of a mass M), we can estimate an upper bound for the frequency of a compact source:

$$f_{\rm GW}(M) < \frac{1}{4\sqrt{2}\pi}\frac{c^3}{GM} \simeq 10^4\,{\rm Hz}\left(\frac{M_\odot}{M}\right). \tag{6.1}$$

This is a rather hard upper limit, since many interesting sources are quite a bit larger than $2GM/c^2$, or else evolve through a range of sizes before terminating their emission at $R \sim 2GM/c^2$. Nonetheless, this frequency gives some idea about the types of compact sources that are likely to be important in each band—for example, high-frequency compact sources are of stellar mass (several solar masses); low-frequency compact sources are thousands to millions of solar masses, or else contain widely separated stellar mass bodies.

6.1. High frequency

The high-frequency band, $1 \lesssim f \lesssim 10^4$ Hz, is targeted by the new generation of ground-based laser interferometric detectors such as LIGO. (It also corresponds roughly to the audio band of the human ear: when converted to sound, LIGO sources are audible to humans.) The low-frequency end of this band is set by the fact that it is extremely difficult to prevent mechanical coupling of the detector to ground vibrations at low frequencies, and probably impossible to prevent gravitational coupling to ground vibrations, human activity and atmospheric motions [46]–[48]. The high end of the band is set by the fact that it is unlikely any interesting GW source radiates at frequencies higher than a few kilohertz. Such a source would have to be a relatively low mass ($\lesssim 1 M_\odot$) but extremely compact (cf equation (6.1)). There are no known theoretical or observational indications that gravitationally collapsed objects in this mass range exist.

The paper by Aufmuth and Danzmann in this volume [21] discusses the detectors relevant to this frequency band in some detail; our discussion here is limited to a brief

survey of these instruments. Several interferometric GW observatories are either operating or being completed in the United States, Europe, Japan and Australia. Having multiple observatories widely scattered over the globe is extremely important: the multiplicity gives rise to cross-checks that increase detection confidence and also aids in the interpretation of measurements. For example, sky location determination and concomitant measurement of the distance to a source follows from triangulation of time-of-flight differences between separated detectors.

The major interferometer projects are

- **LIGO.** The Laser Interferometer Gravitational-wave Observatory [55] consists of three operating interferometers: a single 4 km interferometer in Livingston, Louisiana, as well as a pair of interferometers (4 km and 2 km) in the LIGO facility at Hanford, Washington. The sites are separated by roughly 3000 km and are situated to support coincidence analysis of events.
- **Virgo.** Virgo is a 3 km French–Italian detector under construction near Pisa, Italy [56]. In most respects, Virgo is quite similar to LIGO. A major difference is that Virgo employs a very sophisticated seismic isolation system that promises extremely good low-frequency sensitivity.
- **GEO600.** GEO600 is a 600 m interferometer constructed by a German–English collaboration near Hanover, Germany [57]. Despite its shorter arms, GEO600 achieves sensitivity comparable to the multi-kilometre instruments using advanced interferometry techniques. This makes it an invaluable testbed for technology to be used in later generations of the larger instruments, as well as enabling it to make astrophysically interesting measurements.
- **TAMA300.** TAMA300 is a 300 m interferometer operating near Tokyo. It has been in operation for several years now [58]. The TAMA team is currently designing a 3 km interferometer [59], building on their experiences with the 300 m instrument.
- **ACIGA.** The Australian Consortium for Interferometric Gravitational-Wave Astronomy is currently constructing an 80 m research interferometer near Perth, Australia [60], hoping that it will be possible to extend it to multi-km scale in the future. Such a detector would likely be a particularly valuable addition to the worldwide stable of detectors, since all the Northern Hemisphere detectors lie very near on a common plane. An Australian detector would be far outside this plane, allowing it to play an important role in determining the location of sources on the sky.

The LIGO, GEO and TAMA instruments have now been operating for several years; see [17]–[20] for the results and upper limits from the first set of observations. All of these detectors have or will have sensitivities similar to that illustrated in figure 2 (which shows, in particular, the sensitivity goal of the first generation of LIGO interferometers). This figure also shows the 'facility limits'—the lowest noise levels that can be achieved even in principle within an interferometer facility. The low-level facility limits come from *gravity-gradient noise*: noise arising from gravitational coupling to fluctuations in the local mass distribution (such as from seismic motions in the earth near the test masses [46], human activity near the detector [47] and density fluctuations in the atmosphere [48]). At higher frequencies, the facility limit arises from the residual gas (mostly hydrogen) in the interferometer vacuum system—stray molecules of gas effectively cause stochastic fluctuations in the index of refraction.

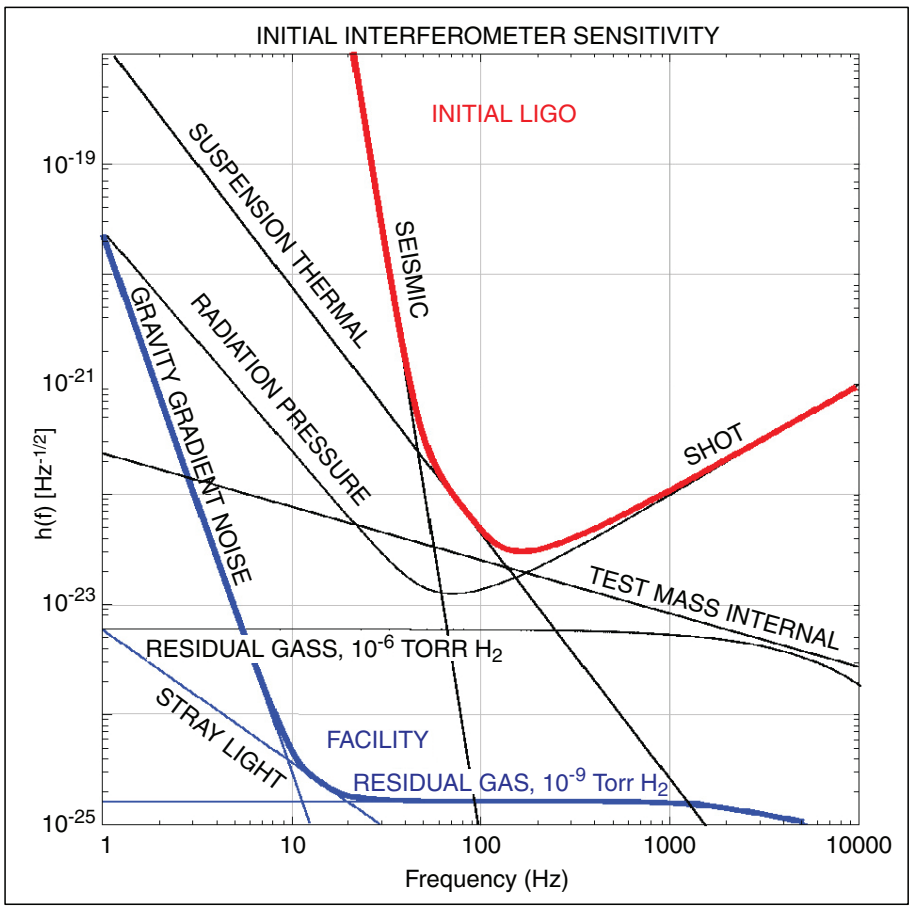

Figure 2. Sensitivity goals of the initial LIGO interferometers, and facility limits on the LIGO sensitivity (taken from [25]).

We now survey the more well-understood possible sources of measurable GWs in the high-frequency band. We emphasize at this point that such a listing of sources can in no way be considered comprehensive: we are hopeful that some GW sources may surprise us, as has been the case whenever we have studied the Universe with a new type of radiation.

