1

The theoretical foundations of gravitational waves: overview and linearized theory

Simone Speziale[1] and Danièle Steer[2]

[1]Aix Marseille Univ., Univ. de Toulon, CNRS, CPT, UMR 7332, 13288 Marseille, France

[2]Laboratoire de Physique de l'ENS, Université Paris Cité, Ecole Normale Supérieure, Université PSL, Sorbonne Université, CNRS, 75005 Paris

GRAVITATIONAL WAVES,
coordinated by Nicola Tamanini (L2IT/CNRS). © ISTE Editions 2019.

1.1. Overview and characteristic scales

1.1.1. *Aims*

While indirect evidence for the existence of gravitational waves (GWs) has existed for many years (see e.g. (Will 2014) for a review), they were detected directly for the first time only in 2015 (Abbott *et al.* 2016), nearly 100 years after GWs were predicted to exist in general relativity (GR). These detections by the LVK collaboration, consisting of the network of LIGO, Virgo and more recently Kagra GW interferometers, are ongoing with new GWs signals being observed on a weekly basis (GraceDB n.d.). In the future more sensitive detectors on earth, together with ones working in different frequency bands such as the Laser Interferometer Space Antenna (LISA) (Amaro-Seoane *et al.* 2017) as well as Pulsar Timing Arrays (PTAs), will lead to new observations of the universe, potential new discoveries, and unprecendented tests of general relativity, cosmology and astrophysics.

The aim of these lectures is to present the basic introductory material required to understand GWs. We will address some of following questions:

- What are GWs? How do they emerge from GR? How does one deal with the symmetries (diffeomorphism invariance) of GR to fix gauges and coordinates, and what do they imply for the stress energy tensor of GWs?
- To what GW frequencies $f_{\rm GW}$ are current and future GW detectors sensitive? Why are those detectors designed to be sensitive to particular GW frequency ranges?
- For a source consisting of two *bound* compact binaries objects (such as black holes) of masses m_1 and m_2 at some distance R from an observer, what is the characteristic frequency, amplitude etc of the GWs emitted? Up to what distances R can such sources be detected?
- Using the quadrupole formula (which we derive) what is the waveform of the emitted GWs and how does it depend for example on the ellipticity of the bound orbit?
- What sources correspond to the GW events detected by LVK? Are there other possible GW sources? We give an example of compact binary sources on unbound orbits and discuss the GW memory effect.
- If we consider sources on cosmological distance scales, how are their amplitude, frequency e.t.c. affected by the cosmological expansion?

1.1.2. *On wave-like solutions and relativity*

Gravitational waves are naturally expected to exist in GR because it is a relativistic theory of gravity. To understand why, let us first take a step back to *non-relativistic* Newtonian gravity. When the famous apple drops on Newton's head, the mass distribution of the Earth changes: thus the gravitation field created by the mass distribution corresponding to the configuration with the "apple on the tree" is different to that created when the apple has fallen. In Newton's time, this variation, however negligible, was assumed to be the effect of some *instantaneous* "action at a distance". After the discovery that the *speed of light is finite*, and that all effects in our universe appear to follow this causal limitation, it seems natural to expect that also the variations of the gravitational field will *not* be felt instantaneously in the whole universe, but will rather be propagated at the speed of light — or less.[1] *The propagation of this perturbation of the gravitational field is intuitively what we call a gravitational wave.*

Conceptually, a gravitational wave is the same thing as a water wave or an electromagnetic (EM) wave. However, while those propagate a modification in the depth of water or the intensities of the electromagnetic field, a GW propagates a modification of the structure of spacetime itself. Just like a charged particle emits EM waves when moving along a closed trajectory, we expect that the Earth emits GWs when orbiting the sun, thus carrying away energy and making the orbit decay.

In practise, however, understanding GWs is very subtle for a number of reasons. First of all, the concept itself of propagation makes reference to a background spacetime, and in GR there is no fixed background structure. Splitting the dynamical spacetime into a reference background and a perturbation on top of it is a delicate process in which potential ambiguities have to be dealt with. In fact many decades passed before a consensus was reached, to the point that, famously, Einstein himself initially doubted the physical existence of GW, for reasons that we will briefly review and clarify below. Secondly, if we think of GWs as waves propagating in a medium, this medium is extraordinarily rigid: the waves go as fast as possible and have very *tiny* amplitudes. To give an idea, the power emitted by the Earth in the form of GW is around 200W a year! This rigidity has to do with the weakness of the gravitational coupling constant. One may think that gravity is strong when for e.g. trying to beat a

1. We will see that Einstein's relativistic theory of gravity, GR, predicts that these variations propagate, whatever their wavelength, at exactly the speed of light, and if some future experiment shows that they propagate at a lesser speed, than this would be an explicit violation of GR.

high jump record, or when skydiving, but this strength is ridiculously small compared to the much much stronger electro-magnetic force that dominates our daily life. These two points — background independence and weakness of the signal — are typical issues that one has to face when studying GWs.

Indeed only very massive and energetic objects can produce produce GWs of amplitudes that are actually detectable. Amongst the most heavy astrophysical objects known are black holes (BH), neutron stars (NS) and white dwarfs (WD). The GW sources detected to date by the LVK collaboration are all 'compact binary systems' made of a bound pair of BH and/or NS. As a result of the energy lost through GW emission, the two bodies making up the bound system approach closer to each other, eventually merging into one final object.In fact the GW signals detected by LVK correspond to the last moments in the life of these systems including their merger — they are known as '*compact binary coalescences*' (CBC's). For comparison with the earth-sun system mentioned above, the energy emitted in GWs by the very first detected GW event GW150914 (Abbott *et al.* 2016), which was due to the coalescence of two BHs of masses $m_1 \sim 36M_\odot$ and $m_1 \sim 29M_\odot$, was almost 10^{50}W.

Another important application of GWs is to cosmology. The gravitational interaction is so weak that the universe is almost completely transparent to a gravitational wave. As a consequence, we can potentially collect pristine information about any cosmological era through GWs, and in particular through the detection and characterization of a stochastic gravitational wave background. Sources relevant to cosmology include primordial GWs produced during inflation but there are also potential new sources to be discovered, such as primordial black holes, cosmic strings, and other exotic objects, see e.g. (Caprini and Figueroa 2018) for a review.

Our aim is not to provide an introduction to the broad set of fascinating GW sources, confirmed or hypothetical, nor to the many creative ideas to detect them that have been proposed, investigated and realized in practise; but only to provide an introduction to the field, and to that end, we decided to focus on the most common type of sources, and most common type of detectors: CBCs and laser interferometers. In the rest of this overview section we review the characteristic properties of GWs emitted by CBCs and the relevant frequency bands of laser interferometers, in particular explaining why LVK detectors are sensitive to the merger of stellar mass BH, whilst LISA for example to that of supermassive BHs. The rest of the chapter will present the theoretical derivation of GWs from GR.

1.1.3. *Detectors and GW frequencies*

The LVK interferometers and future LISA detector are essentially *Michelson-Morley interferometers*, designed to be as sensitive as possible to time-varying changes in the separation between two freely falling test-masses — mirrors in this case. The invariant distance between the test masses varies when a GW passes (see Section 2.1), leading to a change in the observed interference pattern in the detector.

- The LVK interformeters are on earth (in Livingston and Handford in the USA, in Pisa in Europe, and in Kamioka in Japan) and have a typical arm length $L \sim 3$km. They are sensitive to GWs with frequency of order

$$10\text{Hz} \lesssim f_{\text{GW}} \lesssim 5\text{kHz} \qquad (\text{LVK}). \qquad [1.1]$$

- The LISA interferometer (Amaro-Seoane *et al.* 2017) was adopted by ESA on the 25th january 2024, and should be operational in 2037. The distance between the spacecraft which make up arms of LISA is $L \sim 2.5 \cdot 10^6$km. LISA will be sensitive to GWs with frequencies in the range

$$10^{-4}\text{Hz} \lesssim f_{\text{GW}} \lesssim 1\text{Hz} \qquad (\text{LISA}). \qquad [1.2]$$

- There are plans to build new interferometers on earth beyond LVK. These include the Einstein Telescope in Europe (Sathyaprakash *et al.* 2012) and Cosmic Explorer in the USA (Reitze *et al.* 2019), both of which should have $L \sim 10$km, and

$$\text{few Hz} \lesssim f_{\text{GW}} \lesssim 10^4\text{Hz} \qquad (\text{ET, CE...}). \qquad [1.3]$$

- An alternative to interferometers are PTAs which search for GWs by exploiting the variation in distance $L \sim 10^{17}$km between the earth and a typical distant pulsar due to GWs. Pulsars emit EM pulses with extreme regularity Δt, typically of the order of milliseconds. If GWs are present, then as the EM pulses propagate from the pulsar to the earth, the observed Δt_{obs} will be modulated. PTA experiments searching for these modulations are sensitive to GWs in the frequency band of the inverse year,

$$10^{-7}\text{Hz} \lesssim f_{\text{GW}} \lesssim 10^{-9}\text{Hz} \qquad (\text{PTA}). \qquad [1.4]$$

  In 2023 different PTA experiments presented strong evidence for the existence of a stochastic GW background (Antoniadis *et al.* 2023 ; Agazie *et al.* 2023 ; Reardon *et al.* 2023 ; Xu *et al.* 2023).

- Prior to the success of interferometers, there was an effort pioneered by Weber in the 60's to use resonant bars to detect GWs, building material bars whose acoustic modes would resonate at a frequency as near as possible to that expected from the optimal sources (Weber 1960). These experiments would typically have a narrow-band sensitivity around 10^3Hz. In spite of constant experimental evolution throughout the 90's, no observation has occurred in this way.

Table 1.1 summarises the different characteristics of the existing experiments, and in particular the ratio of their characteristic size L to the GW wavelength $\lambda_{\rm GW} = 2\pi c/f_{\rm GW}$.

	Characteristic detector size (km)	GW frequency to which detector sensitive (Hz)	$f_{\rm GW}L$	L vs $\lambda_{\rm GW}$
LVK	~ 1	$10^1 - 10^4$	$f_{\rm GW}L \ll 1$	$L \ll \lambda_{\rm GW}$
LISA	$\sim 10^6$	$10^{-4} - 10^{-1}$	$f_{\rm GW}L \sim 1$	$L \sim \lambda_{\rm GW}$
PTA	$\sim 10^{17}$	$10^{-9} - 10^{-7}$	$f_{\rm GW}L \gg 1$	$L \gg \lambda_{\rm GW}$

Table 1.1: Characteristics of different GW detectors and the corresponding GW wavelength.

1.1.4. *Bound compact binary systems: orders of magnitude and characteristic scales*

LVK and LISA were conceived in order to be sensitive to the particular range of frequencies that are not only within experimental reach, but also that are likely to constitute a rich source according to the known astrophysical data. Amongst those GW sources are compact binary systems. We now focus on orders of magnitude and characteristic scales for such compact binary system, consisting of two masses $m_{1,2}$ at a distance R from the detectors, see figure 1.1. The expressions given here will be derived later in section 2.4. Furthermore, the expansion of the universe, neglected here, is considered in Section 2.5.

As shown in figure 1.2, as a consequence of GW emission, the two masses $m_{1,2}$ approach each other — the *inspiral phase* — until they merge — the *merger phase* — and form a single object. This object will keep radiating GWs, in the so-called *ringdown phase*, until it settles down to an equilibrium state (which for BHs is expected to be Kerr or Schwarzschild, according to theoretical and numerical evidence) after which no further emission occurs. The typical corresponding waveform, related to the GW amplitude, is shown in figure 1.2 as a function of time.

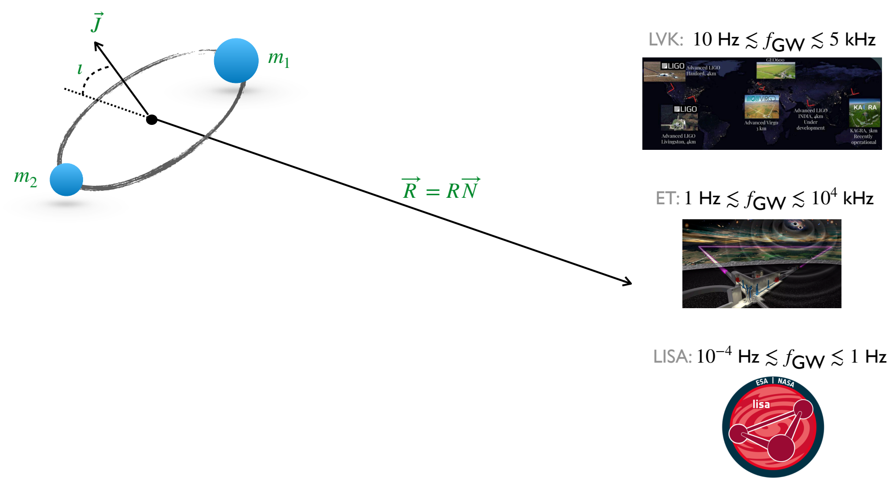

Figure 1.1: Sketch of a binary system of masses $m_{1,2}$ with conserved orbital angular momentum $\vec{J}$ and inclination ι, at a distance R from different detectors (LVK, ET and LISA). The approximate frequency bands of each detector are indicated.

- The *inspiral phase* can be understood with perturbation theory (the "post-Newtonian (PN) expansion" of the Einstein equations) presented below, more details in (Thorne 1980 ; Blanchet 2006 ; Poisson and Will 2014).
- The *merger phase* generally requires numerical relativity other other techniques such as effective one-body techniques, see e.g. (Deruelle and Uzan 2018) for an introduction.
- The *ringdown phase* can also be approached with perturbative methods, namely BH perturbation theory, see e.g. (Kokkotas and Schmidt 1999 ; Santoni 2024).