6.1.1. Coalescing compact binaries. Compact binaries—binary star systems in which each member is a neutron star or black hole—are currently the best-understood sources of GWs. Double neutron stars have been studied observationally since the mid-1970s; five such systems [12]–[16], tight enough to merge within a few 10^8 or 10^9 years have been identified in our Galaxy. Extrapolation from these observed binaries in the Milky Way to the Universe at large [61]–[64] indicates that GW detectors should measure at least several and at most several hundred binary neutron star mergers each year (following detector upgrades; the expected rate for initial detectors is of the order of one event per several years, so that measurement of an event is plausible but of fairly low probability). Population synthesis (modelling evolution of stellar populations) indicates that the measured rate of binaries containing black holes should likewise be interestingly large (perhaps even for initial detectors) [65]–[68]. The uncertainties of population synthesis calculations are rather large, however, due to poorly understood aspects of stellar evolution

and compact binary formation; data from GW detectors is likely to have a large impact on this field.

6.1.2. Stellar core collapse. Core collapse in massive stars (the engine of type II supernova explosions) has long been regarded as likely to be an important source of GWs; see, for example, [69] for an early review. Stellar collapse certainly exhibits all of the *necessary* conditions for strong GW generation—large amounts of mass (1–100$M_\odot$) flow in a compact region (hundreds to thousands of kilometres) at relativistic speeds ($v/c \gtrsim 1/5$). However, these conditions are not sufficient to guarantee strong emission. In particular, the degree of asymmetry in collapse is not particularly well understood.

If the core of a star is very rapidly rotating during collapse, then instabilities may develop which lead to strong GW emission [70]. If such instabilities develop, core collapse GWs could be detected from events as far away as 10 Megaparsecs [71], a distance encompassing enough galaxies that several events per year would be likely. Most models of massive stars, however, indicate that such rapid rotation is not likely (e.g. [72]). Even without the growth of instabilities, the asymmetric dynamics of core collapse is likely to lead to wave emission which would be detectable within the Local Group of galaxies, with perhaps an event every few years detectable by advanced interferometers [73]. The wave strength is likely to correlate strongly with the degree of asymmetry in the supernova. If the GW event has an electromagnetic or neutrino counterpart, we may gain a wealth of knowledge regarding the state of the precollapse core [74].

6.1.3. Periodic emitters. Periodic sources of GWs radiate at constant or nearly constant frequency, like radio pulsars. In fact, the prototypical source of continuous GW is a rotating neutron star, or GW pulsar. A non-axisymmetry in a neutron star crust (caused, for example, by an oblateness that is misaligned with the star's spin axis) will radiate GWs with characteristic amplitude

$$h \sim \frac{G}{c^4}\frac{I f^2 \epsilon}{r}. \tag{6.2}$$

Here I is the star's moment of inertia, f is the wave frequency, r is the distance to the source and ϵ is the dimensionless fractional distortion $\epsilon = (I_{xx} - I_{yy})/I$, where I_{ij} is the moment of inertia tensor. The crucial parameter ϵ characterizes the degree to which the star is distorted; it is rather poorly understood. Upgraded interferometers in LIGO could set an upper limit on ϵ of order 10^{-6} for sources at $\sim$10 kpc [27]. Various mechanisms have been proposed to explain how a neutron star can be distorted to give a value of ϵ that is interesting as a GW source; see [75, 76] for further discussion. Examples of some interesting mechanisms include misalignment of a star's internal magnetic field with the rotation axis [77] and distortion by accreting material from a companion star [78, 79] (discussed in more detail below).

Whatever the mechanism generating the distortion, it is clear that ϵ will be small, so that $h \sim 10^{-24}$ or smaller—quite weak. Measuring these waves will require coherently tracking their signal for a large number of wave cycles. Coherently tracking N cycles boosts the signal-to-noise ratio by a factor $\sim\sqrt{N}$. This is actually fairly difficult, since the signal is strongly modulated by the Earth's rotation and orbital motion, 'smearing' the waves' power across multiple frequency bands. Searching for periodic GWs means demodulating the motion of the detector, a computationally intensive problem since the modulation is different for every sky position. Unless one knows in advance the position of the source, one needs to search over a huge number of sky position

'error boxes', perhaps as many as 10^{13}. One rapidly becomes computationally limited[10]. (Radio pulsar searches face this same problem, with the additional complication that radio pulses are dispersed by the interstellar medium. However, radio observers usually use shorter integration times, and often target their searches on small regions of the sky, so their computational cost is usually not as great.) For further discussion, see [80]; for ideas about doing hierarchical searches that require less computer power, see [81].

As mentioned above, one particularly interesting mechanism for distorting a neutron star comes from accretion of material from a companion star. Accretion provides a spin-up torque to a neutron star,

$$(\mathrm{d}J/\mathrm{d}t)_{\text{spin-up}} \sim R^2 \Omega_* \dot{M} \tag{6.3}$$

(where J is the spin angular momentum, Ω_* is the orbital frequency of the accreting matter as it plunges onto the star, R is the star's radius and $\dot{M}$ is the mass accretion rate). Without any kind of braking mechanism, the neutron star would presumably spin-up until it reaches the 'breakup limit', i.e., the spin frequency at which centrifugal forces would begin to break it apart; the breakup frequency is typically around 2000–3000 Hz.

Observations have shown [82] that accreting neutron stars do, in fact, appear to have a 'speed limit'—no accreting neutron star has been observed to spin faster than 619 Hz [83]. This is consistent with the fact that the fastest radio pulsar[11] has a spin period of 641 Hz [84]. This suggests that some mechanism is removing angular momentum from the neutron star. A plausible and very attractive possibility of how this angular momentum is removed is via GW emission. Because the spin-down torque due to GW emission grows sharply with spin frequency,

$$(\mathrm{d}J/\mathrm{d}t)_{\text{spin-down}} \propto \Omega^5 \quad \text{(quadrupole emission)}, \tag{6.4}$$

the limiting spin obtained by balancing the torques (6.2) and (6.4) is relatively insensitive to the mass accretion rate $\dot{M}$. Such a mechanism was originally suggested by Wagoner [85], and was revived by Bildsten [78] to explain the narrow clustering in the spin frequency of accreting low-mass x-ray binaries (LMXBs). Various mechanisms could provide the spin-down torque—Bildsten originally suggested that a quadrupole moment in the spinning star could be induced by a thermally varying electron capture mechanism, but also noted that the r-mode instability (see, e.g. [86] for a review) could be excited, leading to a similar spin-down law. Whatever the mechanism, accreting neutron stars are obvious and very attractive targets for observing campaigns with GW detectors, particularly given that their sky positions are well known.

[10] This rather large number of patches on the sky is driven by the possible need to search for high-frequency pulsars over several months of observation. The difference Δf between the Doppler frequency shifts for two adjacent sky patches separated by an angle $\delta\theta$ is of order $\Delta f \sim v_\oplus f \delta\theta/c$, where $v_\oplus \sim 3 \times 10^4\,\mathrm{ms}^{-1}$ is the Earth's orbital velocity and f is the gravitational wave frequency. The phase error over an observation time T_{obs} is of order $\Delta f T_{\text{obs}}$. Demanding that this be less than unity yields $\delta\theta \lesssim c/(v_\oplus f T)$. The number of independent sky patches is then $N_{\text{p}} \sim 4\pi\delta\theta^{-2} \sim 4\pi v_\oplus^2 f^2 T_{\text{obs}}^2/c^2 \sim 10^{13}$ for $f = 1000\,\text{Hz}$ and $T_{\text{obs}} = 1/3$ year. Fewer positions would be needed if either the maximum frequency or the integration time is reduced; the figures given here set the maximum values that are plausible. See [80] for more details.

[11] The so-called 'recycled' radio pulsars spin at frequencies $\sim$ several hundred Hertz; they are believed to be the fossils of accreting neutron stars.

6.1.4. Stochastic backgrounds. Stochastic backgrounds are 'random' GWs, arising from a large number of independent, uncorrelated sources that are not individually resolvable. A particularly interesting source of stochastic waves is the dynamics of the early Universe, which could produce an all-sky GW background, similar to the cosmic microwave background; see [87]–[89] for detailed reviews. Stochastic waves can be generated in the early Universe via a variety of mechanisms: amplification of primordial fluctuations in the Universe's geometry via inflation, phase transitions as previously unified interactions separate, a network of vibrating cosmic strings, or the condensation of a brane from a higher-dimensional space. These waves can actually extend over a wide range of frequency bands; waves from inflation, in particular, span all bands, from ultra low frequency to high frequency.