1.1.4.1. *The chirp signal*

During the inspiral phase the GW frequency increases with time according to the well-known *chirp* signal. Using the dominant quadrupolar mode for point masses m_1 and m_2 (with spins set to zero), and assuming circular orbits, it is given by

$$f_{\rm GW} = \frac{1}{\pi}\left(\frac{G\mathcal{M}}{c^3}\right)^{-5/8}\left(\frac{5}{256\tau}\right)^{3/8} \qquad [1.5]$$

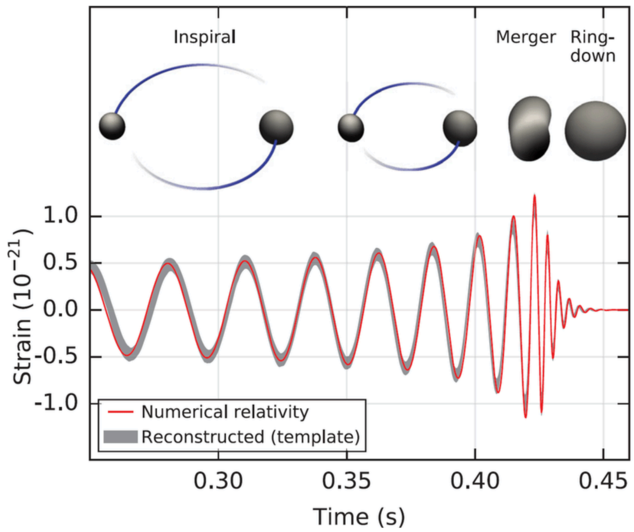

Figure 1.2: Figure showing the inspiral, merger and ringdown of a CBC, with corresponding GW waveform as a function of time, see the first GW detection by the LIGO-Virgo collaboration, GW150914 (Abbott *et al.* 2016) for which $\mathcal{M} \sim 30M_\odot$ and $R \sim 410$ Mpc. The small amplitude of the wave is due to the small coupling in Einstein's equations, see Eq. [1.17].

see Eq. [2.119], section 2.4.4. Here the *chirp mass* is

$$\mathcal{M} \equiv \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \tag{1.6}$$

and

$$\tau = t - t_c \tag{1.7}$$

is the time to coalescence, with t_c the coalescence time. Clearly Eq. [1.5] will break down before $\tau = 0$ (where formally $f_{\rm GW}$ diverges). We refer to this time $t_{\rm merger} < t_c$ as the merger time.

1.1.4.2. *Merger frequency*

Assuming that the two objects are Schwarzschild BHs, and that merger occurs at the innermost stable circular orbit (ISCO) namely a distance $a =$

$6GM/c^2$ with $m = m_1 + m_2$, then it follows from Keplers laws (see Sec. 2.4) together with Eq. [1.5] that

$$f_{\rm merger} = \frac{1}{6^{3/2}} \left(\frac{c^3}{Gm} \right). \qquad [1.8]$$

(Note given a length scale a and a mass m, $\sqrt{Gm/a^3}$ has dimensions of frequency. Setting $a = 6GM/c^2$ gives, modulo factors of 2π, Eq. [1.8].)

- For a binary neutron stars (**BNS**) system, with say $m_{1,2} \sim 1.4M_\odot$ then Eq. [1.8] gives

  $$f_{\rm merger} \simeq 1.5{\rm kHz} \qquad ({\rm BNS}) \qquad [1.9]$$

  This is upper part of the LVK frequency band, see figure 1.3.

- For a stellar mass binary black hole (**BBH**) system with for instance $m_{1,2} \sim 35M_\odot$,

  $$f_{\rm merger} \simeq 60{\rm Hz} \qquad (\text{stellar mass BBH}). \qquad [1.10]$$

  This is right in the frequency band of LVK, see figure 1.3.

- For a super massive black hole binary (**SMBHB**) system with for instance $m_{1,2} \sim 10^6 M_\odot$

  $$f_{\rm merger} \simeq 10^{-3}{\rm Hz} \qquad (\text{supermassive Binary BHs}) \qquad [1.11]$$

  which is in the frequency band of LISA, see figure 1.3.

- Notice that PTA frequencies do not correspond to the merger frequency of any know astrophysical system. Rather, they correspond to the inspiral phase of SMBHB at times much before merger, as can be seen from Eq. [1.8]. Hence these are on broad orbits, with periods of order years.

1.1.4.3. *Time to merger*

If GWs emitted during the inspiral enter the frequency band of a given detector at frequency $f_{\rm low}$, then it is straightforward to integrate Eq. [1.5] from $f_{\rm low}$ to $f_{\rm merger}$ to find the total duration of the GW source as will be detectable by the experiment. Assuming $f_{\rm merger} \gg f_{\rm low}$ for simplicity, one finds

$$T \sim 10^{-3} f_{\rm low}^{-8/3} \left(\frac{c^3}{G\mathcal{M}} \right)^{5/3} \qquad [1.12]$$

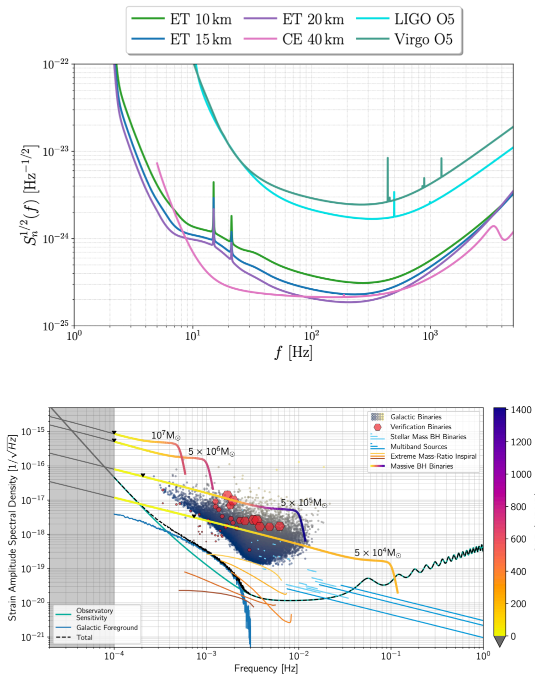

Figure 1.3: Upper panel: LVK, CE and ET sensitivities (y-axis) as a function of frequency, figure from (Maggiore *et al.* 2024). Lower panel in black dashed lines: LISA sensitivity including known astrophysical sources, figure from (Colpi *et al.* 2024).

- For BNS entering the LVK band with $f_{\rm low} \sim 20$Hz, this gives $T \sim 4$ minutes.

- For BNS entering the ET band with $f_{\rm low} \sim 1$Hz, then $T \sim 5$ days.
(This implies for example that (e.g. Doppler) effects of the rotation of the earth cannot be neglected when calculating the GW properties in more detail, see e.g. (Iacovelli *et al.* 2022) and references within. Also one might expect other GW signals to be produced in such a long period, overlapping with the BNS one. This makes data analysis more complex (Samajdar *et al.* 2021).)

- For stellar mass BHs,with say $m_{1,2} \sim 35 M_\odot$ entering the LVK band with $f_{\rm low} \sim 20$Hz, then $T \sim 0.1$ seconds.

- For SMBHB with $m_{1,2} \sim 10^6 M_\odot$ entering the LISA band with $f_{\rm low} \sim 10^{-4}$Hz, then $T \sim 1$ month.

1.1.4.4. *Amplitude and distance*

The dimensionless amplitude of the GW signal scales with distance R to the source and GW frequency $f_{\rm GW}$ as

$$h \sim \frac{4}{R}\left(\frac{G\mathcal{M}}{c^2}\right)^{5/3}\left(\frac{\pi f_{\rm GW}}{c}\right)^{2/3} \qquad [1.13]$$

As an example, consider say stellar mass BBH with $m_{1,2} \sim 35 M_\odot$ for which $f_{\rm merger} \sim 60$Hz. In order to generate (at merger) a signal with amplitude $h \sim 10^{-21}$, which is a couple of orders of magnitude higher than the LVK minimum strain according to figure 1.3 requires from Eq. [1.13] that

$$R \sim 400\,\mathrm{Mpc} \qquad [1.14]$$

which is of the order of galactic scales. For comparison, the observable universe has a scale of $c/H_0 \sim$Gpc, where H_0 is the Hubble constant.

Clearly from Eq. [1.13], given $\mathcal{M}$ and $f_{\rm GW}$, the more sensitive a detector, namely the smaller h can be detected, the further one can detect a given GW source. The "detection volume" of LVK has been steadily increasing with the different observing runs of LVK obviously leading to increasing numbers of detected GW events.

Notice that if such a GW signal is detected, then from the time dependence of the GW frequency one can directly obtain chirp mass Eq. [1.5]. With that, from the amplitude one can obtain the distance through Eq. [1.13]. Distance measurements are thus direct with GW observations from binaries, hence their name *standard sirens* (Schutz 1986 ; Holz and Hughes 2005). This should be contrasted with the case of EM observations (*standard candles*) for which the determination of the distance is particularly difficult.

1.1.4.5. *Distance between objects at merger*

When GWs are emitted with frequency $f_{\rm GW}$, the two bodies in the compact binary are separated by a characteristic scale

$$r \sim \left(\frac{Gm}{f_{\rm GW}^2}\right)^{1/3} \qquad [1.15]$$

(see also the discussion after Eq. [1.8]). Since, from Eq. [1.5], the GW frequency increases during inspiral, the distance r between the two bodies decreases. The minimum distance is at the merger frequency $f_{\rm merger}$. For example, for stellar mass BBH with $m_{1,2} \sim 35 M_{\odot}$ and $f_{\rm merger} \sim 60$Hz then from Eq. [1.15] $r \sim \mathcal{O}(100)$km.

A distance $r \sim \mathcal{O}(100)$km is tiny compared to the characteristic size of a star. WDs for example have a size $\sim 10^3$km and are some of the most dense objects in the universe (bar NSs). Main sequence stars are up to millions of km in size. Thus if GW signals are seen from objects which reach minimum approach distances $\sim \mathcal{O}(100)$km they cannot be stars, as they would already have collided. We must be dealing with BH (or possibly NS) for which the minimum distance will be determined by the Schwarzschild radius.

1.1.5. *Roadmap*

Having gone through the overview and discussed these orders of magnitude, the remainder of this chapter aims to derive formal results on GWs starting from Einsteins equations. Many introductions and reviews on GWs already exist, see for instance (Poisson and Will 2014 ; *E.Poisson, An advanced course on General Relativity* n.d. ; Maggiore 2007, 2018 ; Blanchet 2006 ; Andersson 2019 ; Deruelle and Uzan 2018) to mention a few. It is a rich and intricate topic, and each of these reviews tends to have a different angle on it, whose mutual compatibility may not always be clear to somebody entering the field. We have strived at presenting the material in a way that allows one to understand how the different approaches relate to one another. We have also strived to spend time on some of the subtleties and delicate conceptual aspects of GR and GWs which are often left to the side in gravitational wave introductions, such as gauge dependencies, asymptotic charges and memory effects, and which are becoming more and more relevant as theoretical research and experiments advance into more accurate comparisons. In these notes, a reader will therefore find discussions of questions of such as coordinate invariance, diffeomorphisms and gauge invariance, spin and helicity, the controversies about the stress energy tensor of GWs, GW memory effects, and Noether charges. It is our hope that although most of the more advanced material is not needed for a first introduction, its inclusion here will stimulate the reader, and provide a useful reference for delving further into the topic.

1.2. Einstein's equations: general covariance, Noether's theorem and gauge transformations

1.2.1. *Einstein's equations and general covariance*

Einstein's great discovery about gravitation was that it can be understood as the manifestation of the curvature of spacetime. In Wheeler's words, spacetime tells matter how to move, matter tells spacetime how to bend.[2] This discovery has changed profoundly our understanding of inertia. In Galilean relativity, an inertial observer is defined as one moving on a straight line at constant velocity. Special relativity introduces a non-trivial mixing of space and time, but leaves this notion unaffected: Inertial observers are still moving on a straight line, even though they are now related by Poincaré transformations as opposed to Galilean transformations, so to account for the experimental invariance of the speed of light. But in a curved spacetime, straight lines may no longer exist. The notion that encompasses them is the one of geodesics, which describe free-falling observers. You reading these notes at your desk are inertial in Newton's terms, but accelerated in Einstein's, since you are being held by the ground against Earth's gravitational attraction and not following a geodesic. The understanding offered by general relativity thus has the merit of not only explaining gravity, but also explaining the origin of inertia. Constant motion on a straight line is simply the flat-spacetime version of free falling.

Interpreting gravity as a dynamical spacetime metric has the consequence that the field equations of gravity and matter are *covariant under general coordinate transformation*, as we will review below, introducing a new paradigm that goes under the name of *principle of general covariance*. In a curved spacetime, there are no more preferred Cartesian coordinates, no more Poincaré transformations relating inertial observers, and physical concepts we are familiar with such as global time evolution and energy density lose their meaning. These aspects are often glossed over in lectures aiming at introducing gravitational waves, where one can blissfully rely on the background spacetime introduced by the weak field approximation and ignore most of them. However we believe they are important in order to better appreciate some of the properties of gravitational waves, provide an understanding that is more conceptual and less application-driven, and mostly because frankly who'd need yet another introduction to GWs if we didn't attempt something new? So we will briefly review them below, and use them as benchmark to discuss some conceptual aspects of gravitational waves in the rest of the notes. For instance, the lack of preferred clocks in a curved spacetime is relieved in the weak field approximation, where one can use the flat Minkowski background to introduce a class of Cartesian observers, and select their proper time as preferred time. But the

2. While pictorially charming, this statement is not exactly true: spacetime can be extraordinarily bent even in the absence of matter, as black hole solutions show.

lack of well-defined notion of energy density is a subtlety that persists also in the weak-field approximation, and has to be dealt with.

Let us start by recalling Einstein's equations

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \qquad [1.16]$$

where $G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}$ is the Einstein tensor with $R_{\mu\nu}$ the Ricci tensor and R the Ricci scalar, and $T_{\mu\nu}$ is the (symmetric) stress-energy tensor of matter. We use the definitions and conventions of (Poisson and Will 2014), in particular mostly plus convention for the spacetime metric. The constants G/c^4 and Λ are respectively the relativistic gravitational coupling constant and the cosmological constant. The first can be determined from local gravitational experiments to be

$$\frac{8\pi G}{c^4} \simeq 10^{-43}\,\mathtt{kg}^{-1}\mathtt{m}^{-1}\mathtt{s}^{-2}. \qquad [1.17]$$

This value is stupendously small, and it is the origin of the rigidity mentioned in the overview section. The smallness of this parameter has, on the other hand, a positive side: the gravitational force is so weak that many of the observed phenomena, and virtually all solar system experiments, can be studied using the *weak field approximation*, namely a perturbative expansion around the Minkowski metric. This is quite helpful because Einstein's equations are non-linear and it is not known how to solve them in general. Strong gravity effects occur only near very compact objects, and to study them one has to resort to numerical techniques, or be able to push the perturbative treatment to high orders.

The cosmological constant Λ can be determined from the observed acceleration of the expansion of the universe assuming homogeneity and isotropy on large scales, and turns then out to be $\Lambda \simeq 10^{-52}\,\mathtt{m}^{-2}$. This coupling constant can be interpreted as a sort of 'vacuum energy', often referred to as *dark energy*, whose value is $\rho_{DE} = \Lambda c^2/G \sim 10^{-28}\mathrm{kg/m}^3$. The presence of Λ affects the propagation of gravitational waves on cosmological distances, but it can be ignored for a first understanding of the perturbative treatment. We will set $\Lambda = 0$ for now, and restore it below in Section 2.5 when discussing cosmological effects.

Analysis of the 10 field equations in Eq. [1.16] shows (see subsection 1.3.6) that: four are redundant, because of the Bianchi identities; four are elliptic, hence describe gravitational degrees of freedom constrained by the sources; two are hyperbolic, hence contain independent degrees of freedom. This three-sided structure is a common feature to Maxwell and Yang-Mills theories, with the role of the Gauss constraint generating gauge transformation replaced by the so-called Hamiltonian and vector constraints generating diffeomorphisms, and it is our first indication that coordinate transformations are a gauge symmetry. A second indication comes from Noether's theorem, but before talking about it, let us review how coordinate transformations act.

Recall that a tensor is a quantity that transforms homogeneously under general coordinate transformations $x^\mu \to x'^\mu(x^\nu)$. For instance a scalar field transforms as $\phi'(x') = \phi(x)$, a vector field as (a contravariant tensor of order one) $v'^\mu(x') = \frac{\partial x'^\mu}{\partial x^\nu} v^\nu(x)$, and the metric as (a covariant tensor of order two)

$$g'_{\mu\nu}(x') = \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu} g_{\rho\sigma}(x). \qquad [1.18]$$

Since coordinate transformations are typically restricted to be differentiable, namely continuous and connected to the identity, they are also invertible, and correspond to mathematical transformations called diffeomorphisms. In this language, [1.18] is a diffeomorphism of the metric.

If the coordinate transformation is infinitesimal, we can write it as $x'^\mu = x^\mu - \xi^\mu(x)$, and compare the result of the transformation at the same point. This defines the infinitesimal generator $\delta_\xi \phi(x) := \phi'(x) - \phi(x)$, and

$$\delta_\xi g_{\mu\nu}(x) := g'_{\mu\nu}(x) - g_{\mu\nu}(x) \qquad [1.19]$$

for the metric.[3] Expanding both sides of [1.18] in Taylor series, we find that the first order effect of the transformation is the Lie derivative

$$\delta_\xi g_{\mu\nu}(x) = \pounds_\xi g_{\mu\nu} = \xi^\rho \partial_\rho g_{\mu\nu} + 2 g_{\rho(\mu}\partial_{\nu)}\xi^\rho = 2\text{NewA}_{(\mu}\xi_{\nu)}. \qquad [1.20]$$

3. The transformation law of a scalar is such that its value at one point P is the same after the diffeomorphism, since both x and x' identify the same point, just in different coordinates. This is sometimes misstated by saying that scalars are invariant under coordinate transformations, which is not true. A quantity is invariant under coordinate transformations if it satisfies the stronger property that its value does not depend on the coordinates used, which for the scalar field would be the equation $\phi(x') = \phi(x)$. This is not true in general, but only for isometries — more on this below. A typical example of coordinate invariance is the integral over the whole manifold of a scalar times the volume form, as experience from solving integrals via change of coordinates should teach.

The fact that the infinitesimal diffeomorphism of a tensor is the Lie derivative of that tensor is a general result, for instance for a scalar field we find $\delta_\xi\phi = \pounds_\xi\phi = \xi^\mu\partial_\mu\phi$, and so on. The second equality in [1.20] is on the other hand special to the metric tensor, and follows from the expression of the connection in terms of the metric.

Being written in terms of tensors, Einstein's equations are automatic covariant under general coordinate transformations. This is the principle of general covariance, that played a key role in guiding Einstein to formulate his theory. One immediate implication is that locally we can always find coordinates such that the metric takes the Minkowski expression at a point, which is one version of the principle of equivalence. A more subtle implication is that coordinate transformations must be *symmetries* of the theory, in other words a solution can be equivalently written in any coordinate system. To understand this point, let us consider the Lagrangian description of the dynamics. The field equations [1.16] are Euler-Lagrange equations of $\mathcal{L} = \mathcal{L}_{\text{EH}} + \mathcal{L}_{\text{M}}$, where

$$\mathcal{L}_{\text{EH}} = \frac{c^3}{16\pi G}(R - 2\Lambda)\sqrt{-g} \qquad [1.21]$$

is the Einstein-Hilbert Lagrangian density (with $g = \det(\mathrm{g}_{\mu\nu})$), and $\mathcal{L}_M$ the matter contribution, left arbitrary for the moment. Here the word density has a double meaning: in the physical sense, since $c\mathcal{L}$ has the dimensions $\mathtt{J\ m}^{-3}$ of an energy density, but also in the mathematical sense, since $\sqrt{-g}$ makes it transform not a scalar but a scalar density of weight 1. This means the following. The general formula for the variation of a determinant is $\delta g = g g^{\mu\nu}\delta g_{\mu\nu}$, which implies that $\pounds_\xi\sqrt{-g} = \frac{1}{2}\sqrt{-g}g^{\mu\nu}\pounds_\xi g_{\mu\nu} = \sqrt{-g}\mathrm{NewA}_\mu\xi^\mu$, whence $\pounds_\xi(\sqrt{-g}\phi) = \partial_\mu(\sqrt{-g}\xi^\mu\phi)$ for any scalar ϕ, and thus

$$\pounds_\xi\mathcal{L} = \partial_\mu(\xi^\mu\mathcal{L}). \qquad [1.22]$$

The density has introduced an extra contribution to the Lie derivative of a scalar, and the result is now a total derivative, hence a boundary term upon integration. Since boundary terms in the Lagrangian do not affect the field equations, the transformed solutions are still solutions: *diffeomorphisms are thus a symmetry of the system.*

This fact requires crucially that the metric be a dynamical field, and not a fixed background. To appreciate this point and the difference with non-general relativistic physics, let us consider the matter Lagrangian, which depends on both the metric g and the matter fields, which we denote them collectively as ψ. Applying the chain rule and using the short-hand notation E for the Euler-Lagrange equations, we find

$$\delta_\xi \mathcal{L}_{\rm M} = E_\psi\, \delta_\xi \psi + E_g\, \delta_\xi g = E_\psi\, \pounds_\xi \psi + E_g\, \pounds_\xi g = \pounds_\xi \mathcal{L}_{\rm M} = \partial_\mu(\xi^\mu \mathcal{L}_{\rm M}). \quad [1.23]$$

This means that diffeomorphisms are symmetries also of the matter sector. But this relies crucially on treating the metric as a dynamical variable! In non-general relativistic physics the metric is a non-dynamical, 'background' field, and $\delta g_{\mu\nu} = 0$. Accordingly, also $\delta_\xi g_{\mu\nu} = 0$, and therefore the second equality in the equation above breaks down and $\delta_\xi \mathcal{L}_{\rm M}$ is no longer a boundary term. Therefore a diffeomorphism does not map a solution into a new solution, and physics is not invariant under general coordinate transformations.

The discussion highlights why general covariance is often referred to as background independence, namely the absence of any fixed background metric in the theory, or as diffeomorphism invariance, since every physical observable should be independent of the coordinate used to describe it. General covariance, background independence, or diffeomorphism invariance, are thus different terms used to capture the same underlying property of general relativity.

The fact that every solution can be equivalently described in any coordinate system has a useful analogy with electromagnetism, where every solution can be described in any choice of gauge for the Maxwell potential. It is actually much more than an analogy, there is in fact a precise mathematical sense in which gauge transformations in Maxwell and Yang-Mills theories have the same property of coordinate transformations in general relativity, which we discuss next.

1.2.2. *Noether's theorem and diffeomorphisms as gauge symmetries*

We have seen that diffeomorphisms are symmetries of a general covariant Lagrangian. They are furthermore differentiable, namely continuous and connected to the identity. Noether's theorem proves that every differentiable symmetry defines a current which is conserved on solutions, or 'on-shell' in theoretical physics jargon. When this occurs, it is extremely useful to study the properties of the dynamics of the system, and to extract general physical predictions. However, there is an important difference between 'proper symmetries', which map solutions into distinguishable physical solutions, and 'gauge symmetries', which merely reflect a redundancy of the field equations and which introduce free functions of arbitrary time dependence irrelevant to the dynamics. Noether's theorem provides a simple test to distinguish the two cases: in the latter, the Noether current itself vanishes on shell, and not just its divergence.

The conserved current associated to diffeomorphisms of the Einstein-Hilbert Lagrangian [1.21] is given by

$$j^\mu_\xi = \frac{c^3}{8\pi G}\Big((G^\mu{}_\nu + \Lambda\delta^\mu_\nu)\xi^\nu - \mathrm{NewA}_\nu \mathrm{NewA}^{[\mu}\xi^{\nu]}\Big). \qquad [1.24]$$

One can immediately verify using the Bianchi identities that $\mathrm{NewA}_\mu j^\mu_\xi \hat{=} 0$, where $\hat{=}$ means on-shell. This current has a very peculiar property: the first term above vanishes on-shell, and the the second term is a total derivative. Therefore, the Noether current itself vanishes on-shell (in the absence of boundaries), and there are no conserved quantities. Having no conserved quantities is the hallmark of a gauge symmetry as opposed to a physical symmetry, hence the results provides a precise mathematical sense in which coordinate transformations in general relativity have the same status as gauge transformations in Maxwell and Yang-Mills theories.[4] For this reason, diffeomorphisms are also referred to as the gauge symmetry of general relativity, and fixing a coordinate choice as *fixing the gauge* in general relativity.

Having said so, there is a special situation that stands out: when the diffeomorphism corresponds to an *isometry*, namely a transformation that does not change the metric. This requires that [1.20] vanishes, and the corresponding equation $\mathrm{NewA}_{(\mu}\xi_{\nu)} = 0$ is called Killing equation, and ξ a Killing vector. A generic metric does not admit isometries: Isometries occur only for very special metrics. These special metrics are, however, important for physical applications,[5] hence Killing vectors play an important role. First of all, anyone who is familiar with the study of geodesics on spacetimes with isometries knows that there are conserved quantities associated with the Killing vectors, and which can be derived as Noether charges for the test particles' dynamics.

Second, the Noether current [1.24] evaluated on a spacetime with isometries gives rise to a useful conservation law analogue to the Gauss law in electromagnism, where the total charge in a region is equal to the flux of the electric field. To see this, we first observe that

$$\mathrm{NewA}_\nu \mathrm{NewA}^{[\mu}\xi^{\nu]} = \frac{1}{2}(R^\mu{}_\nu\xi^\nu - \Box\xi^\mu + \mathrm{NewA}^\mu \mathrm{NewA}_\nu \xi^\nu), \qquad [1.25]$$

4. A more rigorous approach is to look at the symplectic 2-form, and show that it is degenerate along gauge transformations and diffeomorphisms.

5. A cow is always a spherical object in the initial investigations of a theoretical physicists.

which follows from the definition of the Riemann tensor as the commutator of two covariant derivatives. If ξ^ν is a Killing vector, the last two terms vanish. Then integrating both sides of the equation over a 3d portion of space V delimited by two boundaries S_1 and S_2, and using Stokes' theorem, we find

$$Q_\xi[S] = \oint_S \mathrm{NewA}_\nu \mathrm{NewA}^{[\mu} \xi^{\nu]} dS_\mu, \quad [1.26]$$

$$Q_\xi[S_2] - Q_\xi[S_1] = -\frac{1}{2}\int_V R^\mu{}_\nu \xi^\nu dV_\mu \hat{=} -\frac{1}{2}\int_V \left(T^{\mu\nu}\xi_\nu - (\Lambda + \frac{T}{2})\xi^\mu \right) dV_\mu. \quad [1.27]$$

The Noether charge [1.26] obtained in this way is known as Komar charge. If the right-hand side of [1.27] vanishes, the Komar charge is conserved in the sense that it has the same value no matter which surface S is used, and its value changes only when the deformations of S include some source terms. If the right-hand side does not vanish, the Noether charge varies by an amount determined by the total quantity of energy-momentum in the enclosed region, see Fig 1.4. As an example, one can consider the Kerr solution, which possesses two Killing vectors corresponding to stationarity and axial symmetry. Evaluating [1.26] on an arbitrary 2-sphere S encompassing the singularity gives respectively the mass and angular momentum (up to numerical coefficients to be fixed), independently of the coordinate used and independently of deformations of S.

While Komar charges are limited to isometries, it is possible to generalize the construction of Noether charges and canonical generators to arbitrary spacetimes, at least in so far as they admit boundaries with non-trivial residual diffeomorphisms.[6] Let us mention three important examples. First, spacetimes that are asymptotically flat at spatial infinity. The residual diffeomorphisms compatible with the boundary conditions are the Poincaré transformations of the flat boundary metric. One can construct Noether charges and canonical generators for these boundary diffeomorphisms (see e.g. (Iyer and Wald 1994)), and the result coincides with the Arnowit-Deser-Misner (ADM) charges that were previously derived with canonical methods. Second, spacetimes that are asymptotically flat a null infinity. This case is particularly relevant to understand gravitational waves at the non-perturbative level. The residual transformations are a generalization of Poincaré transformations in which translations are angle-dependent, an infinite-dimensional extension known as Bondi-Van

6. There is also an extensive literature on constructing charges on arbitrary regions in arbitrary spacetimes, but the results there are much less clear.