Stochastic backgrounds are usually idealized as being stationary, isotropic and homogeneous. They are thus characterized by their energy density per unit frequency, $\mathrm{d}\rho_{\mathrm{gw}}/\mathrm{d}f$. This is often parametrized in terms of the energy density per unit logarithmic frequency divided by the critical energy density to close the Universe

$$\Omega_{\mathrm{gw}}(f) = \frac{1}{\rho_{\mathrm{crit}}}\frac{\mathrm{d}\rho_{\mathrm{gw}}}{\mathrm{d}\ln f}, \tag{6.5}$$

where $\rho_{\mathrm{crit}} = 3H_0^2/8\pi G$ is the critical density and H_0 is the value of the Hubble constant today. Different cosmological sources produce different levels of $\Omega_{\mathrm{gw}}(f)$, centred in different bands. In the high-frequency band, waves produced by inflation are likely to be rather weak: estimates suggest that the spectrum will be flat across LIGO's band, with magnitude $\Omega_{\mathrm{gw}} \sim 10^{-15}$ at best [90]. Waves from phase transitions can be significantly stronger, but are typically peaked around a frequency that depends on the temperature T of the phase transition [87, 91]:

$$f_{\mathrm{peak}} \sim 100\,\mathrm{Hz}\left(\frac{T}{10^5\,\mathrm{TeV}}\right). \tag{6.6}$$

Because of their random nature, stochastic GWs look just like noise. Ground-based detectors will measure stochastic backgrounds by comparing data at multiple sites and looking for 'noise' that is correlated [88, 92]. For comparing to a detector's noise, one should construct the characteristic stochastic wave strain,

$$h \propto f^{-3/2}\sqrt{\Omega_{\mathrm{gw}}(f)\Delta f}, \tag{6.7}$$

where Δf is the frequency band across which the measurement is made. For further discussion and the proportionality constants, see [88]. Note that if $\Omega_{\mathrm{gw}}(f)$ is constant, this strain level grows sharply with decreasing frequency—the most interesting limits are likely to be set by measurements at low frequencies.

Early detectors will have fairly poor sensitivity to the background, constraining it to a level $\Omega_{\mathrm{gw}} \sim 5 \times 10^{-6}$ in a band from about 100 to 1000 Hz. This is barely more sensitive than known limits from cosmic nucleosynthesis [87]. Later upgrades will be significantly more sensitive, able to detect waves with $\Omega_{\mathrm{gw}} \sim 10^{-10}$, which is good enough to place interesting limits on backgrounds from some phase transitions.

6.2. Low frequency

There is no hope of measuring GWs in the low-frequency band, $10^{-5} \lesssim f \lesssim 1\,\mathrm{Hz}$, using a ground-based instrument: even if it were possible to completely isolate one's instrument from

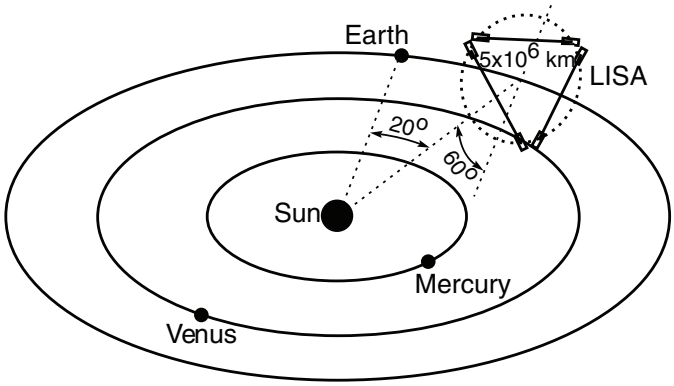

Figure 3. Orbital configuration of the LISA antenna.

local ground motions, gravitational coupling to fluctuations in the local mass distribution ultimately limits the sensitivity to frequencies $f \gtrsim 1\,\mathrm{Hz}$. Nonetheless, many extremely interesting sources radiate in this band. The only way to measure these waves is to build a detector in the quiet environment of space, far removed from low-frequency noise sources.

Such an instrument is currently being designed jointly by NASA in the United States and ESA, the European Space Agency: LISA, the Laser Interferometer Space Antenna. If all goes well, LISA will be launched into orbit in 2013 or so. LISA will be a laser interferometer, similar in concept to the ground-based detectors: changes in the distance between widely separated test masses will be monitored for GWs. However, LISA's scale is vastly larger than that of its ground-based cousins, and so details of its operations are quite different. In particular, LISA will have armlengths $L \simeq 5 \times 10^6\,\mathrm{km}$. The three spacecrafts which delineate the ends of LISA's arms are placed into orbits such that LISA forms a triangular constellation orbiting the sun, inclined 60° with respect to the plane of the ecliptic and following the Earth with a 20° lag. This configuration is sketched in figure 3. Since it essentially shares Earth's orbit, the constellation orbits the sun once per year, 'rolling' as it does so. This motion plays an important role in pinpointing the position of sources by modulating the measured waveform—the modulation encodes the source location and makes position determination possible.

Each of the three spacecrafts contains two optical assemblies, each of which houses a 1 W laser and a 30 cm telescope. Because of the extreme lengths of the interferometer's arms, Fabry–Perot interferometry as in the ground-based detectors is not possible: diffraction spreads the laser beam over a diameter of about 20 km as it propagates from one spacecraft to the other. A portion of that 20 km wavefront is sampled with the telescope. That light is then interfered with a sample of light from the on-board laser. Each spacecraft thus generates two interference data streams; six signals are generated by the full LISA constellation. From these six signals, we can construct the time variations of LISA's armlengths and then build both GW polarizations. More information and details can be found in [93].

Note that the LISA armlengths are not constant—as the constellation orbits, the distances between the various spacecrafts vary by about 1% (including effects such as planetary perturbations). These variations are far larger than the displacements produced by GWs, which are of the order of picometers. However, these variations occur over timescales of the order of months, and are extremely smooth and well modelled. It will not be difficult to remove them from the data leaving clean data in the interesting frequency band. Picometer scale variations are not too

difficult to measure in this band by gathering photons for timescales 10 seconds $\lesssim T \lesssim$ 1 day. Even though the bulk of the laser's emitted power is lost by diffraction, enough photons are gathered on this timescale that the phase shift due to the GW can be determined.

The GW signals are actually read out by monitoring the position of the so-called 'gravitational sensor' on each optical assembly; in particular, the position of a 'proof mass' which floats freely and constitutes the test mass for the LISA antenna is monitored. Because it is freely floating, the proof mass follows a geodesic of the spacetime. MicroNewton thrusters keep the bulk spacecraft centred on these proof masses, forcing the craft to follow their average trajectory. In this way, the spacecrafts are isolated from low-frequency forces that could impact the ability to measure GWs (e.g. variations in solar radiation pressure).

We now take a quick tour through some interesting LISA sources.

6.2.1. Periodic emitters. In the high-frequency band, the source of most periodic GWs is expected to be isolated neutron stars. LISA's periodic GWs will come primarily from binary star systems in the Milky Way, primarily close white dwarf binaries. Most of these systems do not generate waves strong enough to backreact significantly, so that their frequencies do not change measurably over the course of LISA's observations. Certain systems are well known in advance to be sources of periodic waves for the LISA band. These sources are understood well enough from optical observations that they may be regarded as 'calibrators'—LISA should detect them or else something is wrong!

Aside from these sources that are known in advance, it is expected that LISA will discover a good number of binary systems that are too faint to detect with telescopes. Joint observations by LISA and other astronomical instruments are likely to be more fruitful than observations with a single instrument alone. For example, it is typically difficult for telescopes to determine the inclination of a binary to the line of sight (a quantity needed to help pin down the masses of the binary's members). GWs measure the inclination angle almost automatically, since this angle determines the relative magnitude of the polarizations h_+ and $h_\times$.

The total number of periodic binaries radiating in LISA's band is expected to be so large that they will constitute a confused, stochastic background at low frequencies—there are likely to be several thousand galactic binaries radiating in each resolvable frequency bin. This background will constitute a source of 'noise' (from the standpoint of measuring other astrophysical sources) that is larger than that intrinsic to the instrument noise at $f \lesssim 10^{-3}$ Hz.