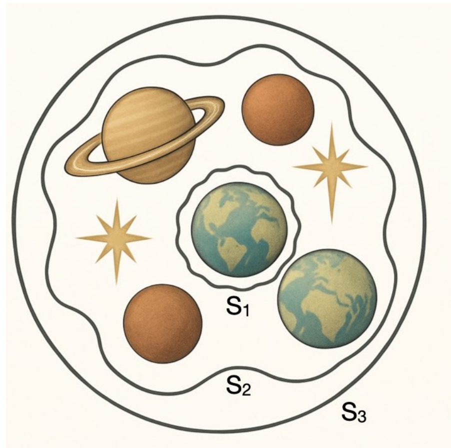

Figure 1.4: Conservation laws on stationary spacetimes. Evaluating the surface integral [1.26] on the innermost surface S_1 gives a quantity proportional to the energy-momentum of the planet encompassed. Evaluating it on S_2 gives the total energy-momentum of all stars and planets. The difference between the two surface integral is proportional to the energy-momentum of the region between them. Finally since there is no source outside S_2, integrating on S_2 or S_3 gives the same result.

der Burg-Metzner-Sachs (BMS) transformations.[7] Noether charges for the BMS symmetry were constructed in (Ashtekar and Streubel 1981 ; Dray and Streubel 1984 ; Wald and Zoupas 2000), and there is a large body of recent literature on the subject motivated by ongoing applications and developments. Among these, the application of the Noether approach to horizons and more generally null boundaries, see e.g. (Chandrasekaran *et al.* 2018 ; Ashtekar *et al.* 2022).

Finally, isometries alter the relation between diffeomorphisms and symmetries of the perturbative expansion, as we will see below, and this is relevant to the understanding of gravitational waves.

1.2.3. *Gauge and observers*

Even though every physical phenomenon can be described in any coordinate system, some choices can stand out because they simplify the description, or

7. Intuitively, the extension comes about because the induced metric on a null hypersurface is degenerate, hence any deformation along that direction leaves the system invariant.

because they are naturally associated with a class of observable of interest. For instance in flat spacetime, Cartesian coordinates make the Christoffel's symbols vanish, and can be associated to inertial observers. To take a simple example in curved spacetime, consider the Schwarzschild black hole solution. The most common coordinates used to describe this solution are the so-called 'static ones', which make time-independence of the solution manifest by making the metric independent of the time coordinate t, or in better terms, making one Killing vector be simply ∂_t. The t coordinate describes a family of non-inertial observers static at a fixed distance outside the black hole, whose time delays are related by $\sqrt{-g_{tt}}$. An alternative choice are Gullstrand-Painlevé coordinates, whose t describes the proper time of observers radially free-falling into the black hole. One could also consider the 'temporal gauge' defined by

$$g_{0\mu} = (-1, 0, 0, 0), \qquad [1.28]$$

in which the radially free-falling observes are all synchronized,[8] whence the alternative name of 'synchronous gauge'. The synchronization may look like a nice feature, but in these coordinates the spatial part of the Schwarzschild metric is explicitly dependent on the coordinate t! Therefore its staticity is hidden, and has to be verified by the existence of a time-translational Killing vector. In this example we can see the analogy between coordinate choices and gauge choices in electromagnetism very clearly. An electrostatic potential is more conveniently described in the Coulomb gauge because it makes the potential manifestly time-independent. But it can be described in any other gauge, and if we use the temporal gauge $A_0 = 0$, the potential acquires an inconvenient time dependence in the potential which is pure gauge and hides the staticity of the system.

The temporal gauge [1.28] can be chosen for *any* spacetime in a given coordinate chart. If it is done, it fixes completely the 4-dimensional diffeomorphism freedom in the chart. It is thus an example of complete gauge fixing,[9] and the resulting coordinates describe free-falling observers with synchronised clocks. It is a non-covariant gauge, because it relies on an initial choice of coordinate to be taken as the time. An example of covariant gauge is the harmonic gauge, which is defined requiring the coordinates to be such that the metric satisfies

$$\Box x^\mu = \Gamma^\mu_{\nu\rho} g^{\nu\rho} = \partial_\nu(\sqrt{-g} g^{\mu\nu}) = 0. \qquad [1.29]$$

8. This requires giving them non-zero energy, as opposed to the Gullstrand-Painlevé observers that have zero energy.

9. Complete here refers to the 4-dimensional picture. There remains the freedom of time-independent 3-dimensional diffeomorphisms, namely of choosing the coordinates on one – and one only – given hypersurface.

The covariance makes this a convenient choice in many dynamical situations. For instance, it is the gauge in which it is easiest to see that the initial value problem is well-posed (Choquet-Bruhat 1952), and the one in which it is easiest to study gravitational waves. Notice that [1.29] is not a complete gauge fixing of the 4-dimensional diffeomorphism freedom, as there are infinitely many solution of the wave equation, hence infinitely many choices of harmonic coordinates for a given metric. To obtain a complete gauge fixing one has to specify a unique set of harmonic coordinates with additional conditions.

As the Schwarzschild example shows, a complete gauge fixing can have the consequence of 'squeezing' physical information in field components one would not naturally expect. For that reason, partial gauges can also be very useful when extracting certain physical information. For instance, one version of the equivalence principle states that purely local experiments can not distinguish the presence of gravity. In the formalism of general relativity, this is embodied in the fact that at any given point in spacetime it is possible to find coordinates so that the metric is flat and its first derivative vanish. A coordinate system that achieves this is called a *local inertial frame*. Only an experiment that can probe second-order variations in the metric would be able to see the effect of gravity in a coordinate-independent way, and these variations are the tidal forces that show up in the geodesic deviation equation. An example of local inertial frame is provided by Riemann normal coordinates, which are constructed around a given point so that the Taylor expansion of the metric components around that point taken as the origin gives

$$g_{\mu\nu} = \eta_{\mu\nu} + c_{\mu\nu} R_{\mu\nu\rho\sigma} x^\rho x^\sigma + O(x^3), \tag{1.30}$$

where $c_{\mu\nu} = -1/3 \ \forall \mu, \nu$.[10] This gauge fixing specifies coordinates only in the neighbourhood of a point and not in a full coordinate chart, but it is still perfectly sufficient to describe the physics of a local inertial frame around that point. It is also possible to find coordinates such that the Christoffel's symbols vanish everywhere along a chosen time-like geodesic. Such coordinates are known as Fermi normal coordinates, and describe a free-falling local inertial frame. The metric takes the same form [1.30], but with $c_{00} = -1, c_{0a} = -2/3, c_{ab} = -1/3$. The difference between temporal gauge and Fermi normal coordinates is that in the first case every observer at constant spatial coordinates is free falling, whereas in the second case only the observer going through the origin along the ∂_t geodesic is free falling.

10. This formula is only valid in certain coordinate choices, hence the non-covariant notation with the $\mu\nu$ indices repeated but not summed over.

1.3. Perturbative treatment of Einstein's equations

1.3.1. *The idea: general background space-time*

Perturbation theory can be set up choosing a background metric $\bar{g}_{\mu\nu}$ and writing

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}, \tag{1.31}$$

with the assumption that $|h_{\mu\nu}| \ll 1$ in some chosen coordinate system, so that it can be treated as a perturbation. One can then systematically Taylor-expand all metric functionals around the background, starting with the inverse metric $g^{\mu\nu} = \bar{g}^{\mu\nu} - h^{\mu\nu} + O(h^2)$ and Levi-Civita connection

$$\Gamma^{\mu}_{\nu\rho} = \bar{\Gamma}^{\mu}_{\nu\rho} + \Gamma^{(1)}{}^{\mu}_{\nu\rho} + O(h^2), \qquad \Gamma^{(1)}{}^{\mu}_{\nu\rho} = \frac{1}{2}\bar{g}^{\mu\sigma}(2\partial_{(\nu}g_{\rho)\sigma} - \partial_\sigma g_{\nu\rho}), \tag{1.32}$$

and attempt to solve the field equations order by order:

$$\bar{G}_{\mu\nu} + G^{(1)}_{\mu\nu} + G^{(2)}_{\mu\nu} + \ldots = \frac{8\pi G}{c^4}(\bar{T}_{\mu\nu} + T^{(1)}_{\mu\nu} + T^{(2)}_{\mu\nu}) + \ldots \tag{1.33}$$

Here $\bar{G}_{\mu\nu} = G_{\mu\nu}(\bar{g})$, $G^{(1)}_{\mu\nu} = G_{\mu\nu}(\bar{g}; h)$, and so on. Explicit expressions will not be needed here, but are given in Appendix 1.4 for completeness. The lowest order of the procedure is straightforward, one just has to be consistent with the treatment of the matter energy-momentum tensor. For a perturbative expansion around a spacetime which is a vacuum solution, $\bar{G}_{\mu\nu} = 0$ and we treat matter as a first order perturbation. Then the first order equation to solve is

$$G^{(1)}_{\mu\nu} = \frac{8\pi G}{c^4}\bar{T}_{\mu\nu}. \tag{1.34}$$

Solving this equation determines $h_{\mu\nu}$ in terms of its independent degrees of freedom, the background solution and the matter content. To go to second order, we add a second perturbation writing

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu} + h^{(2)}_{\mu\nu}. \tag{1.35}$$

Then $G^{(2)}_{\mu\nu} = G^{(2)}_{\mu\nu}(h) + G^{(1)}_{\mu\nu}(h^{(2)})$ has two contributions, and we solve for $h^{(2)}$ using

$$G^{(1)}_{\mu\nu}(h^{(2)}) = \frac{8\pi G}{c^4}\left(T^{(1)}_{\mu\nu} + t^{\text{G}}_{\mu\nu}\right), \qquad t^{\text{G}}_{\mu\nu} := -\frac{c^4}{8\pi G}G^{(2)}_{\mu\nu}(h). \tag{1.36}$$

Notice that the first order solution feeds back as a source for the second order solution, a standard procedure from perturbatively solving non-linear equations. The same procedure applies also to the case when matter fields contribute to the background as well. If the background chosen is flat, then the perturbative approximation describes a weak-field expansion, the order is controlled by the Newton's constant as $h^{(\mathrm{n})} \sim G^n$, and it is also referred to as *post-Minkowskian* (PM) expansion. In many cases the equations are still to hard to solve at each order, and one needs to look for additional approximations. A very common one is the non-relativistic approximation, also known as *post-Newtonian* (PN) expansion, in which one starts from a source that moves at small velocity with respect to the background flat metric, namely with $v \ll c$, and expands in powers of v^2/c^2. Alternative approximation schemes include the Bondi asymptotic expansion which is perturbative in the inverse distance from the source but valid at all orders in G and v^2/c^2 (Bondi *et al.* 1962 ; Sachs 1962); the extremal mass ratio inspiral (EMRI) and self-force expansion for a two-body system where the small parameter is the mass-ratio (Barack and Pound 2019); the effective one-body approach based on resummed PN results (Buonanno and Damour 1999), and more recently the effective field theory approach based on tools from quantum field theory (Goldberger and Rothstein 2006).

Apart from technical difficulties, there is also a conceptual difficulty: *it is not possible to simply truncate the perturbative expansion to a fixed order (except the lowest) and preserve covariance.* Expanding both sides of [1.20] with [1.31], we obtain

$$\delta_\xi \bar{g}_{\mu\nu} + \delta_\xi h_{\mu\nu} = \pounds_\xi \bar{g}_{\mu\nu} + \pounds_\xi h_{\mu\nu}. \tag{1.37}$$

We assume that the background metric is fixed once and for all and unaffected by diffeomorphisms, namely $\delta\bar{g}_{\mu\nu} = 0$ and $\delta_\xi \bar{g}_{\mu\nu} = 0$,[11] and define the effect of the linearized diffeomorphism on the perturbation to be

$$\delta_\xi h_{\mu\nu} = \pounds_\xi \bar{g}_{\mu\nu} + \pounds_\xi h_{\mu\nu}. \tag{1.38}$$

The key point is that the right-hand side contains terms of different order in h. This is why different orders of the perturbative expansion must be included in order for diffeomorphisms to be a symmetry. It is instructive to prove this

11. It is also possible to interpret [1.37] as $\delta_\xi \bar{g}_{\mu\nu} = \pounds_\xi \bar{g}_{\mu\nu}$ and $\delta_\xi h_{\mu\nu} = \pounds_\xi h_{\mu\nu}$. The proof that this is a symmetry is identical. However this definition of perturbed transformations is less interesting physically, because it is more natural to compare perturbations when they are defined with respect to the same background.

in detail, because it highlights the special features that occur for the quadratic Lagrangian, namely the free theory.

To study the symmetries of the perturbative expansion, we write the perturbed Lagrangian as follows

$$\mathcal{L}(\bar{g}+h) = \bar{\mathcal{L}} + \bar{\mathcal{L}}^{(1)\mu\nu} h_{\mu\nu} + \frac{1}{2} h_{\mu\nu} \bar{\mathcal{L}}^{(2)\mu\nu\rho\sigma} h_{\rho\sigma} + \ldots, \qquad [1.39]$$

where barred quantities only depend on $\bar{g}$ and not on the perturbation. Using [1.38], the infinitesimal variation gives

$$\begin{aligned} \delta_\xi \mathcal{L} = {} & \bar{\mathcal{L}}^{(1)\mu\nu} \pounds_\xi \bar{g}_{\mu\nu} + \bar{\mathcal{L}}^{(1)\mu\nu} \pounds_\xi h_{\mu\nu} + h_{\mu\nu} \bar{\mathcal{L}}^{(2)\mu\nu\rho\sigma} \pounds_\xi \bar{g}_{\rho\sigma} \\ & + h_{\mu\nu} \bar{\mathcal{L}}^{(2)\mu\nu\rho\sigma} \pounds_\xi h_{\rho\sigma} + \frac{1}{2} h_{\mu\nu} h_{\rho\sigma} \bar{\mathcal{L}}^{(3)\mu\nu\rho\sigma\tau\lambda} \pounds_\xi \bar{g}_{\tau\lambda} + \ldots . \end{aligned} \qquad [1.40]$$

All terms can be collected into Lie derivatives:

$$\bar{\mathcal{L}}^{(1)\mu\nu} \pounds_\xi \bar{g}_{\mu\nu} = \pounds_\xi \bar{\mathcal{L}}, \qquad [1.41]$$

$$\bar{\mathcal{L}}^{(1)\mu\nu} \pounds_\xi h_{\mu\nu} + h_{\mu\nu} \bar{\mathcal{L}}^{(2)\mu\nu\rho\sigma} \pounds_\xi \bar{g}_{\rho\sigma} = \pounds_\xi (\bar{\mathcal{L}}^{(1)\mu\nu} h_{\mu\nu}), \qquad [1.42]$$

$$h_{\mu\nu} \bar{\mathcal{L}}^{(2)\mu\nu\rho\sigma} \pounds_\xi h_{\rho\sigma} + \frac{1}{2} h_{\mu\nu} h_{\rho\sigma} \bar{\mathcal{L}}^{(3)\mu\nu\rho\sigma\tau\lambda} \pounds_\xi \bar{g}_{\tau\lambda} = \frac{1}{2} \pounds_\xi (\bar{\mathcal{L}}^{(2)\mu\nu\rho\sigma} h_{\mu\nu} h_{\rho\sigma}), \qquad [1.43]$$

and so on. Each Lie derivative gives a boundary term through [1.22], hence [1.38] is indeed a symmetry of the full Lagrangian. However, in every case except the lowest one, getting a boundary term for $\bar{\mathcal{L}}^{(\mathrm{n})}$ requires both $\bar{\mathcal{L}}^{(\mathrm{n})}$ and $\bar{\mathcal{L}}^{(\mathrm{n}+1)}$, hence if we truncate the series at a fixed order, we lose covariance.

There are two special features that occur at the quadratic order. First, if we take the background to be a solution, the first term in [1.42] vanishes, hence

$$\delta_\xi h_{\mu\nu} = \pounds_\xi \bar{g}_{\mu\nu} \qquad [1.44]$$

is a symmetry of the quadratic Lagrangian. Second, if the background has isometries, then the term proportional to $\bar{\mathcal{L}}^{(3)}$ in [1.43] drops out, and then

$$\delta_\xi h_{\mu\nu} = \pounds_\xi h_{\mu\nu} \qquad [1.45]$$

is a symmetry of the quadratic Lagrangian. This shows that isometries play a special role in perturbation theory. For a generic on-shell background the symmetry of the quadratic Lagrangian is [1.44]. But if the background has isometries, we have two different realization of the diffeomorphism symmetry in the quadratic Lagrangian: [1.44] for a generic diffeomorphism, and [1.45] for a Killing vector. To make this consistent with the perturbative expansion, we treat a Killing ξ as zero-th order, and a non-Killing ξ as first order.

We stress that [1.44] and [1.45] are symmetries only for the quadratic Lagrangian. From the cubic Lagrangian onwards, there is no symmetry at fixed order in h. The only symmetry is the combined [1.38] and requires two different perturbative orders of the Lagrangian. This fact has immediate consequences for the expanded Einstein tensor [1.33]. Since the leading order $G^{(1)}$ comes from the quadratic Lagrangian, it is invariant under [1.44]. But $G^{(2)}_{\mu\nu}$ which is derived from the cubic Lagrangian is not invariant. *As a consequence, the quantity $t_{\mu\nu}$ appearing in* [1.36] *is not gauge invariant.* This means that by itself, it cannot be taken as any meaningful notion of gravitational energy. Indeed, one can see from its explicit form that is it always possible to set it to zero locally with a coordinate transformation.

Even though both symmetries [1.44] and [1.45] descend from the same diffeomorphism invariance of the theory, which is a gauge symmetry, they have a different status at the perturbative level. Applying Noether's theorem to the quadratic Lagrangian, we find that the generic diffeomorphisms still have vanishing conserved current, and therefore maintain their status of gauge symmetries. However the diffeomorphisms corresponding to isometries of the background have on-vanishing Noether charges, indeed just like a standard theory on flat spacetime. A related difference is that the field equations are *invariant* under [1.44], hence there are linear dependencies in the equations and some field components are left undetermined, and *covariant* under [1.45], and there are no constraints associated with them.

1.3.2. *Weak-field approximation*

Spacetime in the solar system is nearly flat, hence it can be described using [1.31] with the background being the Minkowski spacetime,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}. \qquad [1.46]$$

We refer to this particular case of perturbative expansion as the weak-field approximation. We further use Cartesian coordinates, so all covariant derivatives become partial derivatives. In this case the linearized Einstein equations [1.34] take the simple form

$$\Box h_{\mu\nu} - 2\partial_{(\mu}\partial_\rho h^\rho{}_{\nu)} + \partial_\mu\partial_\nu h + \eta_{\mu\nu}(\partial_\rho\partial_\sigma h^{\rho\sigma} - \Box h) = -\frac{16\pi G}{c^4}T_{\mu\nu}, \quad [1.47]$$

where $\Box$ is the d'Alembertian in flat spacetime.

The Minkowski background has isometries. These are the ten Poincaré transformations, which we parametrize with vector fields

$$\xi^\mu = a^\mu{}_\nu x^\nu + b^\mu, \tag{1.48}$$

where $a^\mu{}_\nu$ and b^μ are constants, and $a_{(\mu\nu)} = 0$. According to the general discussion of the previous section, we expect that diffeomorphism invariance induces two different types of symmetries for [1.47]. For diffeomorphisms corresponding to isometries, [1.45] gives

$$\delta_\xi h_{\mu\nu} = \pounds_\xi h_{\mu\nu} = (a^\rho{}_\sigma x^\sigma + b^\rho)\partial_\rho h_{\mu\nu} + 2a^\rho{}_{(\mu} h_{\nu)\rho}. \tag{1.49}$$

It is easy to check that this is a symmetry that changes the equations covariantly, provided $T_{\mu\nu}$ is also transformed. For generic diffeomorphisms that are not isometries, [1.44] gives

$$\delta_\xi h_{\mu\nu} = \pounds_\xi \eta_{\mu\nu} = 2\partial_{(\mu}\xi_{\nu)}. \tag{1.50}$$

It is also easy to check that this is a symmetry that leaves invariant the equations. These statements are consistent with the right-hand side. Assuming the background η to be a solution requires that we are treating the matter fields as first order in perturbation theory. Then it transforms as $\pounds_\xi T_{\mu\nu}$ for a Killing ξ, whereas its transformation under generic diffeomorphisms is second order, hence it does not affect the linearized equations.

The presence of the d'Alembertian in the linearized field equations suggests that wave solutions are indeed possible. However, there is an intricate tensorial structure that needs to be dealt with. Before doing so, let us discuss what it means for a wave to be 'tensorial'. The waves that we are most familiar with, such as water waves or sound waves, are *scalar* waves, namely the quantity whose perturbation propagates following the wave equation is a scalar function, such as the height of the water or the pressure of the air. Electromagnetic waves on the other hand propagate changes in the electric and magnetic field which are described by vectors, and are thus 'vectorial' waves. The difference between scalar, vectorial and tensorial waves can be described in terms of *spin*. The reader already familiar with these concepts, or the reader interested in looking first at the explicit solutions without these details, can skip the next subsection.

1.3.3. *Spin and helicity*

Tensors in Minkowski spacetime belong to finite-dimensional representations of the Lorentz group. This property can be used to decompose each tensor into irreducible parts, namely parts that are not mixed with one another by a Lorentz transformation. The irreducible parts can be labelled by a pair of half-integers (j_1, j_2), and contain $(2j_1+1)(2j_2+1)$ components. For instance, a 4-vector v^μ transforms under the irreducible Lorentz representation $(\mathbf{\frac{1}{2}}, \mathbf{\frac{1}{2}})$ with 4 components, and a symmetric tensor $h_{\mu\nu}$ transforms under the reducible Lorentz representation $(\mathbf{1}, \mathbf{1}) \oplus (\mathbf{0}, \mathbf{0})$ with $9+1=10$ components.

The representations can be further subdivided if we pick a time direction τ^μ, and restrict attention to the rotation subgroup of the Lorentz group that preserves it. The subsets of the tensor which are irreducible with respect to the rotation subgroup are called *spin* representations, and their allowed values given by the Clebsch-Gordan addition rule $(j_1+j_2, \dots, |j_1-j_2|)$. For instance a vector contains the two spin representations $\mathbf{1} \oplus \mathbf{0}$, and a symmetric tensor the four spin representations $\mathbf{2} \oplus \mathbf{1} \oplus \mathbf{0} \oplus \mathbf{0}$. To be more explicit let us take $\tau^\mu = (1, 0, 0, 0)$. The rotation subgroup that preserves it has the form

$$R^\mu{}_\nu = \begin{pmatrix} 1 & \vec{0} \\ \vec{0} & R^a{}_b \end{pmatrix}, \qquad R^a{}_b \in \mathrm{SO}(3). \qquad [1.51]$$

It is then immediate to see that the spin-1 and spin-0 representation of the 4-vector are the spatial vector v^a and the spatial scalar v^0 respectively. Similarly for a 1-form v_μ, the spatial and time components v_a and v_0. For a symmetric tensor, the four spin representations are

$$h_{\langle ab \rangle} = h_{ab} - \frac{1}{3}\delta_{ab} h^c_c, \quad h_{0a}, \quad h_{00}, \quad h_s := h^c_c. \qquad [1.52]$$

We see that the spin 0 representation has one component, the spin 1 has three, and the spin 2 has five. These different components can be classified choosing a reference spatial axis and looking at the eigenmodes of the rotation generator along that axis. To fix ideas let us choose the z axis. The rotation matrix that preserves it is

$$R^a{}_b = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad [1.53]$$

Inserting this in [1.51] and acting on a 4-vector we obtain $v'^{\mu} = R^{\mu}{}_{\nu} v^{\nu}$, where

$$v'^0 = v^0, \qquad v'^x = \cos\theta v^x - \sin\theta v^y, \\ v'^y = \sin\theta v^x + \cos\theta v^y, \qquad v'^z = v^z, \tag{1.54}$$

and

$$v^{\pm} := v^x \pm i v^y, \qquad v'^{\pm} = e^{\pm i\theta} v^{\pm}. \tag{1.55}$$

Introducing a canonical basis $e^{\mu}_{I} = \delta^{\mu}_{I}$, the eigenvectors are

$$\epsilon^{\mu}_{0} = e^{\mu}_{0}, \qquad \epsilon^{\mu}_{\rm L} = e^{\mu}_{3}, \qquad \epsilon^{\mu}_{\pm} = \frac{1}{\sqrt{2}}(e^{\mu}_{1} \mp i e^{\mu}_{2}), \tag{1.56}$$

with eigenvalues

$$R^{\mu}{}_{\nu}\epsilon^{\nu}_{0} = \epsilon^{\mu}_{0}, \qquad R^{\mu}{}_{\nu}\epsilon^{\nu}_{\rm L} = \epsilon^{\mu}_{\rm L}, \qquad R^{\mu}{}_{\nu}\epsilon^{\nu}_{\pm} = e^{\pm i\theta}\epsilon^{\mu}_{\pm}. \tag{1.57}$$

The integers 0 and ± 1 that appear in front of the rotation angle are the projection of the spin along the z axis.

Acting on a symmetric tensor $h_{\mu\nu}$ we obtain $h'_{\mu\nu} = R^{\rho}{}_{\mu} R^{\sigma}{}_{\nu} h_{\rho\sigma}$, where

$$h'_{00} = h_{00}, \qquad h'_{0z} = h_{0z}, \qquad h'_{s} = h_{s}, \qquad h'_{zz} = h_{zz} \tag{1.58a}$$

$$h'_{0x} = \cos\theta h_{0x} + \sin\theta h_{0y}, \qquad h'_{0y} = -\sin\theta h_{0x} + \cos\theta h_{0y}, \tag{1.58b}$$

$$h'_{xz} = \cos\theta h_{xz} + \sin\theta h_{yz}, \qquad h'_{yz} = -\sin\theta h_{xz} + \cos\theta h_{yz}, \tag{1.58c}$$

$$h'_{+} = \cos 2\theta h_{+} + \sin 2\theta h_{\times}, \qquad h'_{\times} = -\sin 2\theta h_{+} + \cos 2\theta h_{\times}, \tag{1.58d}$$

and we defined

$$h_{+} := \frac{1}{2}(h_{xx} - h_{yy}), \qquad h_{\times} := h_{xy}, \qquad h_{s} := h_{xx} + h_{yy} + h_{zz}. \tag{1.59}$$

To write the eigenvectors, we introduce a canonical basis in the space of symmetric 4×4 matrices,

$$e^{IJ}_{\mu\nu} = \begin{cases} \delta^{I}_{\mu}\delta^{J}_{\nu} & I = J \\ \frac{1}{\sqrt{2}}(\delta^{I}_{\mu}\delta^{J}_{\nu} + \delta^{J}_{\mu}\delta^{I}_{\nu}) & I \neq J \end{cases} \tag{1.60}$$

Then

$$\epsilon^{0}_{\mu\nu} = e^{00}_{\mu\nu}, \qquad \epsilon^{\mathrm{s}}_{\mu\nu} = \frac{1}{\sqrt{3}}(e^{11}_{\mu\nu} + e^{22}_{\mu\nu} + e^{33}_{\mu\nu}),$$

$$w^{\pm 1}_{\mu\nu} = \frac{1}{\sqrt{2}}(e^{01}_{\mu\nu} \pm i e^{02}_{\mu\nu}), \qquad w^{\mathrm{L}}_{\mu\nu} = e^{03}_{\mu\nu}, \tag{1.61a}$$

$$\epsilon^{\pm}_{\mu\nu} = \frac{1}{2}(e^{11}_{\mu\nu} - e^{22}_{\mu\nu}) \pm \frac{i}{\sqrt{2}} e^{12}_{\mu\nu}, \qquad \epsilon^{\mathrm{L}\pm}_{\mu\nu} = \frac{1}{\sqrt{2}}(e^{13}_{\mu\nu} \pm i e^{23}_{\mu\nu}),$$

$$\epsilon^{\mathrm{LL}}_{\mu\nu} = \frac{1}{\sqrt{6}}(e^{11}_{\mu\nu} + e^{22}_{\mu\nu} - 2e^{33}_{\mu\nu}) \tag{1.61b}$$

with eigenvalues

$$R^{\rho}{}_{\mu} R^{\rho}{}_{\nu}\, \epsilon^{\mathrm{i}}_{\rho\sigma} = \epsilon^{\mathrm{i}}_{\mu\nu}, \quad \text{for } i = 0, s, \mathrm{LL} \qquad R^{\rho}{}_{\mu} R^{\rho}{}_{\nu}\, w^{\mathrm{L}}_{\rho\sigma} = w^{\mathrm{L}}_{\mu\nu}, \tag{1.62}$$

$$R^{\rho}{}_{\mu} R^{\rho}{}_{\nu}\, \epsilon^{\pm}_{\rho\sigma} = e^{\pm 2i\theta} \epsilon^{\pm}_{\mu\nu}, \quad R^{\rho}{}_{\mu} R^{\rho}{}_{\nu}\, w^{\pm}_{\rho\sigma} = e^{\pm i\theta} w^{\pm}_{\mu\nu}, \quad R^{\rho}{}_{\mu} R^{\rho}{}_{\nu}\, \epsilon^{\mathrm{L}\pm}_{\rho\sigma} = \epsilon^{\pm i\theta} w^{\mathrm{L}\pm}_{\mu\nu}. \tag{1.63}$$

The spin-2 components $e^{\pm}$, $e^{\mathrm{L}\pm}$ and e^{LL} carry respectively the $\pm 2, \pm 1$ and 0 modes of the z-projection. The spin-2 components $w^{\pm}$ and w^{L} carry ± 1 and 0 modes as in [1.57], and $e^{0,\mathrm{s}}$ are the remaining two spin-0 modes from [1.52].