6.2.2. Coalescing binary systems containing black holes. Coalescing binary black hole systems will be measurable by LISA to extremely large distances—essentially to the edge of the observable Universe. Even if such events are very rare, the observed volume is enormous so an interesting event detection rate is very likely. One class of such binaries consists of systems in which the member holes are of roughly equal mass ($\sim 10^5$–$10^8 M_\odot$). These binaries can form following the merger of galaxies (or pregalactic structures) containing black holes in their cores. Depending on the mass of the binary, the waves from these coalescences will be detectable to fairly large redshifts ($z \sim 5$–10), possibly probing an early epoch in the formation of the Universe's structure [94].

The other major class of binary systems consists of relatively small bodies (black holes with mass $\sim$10–$100 M_\odot$, neutron stars or white dwarfs) that are captured by larger black holes ($M \sim 10^5$–$10^7 M_\odot$). These 'extreme mass ratio' binaries are created when the smaller body is captured onto an extremely strong field, highly relativistic orbit, generating strong GWs. Such

systems are measurable to a distance of a few Gigaparsecs if the in-spiralling body is a $10M_{\odot}$ black hole, and to a distance of a few hundred Megaparsecs if the body is a neutron star or white dwarf. LISA will measure the waves that come from the last year or so of the smaller body's in-spiral, and thence probe the nature of the larger black hole's gravitational field deep within the hole's potential. The rates for such events are not too well understood and depend on the details of stellar dynamics in the cores of galaxies. Extremely conservative estimates typically find that the rate of measurable events for LISA should be at least several per year [95, 96]. Recent thinking suggests that these rates are likely to be rather underestimated—black holes (which are measurable to much greater distances) are likely to dominate the measured rate, perhaps increasing the rate to several dozen or several hundred per year [97].

Finally, it is worth noting that many events involving intermediate mass black holes—those with masses in the band running from a few 10^2 to a few $10^5 M_{\odot}$—would generate GWs in LISA's sensitive band. There is a large body of tentative evidence for the existence of black holes in this mass band (see, e.g. [98] for a review), though as yet we have no 'smoking gun' unambiguous signature for such a hole. If such black holes do exist and undergo mergers in sufficient numbers, measurement of their waves will make possible a wealth of interesting tests of relativity [99], and could untangle some of the mysteries surrounding supermassive black hole formation and growth.

6.2.3. Stochastic backgrounds. As discussed in section 6.1.4, ground-based detectors can measure a stochastic background by correlating the data streams of widely separated detectors. LISA will use a slightly different technique: by combining its six data streams in an appropriate way, one can construct an observable that is completely insensitive to GWs, measuring noise only [100]. This makes it possible to distinguish between a noise-like stochastic background and true instrumental noise, and thereby to learn about the characteristics of the background [101].

The sensitivity of LISA will not be good enough to set interesting limits on an inflationary GW background: LISA will only reach $\Omega_{\rm gw} \sim 10^{-11}$, about four orders of magnitude too large to begin to say something about inflation [90]. However, LISA's band is well placed for other possible sources of cosmological backgrounds. In particular, an electroweak phase transition at temperature $T \sim$100–1000 GeV would generate waves in LISA's band (cf equation (6.6)). These waves are likely to be detectable if the phase transition is strongly first order (a scenario that does not occur in the standard model, but is conceivable in extensions to the standard model [91]).

6.3. Very low frequency

The very low-frequency band, $10^{-9} \lesssim f \lesssim 10^{-7}$ Hz, corresponds to waves with periods ranging from a few months to a few decades. Our best limits on waves in this band come from observations of millisecond pulsars. First suggested by Sazhin [102] and then carefully analysed and formulated by Detweiler [103], GWs can drive oscillations in the arrival times of pulses from a distant pulsar. Millisecond pulsars are very good 'detectors' for measurements in this band because they are exquisitely precise clocks. The range of frequencies encompassed by the very low-frequency band is set by the properties of these radio pulsar measurements: the high end of the frequency band comes from the need to integrate the radio pulsar data for at least several months; the low end comes from the fact that we have only been observing millisecond pulsars for a few decades. (One cannot observe a periodicity shorter than the span of one's dataset.) A

recent upper limit derived from this technique is [104]

$$\Omega_{\rm gw} h_{100}^2 < 9.3 \times 10^{-8}, \qquad 4 \times 10^{-9} < f < 4 \times 10^{-8}\,{\rm Hz} \tag{6.8}$$

(where the limit is a 95% confidence limit and h_{100} is the Hubble constant in units of 100 km second^{-1} Mpc^{-1}).

The upper limit (6.8) already places constraints on some cosmological models (in particular those involving cosmic strings). With further observations and the inclusion of additional pulsars in the datasets, it is likely to improve quite soon. It is possible that the background in this band will be dominated by many unresolved coalescing massive binary black holes [105]—binaries that are either too massive to radiate in the LISA band, or else are in-spiralling towards the LISA band en route to a final merger several centuries or millenia hence. Constraints from pulsar observations in this band will remain an extremely important source of data on stochastic waves in the future—the limits they can set on $\Omega_{\rm gw}$ are likely to be better than can be set by any of the laser interferometric detectors.

6.4. Ultra low frequency

The ultra low-frequency band, $10^{-18} \lesssim f \lesssim 10^{-13}$ Hz, is better described by converting from frequency to wavelength: for these waves, $10^{-5}\, H_0^{-1} \lesssim \lambda \lesssim H_0^{-1}$, where $H_0^{-1} \sim 10^{10}$ light years is the Hubble length today. Waves in this band oscillate on scales comparable to the size of the Universe. They are most likely to be generated during inflation: quantum fluctuations in the spacetime metric are parametrically amplified during inflation to relatively high amplitude. The RMS amplitude to which the waves are amplified depends upon the energy scale $E_{\rm infl}$ of inflation:

$$h_{\rm rms} \propto \left(\frac{E_{\rm infl}}{m_{\rm P}}\right)^2, \tag{6.9}$$

where $m_{\rm P}$ is the Planck mass. Measuring these GWs would be a direct probe of inflationary physics and would determine the inflation energy scale, which is currently unknown to within many orders of magnitude. These waves have been described as the 'smoking gun' signature of inflation [106].

During inflation, quantum fluctuations impact both the scalar field which drives inflation (the inflaton ϕ) and the metric of spacetime. There exist independent scalar fluctuations (coupled fluctuations in the inflaton and scalar-type fluctuations in the metric) and tensor fluctuations (tensor-type fluctuations in the metric). The Fourier modes of these scalar and tensor perturbations are describable as harmonic oscillators in the expanding Universe [107]. Each mode undergoes zero-point oscillations in the harmonic potential. However, the potential itself is evolving due to the expansion of the Universe. The evolution of this potential parametrically amplifies these zero-point oscillations, creating quanta of the field [87]. During inflation, the Universe's scale factor $a(t)$ grows faster than the Hubble length H^{-1}, and so each mode's wavelength likewise grows faster than the Hubble length. The mode's wavelength eventually becomes larger than the Hubble length, or the mode 'leaves the horizon'. After inflation ends the mode subsequently re-enters the horizon. For gravitational perturbations, the number of quanta generated in the mode is proportional to the factor by which the Universe expands between the two different horizon crossings. Fluctuations in the inflaton seed density fluctuations, $\delta\rho(\vec{r}) = \delta\phi(\vec{r})(\partial V/\partial\phi)$ (where

$V(\phi)$ is the potential that drives the inflaton field). The tensor-type fluctuations in the spacetime metric are GWs.

Density fluctuations and GWs both leave an imprint upon the cosmic microwave background (CMB). First, each contributes to the CMB temperature anisotropy. However, even a perfectly measured map of temperature anisotropy cannot really determine the contribution of GWs very well because of cosmic variance: since we only have one Universe to observe, we are sharply limited in the number of statistically independent influences upon the CMB that we can measure. Large angular scales are obviously most strongly affected by this variance, and these scales are the ones on which GW most importantly impact the CMB [108].