The discussion so far concerned global vectors and tensors. In the case of electromagnetism and gravity linearized around Minkowski we work with vector and tensor *fields*, $A_\mu(x)$ and $h_{\mu\nu}(x)$. The mode decompositions described above are then applied locally in Fourier space. Each Fourier mode $\tilde{A}_\mu(k)$ and $\tilde{h}_{\mu\nu}(k)$ is characterized by a wave vector k, and we can use its spatial direction to classify the spin components. With some abuse of language, we will often refer to the wave vector as the momentum of the wave. The eigenvalues of the projection of the spin along the momentum are called *helicities*, hence when this basis is chosen the mode decomposition is called helicity decomposition. The notion of helicity is closely related to the notion of polarization. More precisely, modes of helicity ± 1 describe waves of right-handed and left-handed circular polarization respectively, and different linear combinations can be taken to describe for instance linear or elliptic polarizations. In particular, the helicity-2, linear polarizations are

$$\epsilon^{+}_{\mu\nu} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \epsilon^{\times}_{\mu\nu} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{1.64}$$

The labels stand for ‘plus’ (not to be confused with the plus used in the circular polarizations, and which will not be used in the following) and ‘cross’, and the reason will become clear below.

The helicity decomposition with an arbitrary momentum $\vec{k}$ can be conveniently described introducing the transverse and longitudinal projectors

$$T^a_b := \delta^a_b - \frac{k^a k_b}{\vec{k}^2}, \qquad L^a_b := \frac{k^a k_b}{\vec{k}^2}. \qquad [1.65]$$

Using these, we can decompose a spin-1 vector as follows,

$$A^a = A^a_{\mathrm{T}} + A^a_{\mathrm{L}}, \qquad A^a_{\mathrm{T}} = P^{(\mathbf{1})a}_{\mathrm{T}\ b} A^b, \qquad A^a_{\mathrm{L}} = P^{(\mathbf{1})a}_{\mathrm{L}\ b} A^b, \qquad [1.66]$$

where

$$P^{(\mathbf{1})a}_{\mathrm{T}\ b} = T^a_b, \qquad P^{(\mathbf{1})a}_{\mathrm{L}\ b} = L^a_b, \qquad P^{(\mathbf{1})}_{\mathrm{T}} + P^{(\mathbf{1})}_{\mathrm{L}} = P^{(\mathbf{1})}; \qquad [1.67]$$

and a spin-2 tensor as follows,

$$\begin{aligned} h_{\langle ab\rangle} &= h^{\mathrm{TT}}_{ab} + h^{\mathrm{L}}_{ab} + h^{\mathrm{LL}}_{ab}, \qquad h^{\mathrm{TT}}_{ab} = P^{(\mathbf{2})cd}_{\mathrm{TT}\ ab} h_{cd} \\ h^{\mathrm{L}}_{ab} &= P^{(\mathbf{2})cd}_{\mathrm{L}\ ab} h_{cd}, \qquad h^{\mathrm{LL}}_{ab} = P^{(\mathbf{2})cd}_{\mathrm{LL}\ ab} h_{cd}, \end{aligned} \qquad [1.68]$$

where

$$\begin{aligned} &P^{(\mathbf{2})ab}_{\mathrm{TT}\ cd} = T^a_{(c} T^b_{d)} - \frac{1}{2} T^{ab} T_{cd}, \qquad P^{(\mathbf{2})ab}_{\mathrm{L}\ cd} = T^a_{(c} L^b_{d)} + T^b_{(c} L^a_{d)}, \qquad [1.69] \\ &P^{(\mathbf{2})ab}_{\mathrm{LL}\ cd} = \frac{1}{3}\left(\frac{1}{2} T^{ab} T_{cd} + 2L^{ab} L_{cd} - T^{ab} L_{cd} - L^{ab} T_{cd}\right), \\ &P^{(\mathbf{2})} = P^{(\mathbf{2})}_{\mathrm{TT}} + P^{(\mathbf{2})}_{\mathrm{L}} + P^{(\mathbf{2})}_{\mathrm{LL}}. \end{aligned}$$

We can then identify the helicities of the different projectors studying how they transform under a rotation with axis $\vec{k}$. For simplicity let us consider the case when $\vec{k}$ is in the z direction, so that we can use the formulas already derived for the eigenvectors. In this case,

$$T^a{}_b = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad L^a{}_b = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad [1.70]$$

Then [1.66] reduces to $A^{\mathrm{T}}_a = (A_x, A_y, 0)$ and $A^{\mathrm{T}}_a = (0, 0, A_z)$. Comparing these to the earlier decomposition [1.57], we conclude that the transverse projector contains the ± 1 helicity modes of a spin-1 field, and the longitudinal projector the 0 helicity mode. The spin-2 projectors based on [1.70] give

$$P^{(\mathbf{2})\,ab}_{\mathrm{TT}\;\;cd}h^{cd} = (ThT)^{ab} - \frac{1}{2}\mathrm{Tr}(Th)T^{ab}$$
$$= \begin{pmatrix} \frac{1}{2}(h_{xx}-h_{yy}) & h_{xy} & 0 \\ h_{xy} & -\frac{1}{2}(h_{xx}-h_{yy}) & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} h_+ & h_\times & 0 \\ h_\times & -h_+ & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad [1.71a]$$

$$P^{(\mathbf{2})\,ab}_{\mathrm{L}\;\;cd}h^{cd} = \begin{pmatrix} 0 & 0 & h_{xz} \\ 0 & 0 & h_{yz} \\ 0 & 0 & 0 \end{pmatrix}, \qquad P^{(\mathbf{2})\,ab}_{\mathrm{LL}\;\;cd}h^{cd} = \frac{1}{6}(h_{xx}+h_{yy}-2h_{zz})\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \qquad [1.71b]$$

Comparing these to the earlier decomposition [1.61], we conclude that the spin-2 part of the gravitational perturbation can be decomposed into 5 helicity modes $\pm 2, \pm 1, 0$, which are carried respectively by the TT, L and LL components.

The spin-helicity interpretation of the components of $h_{\mu\nu}$ is a kinematical classification based on a choice of reference frame given by the time direction[12] and the spatial direction $\vec{k}$. The next question is which of these components are dynamical. If the field satisfied the simple wave equation $\Box h_{\mu\nu} = 0$, then all components would be dynamical, and the field would carry 10 independent degrees of freedom corresponding to all the helicity states described above. But this is not the case of [1.47], because of the additional derivative operators present, and the gauge redundancy. One way to identify the degrees of freedom is to fix the gauge and study the resulting solutions. This is what we do next.

12. Which we have chosen to coincide with t of the background Cartesian coordinates. However the whole spin-helicity description is perfectly Lorentz covariant. The spin projectors for an arbitrary time-direction τ^μ, $\tau^2 = -1$, are

$$P^{(\mathbf{1})\mu}{}_{\nu} = q^\mu_\nu := \delta^\mu_\nu + \tau^\mu\tau_\nu, \qquad P^{(\mathbf{0})\mu}{}_{\nu} = -\tau^\mu\tau_\nu, \qquad 1^{(\frac{1}{2},\frac{1}{2})} = P^{(\mathbf{1})} + P^{(\mathbf{0})}, \; [1.72]$$

for a vector, and

$$P^{(\mathbf{2})\mu\nu}{}_{\rho\sigma} = \delta^\mu_{(\rho}\delta^\nu_{\sigma)} + \tau^\mu\tau_{(\rho}\delta^\nu_{\sigma)} - \frac{1}{2}q^{\mu\nu}q_{\rho\sigma}, \qquad P^{(\mathbf{1})\mu\nu}{}_{\rho\sigma} = -\tau^\mu\tau_{(\rho}q^\nu_{\sigma)},$$
$$P^{(\mathbf{0})\mu\nu}{}_{\rho\sigma} = \tau^\mu\tau_\rho\tau^\nu\tau_\sigma, \qquad P_s^{(\mathbf{0})\mu\nu}{}_{\rho\sigma} = \frac{1}{2}q^{\mu\nu}q_{\rho\sigma}, \qquad [1.73]$$
$$1^{(\mathbf{1},\mathbf{1})\oplus(\mathbf{0},\mathbf{0})} = P^{(\mathbf{2})} + P^{(\mathbf{1})} + P^{(\mathbf{0})} + P_s^{(\mathbf{0})},$$

for a symmetric tensor.

1.3.4. *De Donder and TT gauges*

The gauge symmetry can be exploited to simplify the linearized field equations. This can be done using [1.50] to put the metric perturbation in a form where it satisfies

$$\partial_\mu h^\mu{}_\nu - \frac{1}{2}\partial_\nu h = 0. \qquad [1.74]$$

Doing so eliminates all terms containing divergences. The linearized field equations [1.47] are thus equivalent to

$$\Box \bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu}, \qquad \partial_\mu \bar{h}^{\mu\nu} = 0, \qquad \bar{h}_{\mu\nu} := h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h. \qquad [1.75]$$

The short-hand notation $\bar{h}_{\mu\nu}$ is introduced for convenience, and it is known as trace-reversed perturbation, since $\bar{h} = -h$. The condition [1.74] is called De Donder gauge in the literature, but also Lorenz gauge, in analogy with electromagnetism, and harmonic gauge, because it preserves harmonic coordinates, as can be seen linearizing [1.29]. The compatibility of the coupled system is guaranteed by the fact that $\partial_\mu T^{\mu\nu}$ vanishes on solutions of the matter field equations, more on this in Section 2.2.1 below.

The condition [1.74] does *not* fix completely the 4-dimensional diffeomorphism symmetry. Intuitively, this occurs because we are not fixing metric coefficients, but only their derivatives. To make the residual freedom explicit, observe that the diffeomorphism required to put an arbitrary metric perturbation in De Donder form is given by a solution of the equation

$$\Box \xi^\mu = -\partial_\nu h^{\mu\nu} + \frac{1}{2}\partial^\mu h. \qquad [1.76]$$

But for a given initial metric, this equation admits infinitely many solutions, parametrized by the zero modes of the d'Alembertian operator. In other words, once the De Donder condition is satisfied, there remains a residual freedom of gauge transformations that satisfy $\Box \xi^\mu = 0$. Thus the De Donder gauge contains in fact an infinite family of inequivalent gauge fixings. For this reason, it may be more appropriate to refer to [1.74] not as a gauge fixing condition, but rather a *family* of gauge fixings. The simpler albeit vaguer term De Donder gauge is, however, the one in use in most of the literature.

A unique representative of the De Donder gauge can be specified fixing the solution of $\Box \xi^\mu = 0$ in terms of initial data at a reference $t = 0$ hypersurface.

Such initial data can always be chosen such that any four components of $h_{\mu\nu}$ and their time derivatives vanish there. For vacuum solutions, this means that the chosen components vanish everywhere. It is convenient to apply this procedure and set to zero the components $\bar{h}$, so that $\bar{h}_{\mu\nu} = h_{\mu\nu}$, and h_{0a}. Then the De Donder condition implies that also $\partial_a h^{ab}$ and $\partial_0 h_{00}$ vanish everywhere. The vacuum equations then imply $\vec{\partial}^2 h_{00} = 0$, hence the only solution with vanishing boundary conditions is $h_{00} = 0$. We conclude that in this gauge, vacuum solutions satisfy

$$h = h_{00} = h_{0a} = \partial_a h^{ab} = 0, \qquad h_{ab} = h^{\rm TT}_{ab}. \tag{1.77}$$

This shows explicitly that there are only two independent degrees of freedom. We call this choice the *transverse-traceless gauge* (TT gauge in short), since in this gauge, vacuum solutions coincide with the transverse-traceless perturbations. We stress that this property and the equations [1.77] are only valid for vacuum solutions. In other words, [1.77] is not the definition of a gauge condition, but rather the specific value that solutions take in a certain gauge. The analogue for solutions with sources will be discussed below. Notice also that $h_{0\mu} = 0$ is the linearized approximation of [1.28]. Therefore the TT gauge of vacuum solutions implies the temporal gauge, hence the TT coordinates describe free-falling observers.

1.3.5. *Vacuum solutions*

Let us flesh out these considerations by looking at the explicit form of the vacuum solutions. This will make it clear that the identification of the degrees of freedom with the transverse-traceless modes is a gauge-invariant statement. The vacuum equations can be solved straightforwardly taking linear combination of plane waves via the Fourier transform

$$h_{\mu\nu}(x) = \mathrm{Re} \int d^4k \, \tilde{h}_{\mu\nu}(k) e^{ik\cdot x}. \tag{1.78}$$

Imposing the vacuum equations and the De Donder condition [1.75] requires

$$k^2 = 0, \qquad k^\mu \tilde{h}_{\mu\nu} = \frac{1}{2} k_\nu \tilde{h}. \tag{1.79}$$

The first equation is solved by $k^0 = \pm|\vec{k}\,|$, and we denote $\omega := k^0 c > 0$ for a future-pointing 4-momentum. The second equation gives a linear system of 4 conditions on the 10 components of the matrix. In the frame where $\vec{k}$ is along the z axis, we take as independent components

$$\tilde{h}_{0\mu}, \qquad h_+ = \tilde{h}_{xx}, \qquad h_\times = \tilde{h}_{xy}, \tag{1.80}$$

and then the system is solved by

$$\tilde{h}_{xz} = -\tilde{h}_{0x}, \qquad \tilde{h}_{yz} = -\tilde{h}_{0y}, \qquad \tilde{h}_{zz} = -\tilde{h}_{00} - 2\tilde{h}_{0z}, \qquad \tilde{h}_{yy} = -\tilde{h}_+ . \tag{1.81}$$

Therefore the general solution for a real, monochromatic wave with frequency ω and propagating along the z axis is a 6-parameter family given by

$$h_{\mu\nu}(x) = \begin{pmatrix} \tilde{h}_{00} & \tilde{h}_{0x} & \tilde{h}_{0y} & \tilde{h}_{0z} \\ & h_+ & h_\times & -\tilde{h}_{0x} \\ & & -h_+ & -\tilde{h}_{0y} \\ & & & -\tilde{h}_{00} - 2\tilde{h}_{0z} \end{pmatrix} \cos(k\cdot x), \qquad k\cdot x = -\omega(t - z/c). \tag{1.82}$$

The solution, however, contains gauge redundancy. Recall in fact that there is residual gauge freedom in the form of diffeomorphisms ξ^μ that satisfy the vacuum wave equation. In Fourier space the gauge transformation [1.50] read $\delta_\xi \tilde{h}_{\mu\nu} = 2ik_{(\mu}\tilde{\xi}_{\nu)}$, where $\tilde{\xi}^\mu(k)$ is the Fourier transform of $\xi^\mu(x)$ with $k^2 = 0$ in order to satisfy the vacuum wave equation and be an admissible residual gauge. Under the residual gauge transformation,

$$\tilde{h}_{0\mu} \to \tilde{h}_{0\mu} + ik_0\tilde{\xi}_\mu + ik_\mu\tilde{\xi}_0, \qquad h_{+,\times} \to h_{+,\times}. \tag{1.83}$$

The second property follows immediately from the fact that k_μ has no transverse components. We conclude that only the two components $h_{+,\times}$ are gauge-invariant and thus physically relevant. These can be recognized as the TT components. The four coefficients $\tilde{h}_{0\mu}$ can be changed arbitrarily, and have no physical meaning. We can in particular set them to zero simultaneously, choosing $i\tilde{\xi}_\mu = \frac{c}{\omega}(\frac{1}{2}\tilde{h}_{00}, \tilde{h}_{0x}, \tilde{h}_{0y}, \tilde{h}_{0z} - \frac{1}{2}\tilde{h}_{00})$. Then

$$h_{\mu\nu}(x) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cos\big(\omega(t - z/c)\big). \tag{1.84}$$

We have thus achieved the TT form [1.77]. In this gauge, the only components of the vacuum solutions are the gauge-invariant transverse-traceless ones. Accordingly, the solution can also be written as a sum over the two TT polarization tensors [1.64], namely

$$\tilde{h}_{\mu\nu}(k) = \sum_p A_p(k)\epsilon^p_{\mu\nu}(k), \qquad [1.85]$$

where $p = +, \times$ and $A_p = h_p$.[13]

The resulting perturbed metric is

$$ds^2 = -dt^2 + (1 + h_+ \cos k\cdot x)dx^2 + (1 - h_+ \cos k\cdot x)dy^2 + 2h_\times \cos k\cdot x\, dxdy + dz^2. \qquad [1.86]$$

Let us pause for a short historical digression, about some of the confusion that hindered historically the understanding of gravitational waves. Let us consider any member of the general solution [1.82] with $h_+ = h_\times = 0$. This looks like a genuine wave, and it is a solution of the linearized Einstein's equations. However, it is a pure gauge solution, and can be set to vanish identically without loss of physical information. In other words, there are wave solutions which are in the end only coordinate artefacts. For the same reason, it is also possible to change coordinates so that the argument $ct - z$ in its cosine is replaced by $vt - z$ for an arbitrary constant v. Hence the pure gauge modes don't really propagate, and if a gauge is chosen so that they look like they are propagating, well one can do this with an arbitrary speed, the speed is not constrained in any way by the dynamics. To use Eddington's words, the non-physical gauge modes propagate at the “speed of thought". This initial confusion was clarified by the identification of gauge-invariant components, and the fact that for the physical modes, the propagation speed is fixed to be the speed of light by Einstein's equations. However, additional doubts persisted, because the metric [1.86] is *not* a solution of the exact Einstein's equations; only of the linearized theory. In other words, there are vacuum solution the take approximately the form [1.86] in some regions of spacetime, but none that has that exact form everywhere in

13. The polarization decomposition can also be used to solve the wave equation. In this approach one writes [1.85] with the sum including all ten polarization modes [1.61] and ten arbitrary coefficients A_i. Then the relations [1.81] between Cartesian components are replaced by relations between the polarization coefficients. The algebra is slightly more involved but the end result is the same. We will see the use of the the polarization mode procedure on cosmological backgrounds in Sec 2.5.

spacetime. This raised the issue of whether gravitational waves existed in the full theory, or were only an artefact of the linearized approximation. This more complicated issue was solved only much later, when a non-perturbative identification of the wave degrees of freedom and their energy was made possible by the work of Bondi, Sachs and many others (Bondi 1960 ; Bondi *et al.* 1962 ; Sachs 1962 ; Newman and Penrose 1968 ; Ashtekar and Streubel 1981 ; Dray and Streubel 1984 ; Wald and Zoupas 2000).

Coming back to the physical solution [1.84], we see that the resulting wave tensor is transverse to the direction of propagation. Comparing with [1.61], we conclude that the physical gauge-invariant modes have helicity ± 2. All modes of helicity ± 1 and 0 have dropped out, either by gauge fixing, or by solving the vacuum field equations. Because the physical modes have helicity ± 2, it takes a rotation of an angle $\pi/4$ to turn one mode into the other. This can be compared with the electromagnetic case, whose modes have helicity ± 1, and it takes a rotation by $\pi/2$ to turn one mode into the other, see Fig. 2.1.[14] For the electromagnetic case, this angle $\pi/2$ corresponds to the orthogonality of the oscillations of the electric and magnetic fields. For the gravitational case, it corresponds to the antipodal symmetry in tidal forces. We will see below in Section 2.1 how this intuition can be made precise by studying the effect of a gravitational wave on test particles, see in particular Fig. 2.1.

The solution [1.84] describes a monochromatic wave. The most general solution has ∞^3 Fourier components (one per choice of $\vec{k}$), and two independent degrees of freedom per component (the values of $h_{+,\times}$), therefore it is described by $2 \times \infty^3$ arbitrary numbers. These are the independent degrees of freedom of gravitational waves. It is the same number of the full theory, so the linearized approximation simplifies the dynamics but preserves the number of independent variations of the gravitational field that can occur in the full theory. Notice that the number of independent degrees of freedom is the same of electromagnetism, or of two scalar fields. A part from the dynamical behaviour, what changes is also the behaviour of these degrees of freedom under Lorentz transformations, because of their different spins.

The analysis has also shown that only the transverse-traceless modes are gauge invariant. We did this using the partial gauge fixing provided by the De Donder condition, but it can be proved in full generality starting from the projector $P^{(\mathbf{2})\mathrm{TT}}$ on transverse-traceless modes defined in [1.69], and observing that it annihilates gauge transformations. Since we are going to use this projector often, we drop the label $(\mathbf{2})$ from now on. We also introduce the notation

14. The general formula is that two modes of helicity $\pm j$ are related by $\pi/2j$.

$\hat{k} = \vec{k}/k^0$, using which we can write the explicit form of the projector as

$$P^{\mathrm{TT}}{}^{ab}_{cd}(\hat{k}) = \delta^a_{(c}\delta^b_{d)} - \frac{1}{2}\delta^{ab}\delta_{cd} - \delta^a_{(c}\hat{k}^b\hat{k}_{d)} - \hat{k}^a\hat{k}_{(c}\delta^b_{d)} + \frac{1}{2}(\delta^{ab}\hat{k}_c\hat{k}_d + \hat{k}^a\hat{k}^b\delta_{cd} + \hat{k}^a\hat{k}^b\hat{k}_c\hat{k}_d). \quad [1.87]$$

Note that The symmetrization on the indices here and in [1.69] is omitted in some books (Poisson and Will 2014 ; Maggiore 2007), under the premises that one is applying it to symmetric tensors only anyway. In conclusion, we do not need to use the TT gauge, nor even the De Donder gauge, in order to identify the independent degrees of freedom. Whatever gauge we are using, we can always extract them via

$$h^{\mathrm{TT}}_{ab} = P^{\mathrm{TT}}{}^{cd}_{ab}h_{cd}, \quad [1.88]$$

and the result of this projection is gauge invariant. If we align the frame so that $\hat{k}$ coincides with the z axis, it is given by [1.71a]. To treat a general direction, we parametrize it using polar coordinates (θ, φ) on the sphere, and write

$$\hat{k}(\theta,\varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\phi, \cos\theta) = R_{\hat{z}}(\varphi)R_{\hat{y}}(\theta)\hat{e}_z =: R(\theta,\varphi)\hat{e}_z. [1.89]$$

Then we can use the inverse of this rotation to align the arbitrary direction back with z and use again [1.71a] to extract the TT components, namely

$$h^{\mathrm{TT}}_{ab} = P^{\mathrm{TT}}{}^{cd}_{ab}(\hat{z})R(\theta,\varphi)^e{}_c h_{ef}(\hat{k})R(\theta,\varphi)^f{}_d. \quad [1.90]$$

We obtain in this way

$$h_+ = \frac{1}{2}\Big(h_{xx}(\cos^2\theta\cos^2\varphi - \sin^2\varphi) + h_{yy}(\cos^2\theta\sin^2\theta - \cos^2\varphi) + h_{zz}\sin^2\theta + h_{xy}(1+\cos^2\theta)\sin 2\varphi - h_{xz}\cos\varphi\sin 2\theta - h_{yz}\sin\varphi\sin 2\theta\Big), \quad [1.91a]$$

$$h_\times = h_{xy}\cos 2\varphi\cos\theta + h_{xz}\sin\varphi\sin\theta - h_{yz}\cos\varphi\sin\theta - \frac{h_{xx}-h_{yy}}{2}\sin 2\varphi\cos\theta. \quad [1.91b]$$

1.3.6. *Gauge-invariant description: independent and constrained degrees of freedom*

The analysis of vacuum solutions has allowed us to identify the independent degrees of freedom of the gravitational field. There are also dependent degrees of freedom, namely components of the field that are gauge invariant, but uniquely determined from the sources. These can be exposed looking at the constraints, namely the 00 and $0a$ components of [1.47] which give rise to elliptic equations, as opposed to hyperbolic ones. The intricacies of the tensorial structure make it however convenient to do first a kinematical analysis of gauge invariance using the transverse and longitudinal projectors. To that end, we first observe that the projectors can be described also in configuration space without doing the Fourier transform. In this case [1.65] is replaced by

$$T^a_b = \delta^a_b - \frac{\partial^a \partial_b}{\vec{\partial}^2}, \qquad L^a_b := \frac{\partial^a \partial_b}{\vec{\partial}^2}. \tag{1.92}$$

These expressions are somewhat implicit because one needs to specify boundary conditions in order to have a well-defined inverse of the Laplace operator $\vec{\partial}^2$. Requiring the fields to vanish at spatial infinity makes this definition equivalent to the one in momentum space. Notice that the transverse-longitudinal decomposition in configuration space can be recognized as the Helmholtz decomposition of a 3d vector field into solenoidal and irrotational parts. This representation also makes it clear that the projectors are local operators in Fourier space, but non-local in spacetime. It has the important consequence that it is not possible to identify exactly the gauge-invariant modes with local observations only.