Fortunately, the scalar and tensor contributions also impact the polarization of the CMB. These two contributions can be detangled from one another in a model-independent fashion. This detangling uses the fact that the polarization tensor $P_{ab}(\hat{\mathbf{n}})$ on the celestial sphere can be decomposed into tensor harmonics. These harmonics come in two flavours, distinguished by their parity properties: the 'E-modes' or 'gradient-type' harmonics $Y^{E}_{(lm)ab}(\hat{\mathbf{n}})$ (which pick up a factor $(-1)^l$ under $\hat{\mathbf{n}} \to -\hat{\mathbf{n}}$), and the 'B-modes' or 'curl-type' harmonics $Y^{C}_{(lm)ab}(\hat{\mathbf{n}})$ (which pick up a factor $(-1)^{l+1}$ under $\hat{\mathbf{n}} \to -\hat{\mathbf{n}}$). These harmonics are constructed by taking covariant derivatives on the sphere of the 'ordinary' spherical harmonics $Y_{lm}(\hat{\mathbf{n}})$; see [109] for details. Because scalar perturbations have no handedness, they only induce gradient-type polarization. GWs induce both gradient- and curl-type polarization. Thus, an unambiguous detection of the curl-type polarization would confirm production of GWs by inflation. (A caveat is that gravitational lensing can convert E-modes to B-modes; this so-called 'cosmic shear' ultimately limits the sensitivity to GWs of CMB polarization studies [110].)

7. Conclusion

This paper has summarized many of the most important topics in the theory of GWs. Due to space and time limitations, we sadly were not able to cover all topics with which students of this field should be familiar. In particular, we had hoped to include a discussion of strong field relativity and GW emission. We confine ourselves, in this conclusion, to a (very) brief discussion of important aspects of this subject for GW science, as well as pointers to the relevant literature.

Linearized theory as described in sections 2 and 5 is entirely adequate to describe the propagation of GWs through our Universe and to model the interaction of GWs with our detectors. In some cases, it is even adequate to describe the emission of waves from a source, as described in section 4 (although for sources with non-negligible self-gravity such as binary star systems, one has to augment linearized theory as described in section 4.2). However, many sources have very strong self-gravity where the linearized treatment is completely inadequate. A variety of formalisms have been developed to handle these cases.

- Post-Newtonian (PN) theory. PN theory is one of the most important of these formalisms, particularly for modelling binary systems. Roughly speaking, PN theory analyses sources using an iterated expansion in two variables: the 'gravitational potential', $\phi \sim M/r$, where M is a mass scale and r characterizes the distance from the source; and velocities of internal motion, v. (In linearized theory, we assume ϕ is small but place no constraints on v.) Newtonian gravity emerges as the first term in the expansion, and higher-order corrections are found as the expansion is iterated to ever higher order. Our derivation of the quadrupole

formula in section 4.2 gives the leading order term in the PN expansion of the emitted radiation. See Blanchet's recent review [111] and references therein for a comprehensive introduction to and explanation of this subject.

- Numerical relativity. Numerical relativity seeks to directly integrate Einstein's equations on a computer. Ideally, we would like to use a well-understood model of a GW source (e.g. a binary system in which the field strengths are small enough that it is well described by post-Newtonian theory) as 'initial data', and then numerically evolve the Einstein equations from that point to some final equilibrium configuration. The form in which we normally encounter Einstein's equation in textbooks is not well suited to this task—the coordinate freedom of general relativity means that there is no notion of 'time' built into the equation $G_{ab} = 8\pi T_{ab}$. One must introduce some notion of time for the concept of 'initial data' to have any meaning. The four dimensions of spacetime are then split into 3+1 dimensions of space and time. Having made this choice, Einstein's equations take on a particular form which is amenable to numerical computation.

  We recommend the reviews of numerical relativity by Lehner [112] and by Baumgarte and Shapiro [37]. For the purpose of our present discussion, it suffices to remark that it has proven to be *extremely* difficult to model some of the most interesting and important GW sources. In particular, the final stage of binary black hole mergers—regarded by many as the 'Holy Grail' of numerical relativity—has proven to be quite a challenge.
- Perturbation theory. In some cases, GW sources can be modelled as nearly, but not quite, identical to some exact solution of the Einstein field equations. For example, the end state of a binary black hole coalescence must be a single black hole. As we approach this final state, the system will be well-modelled as the Kerr black hole solution, plus some distortion that radiates away. Another example is a binary consisting of a stellar mass compact body orbiting a massive black hole. The binary's spacetime will be well described as a single black hole plus a perturbation due to the captured body. These cases can be nicely described using perturbation theory: we treat the spacetime as some exact background, g^{B}_{ab}, plus a perturbation h_{ab}:

$$g_{ab} = g^{\mathrm{B}}_{ab} + h_{ab}. \tag{7.1}$$

  We are in the perturbative regime if $\|h_{ab}\|/\|g^{\mathrm{B}}_{ab}\| \ll 1$. This system can then be analysed by expanding the Einstein equations for this metric and keeping terms to first order in h_{ab} (see section 5.1 for details but without the matter source terms included).

  This approach has proven to be particularly fruitful when the background spacetime is that of a black hole. For the case of a Schwarzschild background, the derivation of the full perturbation equations is rather straightforward; Rezzolla gives a particularly compact and readable summary [113]. Perturbations of Kerr black holes are not nearly so simple to describe, largely due to the lack of spherical symmetry—expanding the metric as in equation (7.1) does not prove to be so fruitful as it is in the Schwarzschild case. Somewhat miraculously, it turns out that progress can be made by expanding the curvature tensor: by expanding the Riemann tensor as $R_{abcd} = R^{\mathrm{B}}_{abcd} + \delta R_{abcd}$ and taking an additional derivative of the Bianchi identity,

$$\nabla_e R_{abcd} + \nabla_d R_{abec} + \nabla_c R_{abde} = 0, \tag{7.2}$$

  one can derive a wave-like equation for the perturbation δR_{abcd}. This analysis was originally performed by Teukolsky; see his original analysis [114] for details.

Acknowledgments

We thank R Price and J Pullin for the invitation to write this paper, and are profoundly grateful to T Smith at the Institute of Physics for patiently and repeatedly extending our deadline as we wrote this paper. We are grateful to two anonymous referees for detailed and helpful comments, as well as to helpful comments from Hongbao Zhang. SAH thanks the Caltech TAPIR group and the Kavli Institute for Theoretical Cosmology at the University of Chicago for their hospitality while this paper was completed. EEF is supported by NSF grant PHY-0140209; SAH is supported by NSF grant PHY-0244424 and NASA grant NAGW-12906.

Appendix A. Existence of TT gauge in local vacuum regions in linearized gravity

In this appendix, we show that one can always find TT gauges in local vacuum regions in linearized gravity. More precisely, suppose that $\mathcal{V}$ is a connected open spatial region and (t_0, t_1) is an open interval of time. Then one can find a gauge on the product $\mathcal{R} \equiv (t_0, t_1) \times \mathcal{V}$ that satisfies $h_{tt} = h_{ti} = \delta^{ij} h_{ij} = \partial_i h_{ij} = 0$, as long as $T_{ab} = 0$ throughout $\mathcal{R}$.

The proof involves a generalization of the gauge-invariant formalism of section 2.2 to finite spacetime regions. We define a decomposition of the metric perturbation h_{ab} in terms of quantities ϕ, β_i, γ, h_{ij}^{TT}, H, ε_i and λ using the same equations (2.29)–(2.35) as before. However we replace the boundary conditions (2.36) with

$$\gamma_{|\partial\mathcal{V}} = \int_{t_0}^{t} \mathrm{d}t\, \phi_{|\partial\mathcal{V}}, \tag{A.1}$$

$$\lambda_{|\partial\mathcal{V}} = 0, \tag{A.2}$$

$$(\nabla^2\lambda)_{|\partial\mathcal{V}} = H_{|\partial\mathcal{V}}, \tag{A.3}$$

$$(\mathbf{n} \times \boldsymbol{\varepsilon})_{|\partial\mathcal{V}} = 2 \int_{t_0}^{t} \mathrm{d}t\, (\mathbf{n} \times \boldsymbol{\beta})_{|\partial\mathcal{V}}, \tag{A.4}$$

where $\mathbf{n}$ is the unit outward-pointing unit normal to $\partial\mathcal{V}$. The reason for this particular choice of boundary conditions will be explained below. These boundary conditions define a unique decomposition of the metric within $\mathcal{R}$.

Next, we compute how the variables ϕ, β_i, γ, h_{ij}^{TT}, H, ε_i and λ transform under general gauge transformations. We use the same parametrization (2.37) of the gauge transformation as before, except that we impose now the boundary condition $C_{|\partial\mathcal{V}} = 0$. We find that the transformation laws (2.38)–(2.44) are replaced by the following equations which contain some extra terms:

$$\phi \to \phi - \dot{A}, \tag{A.5}$$

$$\beta_i \to \beta_i - \dot{B}_i - \partial_i \psi, \tag{A.6}$$

$$\gamma \to \gamma - A - \dot{C} + \psi, \tag{A.7}$$

$$H \to H - 2\nabla^2 C, \tag{A.8}$$

$$\lambda \to \lambda - 2C, \tag{A.9}$$

$$\varepsilon_i \to \varepsilon_i - 2B_i + 2\eta_i - 2(t-t_0)\partial_i\psi, \tag{A.10}$$

$$h_{ij}^{\mathrm{TT}} \to h_{ij}^{\mathrm{TT}} - 2\partial_{(i}\eta_{j)} + 2(t-t_0)\partial_i\partial_j\psi. \tag{A.11}$$

Here ψ is the time-independent, harmonic function defined by $\nabla^2\psi = 0$ and $\psi_{|\partial\mathcal{V}} = A_{|\partial\mathcal{V},t=t_0}$. Similarly η_i is the time-independent, harmonic transverse vector defined by $\nabla^2\eta_i = 0$ and $(\mathbf{n}\times\mathbf{n})_{|\partial\mathcal{V}} = \mathbf{n}\times\mathbf{B}_{|\partial\mathcal{V},t=t_0}$.

We define the variables Φ, Θ and Ξ_i by the same equations (2.45)–(2.47) as before. From the transformation laws (A.5)–(A.11) these variables are still gauge-invariant, while h_{ij}^{TT} is no longer gauge-invariant in the present context. Next, imposing the linearized vacuum Einstein equations using the expressions (2.59)–(2.61) yields

$$\nabla^2\Theta = 0, \qquad \nabla^2\Xi_i = -2\partial_i\dot{\Theta}, \qquad \nabla^2\Phi = \tfrac{3}{2}\ddot{\Theta} \tag{A.12}$$

in $\mathcal{V}$. The boundary conditions (A.1)–(A.4) together with the definitions (2.45)–(2.47) imply that the boundary conditions on the gauge-invariant variables are

$$\Phi_{|\partial\mathcal{V}} = \Theta_{|\partial\mathcal{V}} = \Xi^i_{|\partial\mathcal{V}} = 0. \tag{A.13}$$

(This is why we choose those particular boundary conditions.) Therefore, all the gauge-invariant variables vanish, $\Theta = \Phi = \Xi^i = 0$ in $\mathcal{R}$.

It is now straightforward to show that one can choose a gauge in which $\phi = \beta_i = \gamma = H = \varepsilon_i = \lambda = 0$. From the transformation laws (A.5)–(A.11), we can choose C to make $\lambda = 0$, choose $\dot{A}$ to make $\phi = 0$ and choose $\dot{B}_i$ to make $\beta_i = 0$. The residual gauge freedom is then parametrized by functions A and B_i that are time-independent. Next, from equation (A.13) together with the definitions (2.45)–(2.47) it follows that

$$0 = \Theta = \tfrac{1}{3}H, \tag{A.14}$$

$$0 = \Phi = -2\dot{\gamma}, \tag{A.15}$$

$$0 = \Xi_i = -\tfrac{1}{2}\dot{\varepsilon}_i. \tag{A.16}$$

Thus the only remaining non-zero pieces of the metric other than the TT piece are γ and ε_i, and these are both time-independent. Finally, we can use the residual gauge freedom given by time-independent functions A and B_i to set γ and ε_i to zero, by equations (A.7) and (A.10). (For this purpose A and B_i will vanish on $\partial\mathcal{V}$, by equations (A.1) and (A.4), so ψ and η_i vanish.)

References

[1] Einstein A 1905 On the electrodynamics of moving bodies *Ann. Phys.* **17** 891
[2] Newton I 1687 *Philosophiae Naturalis Principia Mathematica* 2nd edn (London: Streater); from the *General Scholum* added at the end of the 3rd book in the 2nd edn of 1713
[3] Newton I 1687 *Philosophiae Naturalis Principia Mathematica* 1st edn (London: Streater). All quotes from the *Principia* are taken from [4], Box 1.10
[4] Misner C W, Thorne K S and Wheeler J A 1973 *Gravitation* (San Francisco: Freeman)

[5] Einstein A 1915 On the general theory of relativity *Sitzungsberichte Preußische Akademie der Wissenschaften Berlin* 778 (1915) (English translation in *Math. Phys.*) 799 (1915)
[6] Einstein A 1915 On the field equations of gravitation *Sitzungsberichte Preußische Akademie der Wissenschaften Berlin* (*Math. Phys.*) 844 (1915)
[7] Schutz B F 1984 Gravitational waves on the back of an envelope *Am. J. Phys.* **52** 412
[8] Einstein A 1916 Approximative integration of the field equations of gravitation *Sitzungsberichte Preußische Akademie der Wissenschaften Berlin* (*Math. Phys.*) 688
[9] Einstein A 1918 On gravitational waves *Sitzungsberichte Preußische Akademie der Wissen-schaften Berlin* (*Math. Phys.*) 154
[10] Kennefick D 1997 Controversies in the history of the radiation reaction problem in general relativity *PhD Thesis* (part II), California Institute of Technology; short version available as *Preprint* gr-qc/9704002
[11] Hulse R A and Taylor J H 1975 Discovery of a pulsar in a binary system *Astrophys. J.* **195** L51
[12] Weisberg J M and Taylor J H 2004 Relativistic binary pulsar B1913+16 thirty years of observations and analysis *Binary Radio Pulsars, Proc. Aspen Conf, ASP Conf. Series* ed F A Rasio and I H Stairs, at press (*Preprint* astro-ph/0407149)
[13] Stairs I H, Thorsett S E, Taylor J H and Wolszczan A 2002 Studies of the relativistic binary pulsar PSR B1534+12. I. Timing analysis *Astrophys. J.* **581** 501
[14] Deich W T S and Kulkarni S R 1996 The masses of the neutron stars in M15C *Compact Stars in Binaries, Proc. IAU Symp. 165* ed J van Paradijs, E P J van den Heuvel and E Kuulkers (Dordrecht: Kluwer Academic) p 279
[15] Burgay M *et al* 2003 An increased estimate of the merger rate of double neutron stars from observations of a highly relativistic system *Nature* **426** 531
[16] Faulker A J *et al* 2004 PSR J1756-2251: a new relativistic double neutron star system *Astrophys. J.* in press (*Preprint* astro-ph/0411796)
[17] Abbott B *et al* 2004 Setting upper limits on the strength of periodic gravitational waves from PSR J1939+2134 using the first science data from the GEO 600 and LIGO detectors *Phys. Rev.* D **69** 082004
[18] Abbott B *et al* 2004 First upper limits from LIGO on gravitational wave bursts *Phys. Rev.* D **69** 102001
[19] Abbott B *et al* 2004 Analysis of LIGO data for gravitational waves from binary neutron stars *Phys. Rev.* D **69** 122001
[20] Abbott B *et al* 2004 Analysis of first LIGO science data for stochastic gravitational waves *Phys. Rev.* D **69** 122004
[21] Aufmuth P and Danzmann K 2005 *New J. Phys.* **7** 202
[22] Bonazzola S and Marck J-A 1994 Astrophysical sources of gravitational radiation *Ann. Rev. Nucl. Part. Sci.* **44** 655
[23] Schutz B F 1999 Gravitational wave astronomy *Class. Quantum Grav.* **16** A131
[24] Grishchuk L P, Lipunov V M, Postnov K A, Prokhorov M E and Sathyaprakash B S 2001 Gravitational wave astronomy: in anticipation of first sources to be detected *Phys. Uspekhi* **44** 1
Grishchuk L P, Lipunov V M, Postnov K A, Prokhorov M E and Sathyaprakash B S 2001 *Usp. Fiz. Nauk.* **171** 3 (*Preprint* astro-ph/0008481)
[25] Hughes S A, Marká S, Bender P L and Hogan C J 2001 New physics and astronomy with the new gravitational-wave observatories *Proc. APS/DPF/DPB summer study on the future of particle physics* (*Snowmass 2001*) ed R Davidson and C Quigg, eConf C010630, P402 (*Preprint* astro-ph/0110349)
[26] Schutz B F 2001 Lighthouses of gravitational wave astronomy *Lighthouses of the Universe: The Most Luminous Celestial Objects and Their Use for Cosmology, Proc. MPA/ESO Conf. Lighthouses of the Universe* p 207 (*Preprint* gr-qc/0111095)
[27] Cutler C and Thorne K S 2002 An overview of gravitational-wave sources *Proc. 16th Int. Conf. on General Relativity and Gravitation* (GR16) (*Preprint* gr-qc/0204090)
[28] Finn L S 1999 Gravitational radiation sources and signatures (*Gravity: From the Hubble Length to the Planck Length Lectures given at the XXVI SLAC Summer Institute on Particle Physics*) ed L Dixon (Springfield, VA: National Technical Information Service) (*Preprint* gr-qc/9903107)

[29] Hughes S A 2003 Listening to the Universe with gravitational-wave astronomy *Ann. Phys.* **303** 142
[30] Grishchuk L P 2004 Update on gravitational-wave research *Astrophysics Update* ed J W Mason (Springer-Praxis) at press (*Preprint* gr-qc/0305051)
[31] Thorne K S 1983 The theory of gravitational radiation: an introductory review *Gravitational Radiation* ed N Deruelle and T Piran (Amsterdam: North-Holland)
[32] Schutz B F and Ricci F 2001 Gravitational waves, sources and detectors *Gravitational Waves* ed I Ciufolini, V Gorini, U Moschella and P Fré (Bristol: Institute of Physics Publishing)
[33] Allen B 1996 *Proc. Les Houches School on Astrophysical Sources of Gravitational Waves* ed J-A Marck and J-P Lasota (Cambridge: Cambridge University Press)
[34] Ciufolini I, Gorini V, Moschella U and Fré P (ed) 2001 *Gravitational Waves* (Bristol: Institute of Physics Publishing UK)
[35] Bradaschia C (ed) 2004 *Proc. 5th Edoardo Amaldi Conf. on Gravitational Waves (Class. Quantum Grav.* **21**)
[36] Hartle J B 2003 *Gravity: An Introduction to Einstein's General Relativity* (San Francisco: Addison-Wesley)
[37] Baumgarte T W and Shapiro S L 2003 Numerical relativity and compact binaries *Phys. Rep.* **376** 41
[38] Van Bladel J 1991 Lorenz or Lorentz? *IEEE Antennas Prop. Mag.* **33** 69
[39] Griffiths D J 1999 *Introduction to Electrodynamics* 3rd edn (Englewood Cliffs, NJ: Prentice-Hall). Note, the footnote on 'Lorenz vs. Lorentz' does not appear in the 1st printing, but *does* appear in the 4th printing.
[40] Eddington A S 1922 The propagation of gravitational waves *Proc. R. Soc.* A **102** 268
[41] Bardeen J M 1980 Gauge-invariant cosmological perturbations *Phys. Rev.* D **22** 1882
[42] Bertschinger E 1999 Gravitation in the weak-field limit *Physics 8.962* (*MIT's General Relativity Course*) *Lecture Notes* (Spring)
[43] Larson S L, Hiscock W A and Hellings R W 2000 Sensitivity curves for spaceborne gravitational wave interferometers *Phys. Rev.* D **62** 062001
[44] Abramovici A *et al* 1992 *Science* **256** 325
[45] Thorne K S 1987 *300 years of Gravitation* (ed S W Hawking and W Israel) (Cambridge: Cambridge University Press) p 330
[46] Hughes S A and Thorne K S 1998 Seismic gravity-gradient noise in interferometric gravitational-wave Detectors *Phys. Rev.* D **58** 122002
[47] Thorne K S and Winstein C J 1999 Human gravity-gradient noise in interferometric gravitational-wave detectors *Phys. Rev.* D **60** 082001
[48] Creighton T 2000 Tumbleweeds and airborne gravitational noise sources for LIGO *Phys. Rev.* D submitted (*Preprint* gr-qc/0007050)
[49] Jackson J D 1975 *Classical Electrodynamics* 2nd edn, section 6.6 (New York: Wiley)
Jackson J D 1999 *Classical Electrodynamics* 3rd edn, section 6.4 (New York: Wiley)
[50] Thorne K S, Price R H and MacDonald D A 1986 *Black Holes: The Membrane Paradigm* (New Haven, CT: Yale University Press)
[51] Wald R M 1984 *General Relativity* (Chicago, IL: University of Chicago Press)
[52] Grischuk L P 1975 *Nuovo Cimento Lett.* **12** 60
Kolb E W and Turner M S 1990 *The Early Universe* (Redwood, CA: Addison-Wesley) and references therein
[53] Brill D R and Hartle J B 1964 Method of the self-consistent field in general relativity and its application to the gravitational geon *Phys. Rev.* **135** 271
[54] Isaacson R A 1968 Gravitational radiation in the limit of high frequency. I. The linear approximation and geometrical optics *Phys. Rev.* **166** 1263
Isaacson R A 1968 Gravitational radiation in the limit of high frequency. II. Nonlinear terms and the effective stress tensor *Phys. Rev.* **166** 1272
[55] See www.ligo.caltech.edu
[56] Freise A (Virgo Collaboration) 2004 Status of Virgo *Preprint* gr-qc/0406123
[57] Lück H *et al* 2000 *Proc. 3rd Edoardo Amaldi Conf.* (*Melville, NY 2000*) ed S Meshkov (*AIP Conf. Proc.* **523**) 119

[58] Ando M *et al* 2001 Stable operation of a 300-m laser interferometer with sufficient sensitivity to detect gravitational-wave events within our galaxy *Phys. Rev. Lett.* **86** 3950
[59] Kuroda K *et al* 2000 Large-scale cryogenic gravitational wave telescope *Int. J. Mod. Phys.* D **8** 557
[60] McCleland D *et al* 2000 *Proc. 3rd Edoardo Amaldi Conf.* (*Melville, NY 2000*) ed S Meshkov *AIP Conf. Proc.* **523**) 140
[61] Narayan R, Piran T and Shemi A 1991 Neutron star and black hole binaries in the galaxy *Astrophys. J.* **379** L17
[62] Phinney E S 1991 The rate of neutron star binary mergers in the universe—minimal predictions for gravity wave detectors *Astrophys. J.* **380** L17
[63] Kalogera V and Lorimer D R 2000 An upper limit on the coalescence rate of double neutron-star binaries in the galaxy *Astrophys. J.* **530** 890
[64] Burgay M *et al* 2003 An increased estimate of the merger rate of double neutron stars from observations of a highly relativistic system *Nature* **426** 531
[65] Bethe H and Brown G E 1998 Evolution of binary compact objects that merge *Astrophys. J.* **506** 780
[66] Portegies-Zwart S F and Yungel'son L R 1998 Formation and evolution of binary neutron stars *Astron. Astrophys.* **332** 173
[67] Portegies-Zwart S F and McMillan S L W 2000 Black hole mergers in the universe *Astrophys. J.* **528** L17
[68] Belczynski K, Kalogera V and Bulik T 2002 A comprehensive study of binary compact objects as gravitational wave sources: evolutionary channels, rates, and physical properties *Astrophys. J.* **572** 407
[69] Eardley D M 1983 *Gravitational Radiation* ed N Dereulle and T Piran (Amsterdam: North-Holland) p 257
[70] Chandrasekhar S 1969 *Ellipsoidal Figures of Equilibrium* (New Haven, CT: Yale University Press)
[71] Fryer C L, Holz D E and Hughes S A 2002 Gravitational wave emission from core collapse of massive stars *Astrophys. J.* **565** 430
[72] Fryer C L and Heger A 2000 Core-collapse simulations of rotating stars *Astrophys. J.* **541** 1033
[73] Dimmelmeier H, Font J A and Mueller E 2002 Relativistic simulations of rotational core collapse. II. Collapse dynamics and gravitational radiation *Astron. Astrophys.* **393** 523
[74] Fryer C L, Holz D E and Hughes S A 2004 Gravitational waves from stellar collapse: Correlations to explosion asymmetries *Astrophys. J.* **609** 288
[75] Jones D I 2002 Gravitational waves from rotating strained neutron stars *Class. Quantum Grav.* **19** 1255
[76] Owen B J 2002 *Matters of Gravity* vol 20, ed J Pullin, p 8 (available online at *Preprint* gr-qc/0209085)
[77] Cutler C 2002 Gravitational waves from neutron stars with large toroidal B fields *Phys. Rev.* D **66** 4025
[78] Bildsten L 1998 Gravitational radiation and rotation of accreting neutron stars *Astrophys. J.* **501** L89
[79] Ushomirsky G, Cutler C and Bildsten L 2000 Deformations of accreting neutron star crusts and gravitational wave emission *Mon. Not. R. Astron. Soc.* **319** 902
[80] Brady P R, Creighton T, Cutler C and Schutz B F 1999 Searching for periodic sources with LIGO *Phys. Rev.* D **57** 2101
[81] Brady P R and Creighton T 2000 Searching for periodic sources with LIGO. II. Hierarchical searches *Phys. Rev.* D **61** 082001
[82] Chakrabarty D *et al* 2003 Nuclear-powered millisecond pulsars and the maximum spin frequency of neutron stars *Nature* **424** 42
[83] Chakrabarty D 2004 Millisecond pulsars in x-ray binaries *Binary Radio Pulsars ASP Conf. Ser.* ed F A Rasio and I H Stairs (*Preprint* astro-ph/0408004)
[84] Backer D C, Kulkarni S R, Heiles C, Davis M M and Goss W M 1982 A millisecond pulsar *Nature* **300** 615
[85] Wagoner R V 1984 Gravitational radiation from accreting neutron stars *Astrophys. J.* **278** 345
[86] Andersson N and Kokkotas K D 2001 The r-mode instability in rotating neutron stars *Int. J. Mod. Phys.* D **10** 381
[87] Allen B 1996 *Proc. Les Houches School on Astrophysical Sources of Gravitational Waves* ed J-A Marck and J-P Lasota (Cambridge: Cambridge University Press) (*Preprint* gr-qc/9604033)
[88] Maggiore M 2000 Gravitational wave experiments and early universe cosmology *Phys. Rep.* **331** 283
[89] Buonanno A Gravitational waves from the early Universe (*TASI Lectures*) *Preprint* gr-qc/0303085

[90] Turner M S 1997 Detectability of inflation-produced gravitational waves *Phys. Rev.* D **55** 435
[91] Kosowsky A, Mack A and Kahniashvili T 2001 *Astrophysical Sources for Ground-based Gravitational Wave Detectors* (*AIP Conf. Proc.*) (Melville, NY: American Institute of Physics) vol 575, p 191
[92] Allen B and Romano J D 1999 Detecting a stochastic background of gravitational radiation: signal processing strategies and sensitivities *Phys. Rev.* D **59** 102001
[93] Tinto M and Dhurandar S 2004 Time-delay interferometry *Living Rev. Relat.* at press (*Preprint* r-qc/0409034)
[94] Hughes S A 2002 Untangling the merger history of massive black holes with LISA *Mon. Not. R. Astron. Soc.* **331** 805
[95] Sigurdsson S and Rees M J 1997 Capture of stellar mass compact objects by massive black holes in galactic cusps *Mon. Not. R. Astron. Soc.* **284** 318
[96] Sigurdsson S 1997 Estimating the detectable rate of capture of stellar mass black holes by massive central black holes in normal galaxies *Class. Quantum Grav.* **14** 1425
[97] Gair J *et al* 2004 Event rate estimates for LISA extreme mass ratio capture sources *Class. Quantum Grav.* **21** S1595
[98] Colbert E J M and Miller M C 2005 Observational evidence for intermediate-mass black holes in ultra-luminous x-ray sources (invited review talk) *Proc. 10th Marcel Grossmann Meeting on General Relativity* (*Rio de Janeiro, July 20–26, 2003*) ed M Novello, S Perez-Bergliaffa and R Ruffini (Singapore: World Scientific) (*Preprint* astro-ph/0402677)
[99] Miller M C 2005 Probing general relativity with mergers of supermassive and intermediate-mass black holes *Astrophys. J.* **618** 426
[100] Armstrong J W, Estabrook F B and Tinto M 1999 Time-delay interferometry for space-based gravitational wave searches *Astrophys. J.* **527** 814
[101] Hogan C J and Bender P L 2001 Estimating stochastic gravitational wave backgrounds with the Sagnac calibration *Phys. Rev.* D **64** 062002
[102] Sazhin M V 1978 Opportunities for detecting ultralong gravitational waves *Sov. Astron.* **22** 36
[103] Detweiler S 1979 Pulsar timing measurements and the search for gravitational waves *Astrophys. J.* **234** 1100
[104] McHugh M P, Zalamansky G, Vernotte F and Lantz E 1996 Pulsar timing and the upper limits on a gravitational wave background: a Bayesian approach *Phys. Rev.* D **54** 5993
[105] Jaffe A and Backer D C 2003 Gravitational waves probe the coalescence rate of massive black hole binaries *Astrophys. J.* **583** 616
[106] Turner M S 1999 Cosmology solved? Quite possibly! *Pub. Astron. Soc. Pacific* **111** 264
[107] Kamionkowski M and Kosowsky A 1999 The cosmic microwave background and particle physics *Annu. Rev. Nucl. Part. Sci.* **49** 77
[108] Kamionkowski M and Kosowsky A 1998 Detectability of inflationary gravitational waves with microwave background polarization *Phys. Rev.* D **57** 685
[109] Kamionkowski M, Kosowsky A and Stebbins A 1997 Statistics of cosmic microwave background polarization *Phys. Rev.* D **55** 7368
[110] Kesden M, Cooray A and Kamionkowski M 2002 Separation of gravitational-wave and cosmic-shear contributions to cosmic microwave background polarization *Phys. Rev. Lett.* **89** 011304
[111] Blanchet L 2002 Gravitational radiation from post-Newtonian sources and inspiralling compact binaries *Living Rev. Relat.* **5** 3
[112] Lehner L 2001 Numerical relativity: a review *Class. Quantum Grav.* **18** R25
[113] Rezzolla L 2002 'Gravitational waves from perturbed black holes and relativistic stars', *Lectures given at the Summer School on Astroparticle Physics and Cosmology (ICTP, July 2002)* (*Preprint* gr-qc/0302025)
[114] Teukolsky S A 1973 Perturbations of a rotating black hole. I. Fundamental equations for gravitational, electromagnetic, and neutrino-field perturbations *Astrophys. J.* **185** 635

Corrigendum added 6 October 2005

Equation (4.23) of this paper should read

$$h_{ij}^{\rm TT}(t,\mathbf{x}) = \frac{2}{r}\frac{d^2\mathcal{I}_{kl}(t-r)}{dt^2}\left[P_{ik}(\mathbf{n})P_{jl}(\mathbf{n}) - \frac{1}{2}P_{kl}(\mathbf{n})P_{ij}(\mathbf{n})\right] . \qquad (4.23)$$

The crucial second term in the square brackets was left out of the published version. We thank Ryan Lang at MIT for bringing this error to our attention.