It is also convenient to dispose of the projectors for the components with non-maximal helicities, by introducing auxiliary fields with smaller spin. For instance, we denote $W_a := h_{0a}$ the spin-1 part of the gravitational perturbation, and write its longitudinal part as the gradient of a scalar:

$$W^{\mathrm{L}}_a := P^{(1)}_{\mathrm{L}}{}^a_b W^b = \partial_a W, \qquad W = \vec{\partial}^{-2} \partial_a W^a. \tag{1.93}$$

Similarly for the spin-2 part, we can write the mixed transverse-longitudinal and fully longitudinal modes introducing a transverse vector B^a and a scalar

B,

$$P^{(2)\,ab}_{\text{L}\ \ cd}h^{cd} = 2\partial^{(a}T^{b)}_{c}\frac{\partial_d h^{cd}}{\vec{\partial}^2} = 2\partial^{(a}B^{b)}, \qquad B^b := \frac{2}{\vec{\partial}^2}T^b_c\partial_d h^{cd},$$

$$P^{(2)\,ab}_{\text{LL}\ cd}h^{cd} = \frac{3}{2}\left(\frac{\partial^a\partial^b}{\vec{\partial}^2} - \frac{1}{3}\delta^{ab}\right)\frac{\partial_{\langle c}\partial_{d\rangle}}{\vec{\partial}^2}h^{cd} = (\partial_a\partial_b - \frac{1}{3}\delta_{ab}\vec{\partial}^2)B,$$

$$B := \frac{3}{2}\frac{\partial_{\langle c}\partial_{d\rangle}}{\vec{\partial}^4}h^{cd}.$$

Let us summarize. With this new notation at hand, the gravitational perturbation around Minkowski can be decomposed into

spin 0 $\quad h_s$ [1.94a]

spin 0 $\quad h_{00}$ [1.94b]

spin 1 $\quad h_{0a} = W_a = W^{\text{T}}_a + \partial_a W$ [1.94c]

spin 2 $\quad h_{ab} = h^{\text{TT}}_{ab} + 2\partial_{(a}B_{b)} + (\partial_a\partial_b - \frac{1}{3}\delta_{ab}\vec{\partial}^{\,2})B + \frac{1}{3}\delta_{ab}h_s,$ [1.94d]

with

$$W^{\text{T}}_a = T^b_a W_b \qquad \text{helicity} \pm 1 \tag{1.95a}$$

$$W = \vec{\partial}^{-2}\partial_a W^a \qquad \text{helicity } 0, \tag{1.95b}$$

and

$$h^{\text{TT}}_{ab} := P^{(2)\,ab}_{\text{TT}\,cd}h^{cd} = \left(T^a_c T^b_d - \frac{1}{2}T^{ab}T_{cd}\right)h^{cd} \qquad \text{helicity} \pm 2 \tag{1.96a}$$

$$B^b := \frac{2}{\vec{\partial}^2}T^b_c\partial_d h^{cd} \qquad \text{helicity} \pm 1 \tag{1.96b}$$

$$B := \frac{3}{2}\frac{\partial_{\langle c}\partial_{d\rangle}}{\vec{\partial}^4}h^{cd} \qquad \text{helicity } 0 \tag{1.96c}$$

Next, we look at the behaviour of these different helicities under gauge transformation. We have used the Poincaré symmetry of the Minkowski background to organize the ten components of $h_{\mu\nu}$ in terms of spin and helicity. But Minkowski is not invariant under the general diffeomorphism symmetry [1.50], hence there is no reason to expect that this decomposition be gauge-invariant. Consider for instance the spin-0 part h_{00}: this is just a metric component, and manifestly not invariant under diffeomorphisms. Only the TT component is gauge-invariant, as can be seen explicitly replacing $h_{\mu\nu}$ in [1.96a] with its gauge transformation [1.50] and observing that it vanishes thanks to the transverse projection on both indices. This is the gauge-invariance of the two TT components already observed in (the second equation in) [1.81]. The remaining components transform as follows:

$$h'_{00} = h_{00} + 2\partial_0\xi_0, \qquad h'_s = h_s + 2\vec{\partial}^2\xi^{\text{L}}, \qquad \xi^{\text{L}} := \vec{\partial}^{-2}\partial_a\xi^a,$$
$$W_a^{\text{T}\prime} = W_a^{\text{T}} + \partial_0\xi_a^{\text{T}}, \qquad W' = W + \xi_0 + \partial_0\xi^{\text{L}},$$
$$B'_a = B_a + \xi_a^{\text{T}}, \qquad B' = B + 2\xi^{\text{L}}.$$

It is possible to combine them to find a maximal number of 6 gauge-invariant quantities, given by

$$h_{ab}^{\text{TT}}, \qquad c^{-2}\Phi := -\frac{1}{2}h_{00} + \partial_0 W - \frac{1}{2}\partial_0^2 B,$$
$$c^{-3}\Phi_a := \frac{1}{4}(W_a^{\text{T}} - \partial_0 B_a), \qquad c^{-2}\Psi := \frac{1}{6}(\vec{\partial}^2 B - h_s). \qquad [1.97]$$

The numerical factors and powers of c have been chosen for later convenience. We will also use a dot for the derivative with respect to the time coordinate $t = x^0/c$, e.g. $\partial_0 W = c^{-1}\dot{W}$. Our conventions for the physical dimensions are summarized in Table 1.2.

$$[x^\mu] = \mathtt{m} \qquad [t = x^0/c] = \mathtt{s} \qquad [\partial_\mu] = \mathtt{m}^{-1} \qquad [\partial_t = c\partial_0] = \mathtt{s}^{-1}$$
$$[g_{\mu\nu}] = [h_{\mu\nu}] = [h] = [h_s] = [W^{\text{T}}] = 1$$
$$[W] = [B_a] = \mathtt{m} \quad [B] = \mathtt{m}^2 \qquad [\Phi] = [\Psi] = \mathtt{m}^2\mathtt{s}^{-2} \qquad [\Phi_a] = \mathtt{m}^3\mathtt{s}^{-3}$$

Table 1.2: Dimensional conventions used, represented in terms of international system units.

Inserting the parametrization [1.94] in the linearized Einstein's equation [1.47] one finds that gauge dependent quantities drop out and only the gauge-independent ones remain. This allows us to decouple the tensorial equations into two sets, an hyperbolic one featuring the d'Alambertian operator alone, and an elliptic one featuring the Laplace operator alone:

$$\Box h_{ab}^{\text{TT}} = -\frac{16\pi G}{c^4}\sigma_{ab}, \qquad [1.98a]$$

$$\vec{\partial}^2\Psi = 4\pi G\rho, \qquad \vec{\partial}^2\Phi_a = 4\pi G s_a, \qquad \vec{\partial}^2(\Phi - \Psi) = \frac{12\pi G}{c^2}\left(\dot{s} + \frac{1}{3}\tau\right). \qquad [1.98b]$$

The sources on the right-hand side of these equations are the components of $T_{\mu\nu}$ projected in the same way as [1.94], namely

$$T_{00} = c^2\rho, \qquad T_{0a} = c(s_a + \partial_a s),$$
$$T_{ab} = \sigma_{ab} + 2\partial_{(a}\sigma_{b)} + \left(\partial_a\partial_b - \tfrac{1}{3}\delta_{ab}\vec{\partial}^2\right)\sigma + \tfrac{1}{3}\delta_{ab}\tau.$$

Rewriting the linearized Einstein equations [1.47] in the equivalent form [1.98] makes their three-sided structure, mentioned in subsection 1.2.1, manifest. Four equations were redundant and have dropped out, hence the field has four undetermined components, which can be assigned arbitrarily choosing a specific gauge. Four equations are elliptic, and describe constrained degrees of freedom, namely dynamical components of the gravitational field which are uniquely determined by the sources. Finally, two equations are hyperbolic, hence they contain free data, and describe how these independent degrees of freedom propagate and react to the sources. This analysis therefore establishes that the gravitational field has two independent degrees of freedom, which are carried by the two $h^{\rm TT}_{ab}$ components of the metric. These are gauge-invariant, and describe the ± 2 helicities of a spin-2 wave.

As for the remaining gauge-invariant components, we have seen that they satisfy Poisson equations, hence these are degrees of freedom that are entirely determined by the sources. For these reason they are sometimes called ‘Coulombic’ degrees of freedom. Their meaning can be elucidated looking at the post-Newtonian expansion, in which sources are moving slowly with respect to the speed of light. To begin with, let us first consider perfectly static sources.

For static sources in a given frame, $T_{00} = c^2\rho$ is the only non-vanishing component on the right-hand side of Einstein’s equations. Then the second and third equations in [1.98b] imply[15] $\Phi_a = 0$ and $\Psi = \Phi$, whilst the first gives Newton’s equation, and we can identify Φ with Newton’s potential. This is also the way in which one fixes the coupling constant of the full Einstein’s equations in terms of G and c. How about the other potentials? They are sourced by moving bodies, and their existence is a consequence of relativistic invariance, akin to the electromagnetic occurrence of the vector potential next to the Coulomb potential. They describe effects that collectively go under the name of ‘gravito-magnetism’. These include additional contributions to the precessions of equinoxes, light bending, and frame dragging or Lense-Thirring effect. Their effect can be studied looking at the Lagrangian for a massive test

15. Assuming trivial boundary conditions, see App 1.5.

particle, which gives

$$\begin{aligned}\mathcal{L} &= -mc\sqrt{-g_{\mu\nu}u^\mu u^\nu} = -mc\sqrt{-(\eta_{\mu\nu}+h_{\mu\nu})u^\mu u^\nu} \\ &= -mc^2 + \frac{1}{2}mv^2 - m\Phi + \\ &\quad + \frac{m}{c^2}\left(\frac{1}{8}v^4 + \frac{1}{2}\Phi^2 - \frac{1}{2}v^2\Phi - \Psi v^2 + 4\Phi_a v^a + c^2 h^{\text{TT}}_{ab}v^a v^b\right) + O(c^{-4})\end{aligned}$$

where $v^a = dx^a/dt$. The term in round bracket is the *first post-Newtonian correction*. As we will see below, $h^{\text{TT}} \sim c^{-4}$, hence the last term there is higher order: dissipative effects for massive particles only appear at 2PN.

1.3.7. *Gauge-fixed description*

Solving the decoupled equations [1.98] determines the gauge-invariant quantities [1.97] in terms of the sources and the initial conditions. These solutions *do not determine a metric*. Doing so requires the additional step of specifying the coordinates to be used. Only after the coordinates are given, or in other words only after a gauge is chosen, the physical degrees of freedom can be described in terms of a metric tensor.

The gauge-invariant approach is conceptually satisfying because it identifies the physical degrees of freedom and decouples the equations, making them easier to solve in principle. However, it is very limited in applicability. Firstly, the decoupling and simple identification of gauge-invariant quantities occur only for very special backgrounds, such as flat spacetime or homogeneous and isotropic.[16] Secondly, even when the background is Minkowski, a general identification of gauge-invariant quantities at a fixed order in perturbation theory is only possible at the linear level, as explained earlier.

To go beyond these limitations, it is easier to put to the side the gauge-invariant description, and work instead in a fixed gauge. In the gauge-fixed approach, one chooses coordinates that impose restrictions on the metric, and

16. On a general background, one can still build the analogue of the spin-helicity decomposition replacing the partial derivatives with covariant derivatives, although care is needed to invert the Laplacian and handle its non-commutativity with the covariant derivative. However, the metric now enters explicitly the decomposition of the energy-momentum tensor, hence the decoupling will be lost in general. The non-commutativity of covariant derivatives also hinders the identification of gauge-invariant quantities.

then solves for individual metric components in that gauge, like we did in Section 1.3.4. Namely we do not solve [1.98], but the original system [1.47], coupled to additional equations fixing the gauge. The additional equations remove the problem that the field equations are redundant and do not determine all metric components.

A simple example of gauge fixing is the temporal gauge

$$h_{00} = h_{0a} = 0. \tag{1.99}$$

This is the linearized version of the non-perturbative temporal gauge described in Section 1.2.3, and provides a complete gauge fixing of the 4-dimensional diffeomorphism symmetry. In this gauge the only non-trivial components of the metric are the spatial ones, and [1.97] reduces to

$$c^{-2}\Phi = -\frac{1}{2}\partial_0^2 B, \qquad c^{-3}\Phi_a = -\frac{1}{4}\partial_0 B_a, \qquad c^{-2}\Psi = \frac{1}{6}(\vec{\partial}^2 B - h_s). \tag{1.100}$$

All gauge-invariant potentials are encoded in the components of the spin-2 mode [1.94d]. A related choice would be to replace the condition on h_{00} with the trace condition $h = 0$. In this case

$$\begin{aligned} c^{-2}(\Phi + 3\Psi) &= -h_{00} + \frac{1}{2}\Box B, \qquad c^{-3}\Phi_a = -\frac{1}{4}\partial_0 B_a, \\ c^{-2}(\Phi - 3\Psi) &= -\frac{1}{2}(\partial_0^2 + \vec{\partial}^2)B. \end{aligned} \tag{1.101}$$

Again, the four metric components B, B_a and this time h_{00}, are fixed uniquely in terms of the sources.

Another simple gauge fixing is the gravitational equivalent of the Coulomb gauge, defined so that both spin-2 and spin-1 parts are purely transverse, namely

$$B_a = B = W = 0. \tag{1.102}$$

In this gauge the non-trivial components of the metric are

$$h_{00}, \qquad h_{0a}^{\mathrm{T}} = W_a^{\mathrm{T}}, \qquad h_{ab} = h_{ab}^{\mathrm{TT}} + \frac{1}{3}\delta_{ab}h_s, \tag{1.103}$$

and their relation to the gauge invariant potential is given by

$$c^{-2}\Phi = -\frac{1}{2}h_{00}, \qquad c^{-3}\Phi_a = \frac{1}{4}W_a^{\mathrm{T}}, \qquad c^{-2}\Psi = -\frac{1}{6}h_s. \tag{1.104}$$

When fixing the coordinate gauge, one should keep in mind that some choices may be better than others, as we discussed with the Schwarzschild example in Section 1.2.3.[17] The examples above are non-covariant with respect to the Lorentz symmetry of the background, because they make reference to a given time foliation, and treat time and space components differently. In the presence of radiation, it is best to use a covariant gauge, because it simplifies the analysis of the solutions. We can, in fact, remark that while the above examples simplify the relation between the potentials and individual metric components, the remaining field equations for the propagating degrees of freedom are complicated. Whereas with the covariant De Donder gauge, all field equations took the simpler form [1.75]. However while the De Donder condition [1.74] can be also imposed in the presence of sources, we can no longer select a unique representative satisfying the TT condition [1.77]. In the presence of sources, a unique representative of the De Donder gauge can be specified as follows. The general solution of [1.75] is

$$\bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4}\int d^4x' G(x,x')T_{\mu\nu}(x') + \bar{h}^{\circ}_{\mu\nu}, \qquad [1.105]$$

where $G(x,x')$ is the d'Alembertian's Green function (see Appendix 1.5), $\bar{h}^{\circ}_{\mu\nu}$ any solution of the homogeneous equation, and the gauge condition is maintained via $\partial_\mu T^{\mu\nu} = \partial_\mu \bar{h}^{\circ\mu\nu} = 0$. Since both $\bar{h}^{\circ}_{\mu\nu}$ and the residual diffeomorphism parameters ξ^μ satisfy the vacuum wave equation, we can use the residual freedom to set to zero any four components of $\bar{h}^{\circ}_{\mu\nu}$. For instance, we can choose $\bar{h}^{\circ} = 0$, so that $\bar{h}^{\circ}_{\mu\nu} = \bar{h}^{\circ}_{\mu\nu}$, as well as $\bar{h}^{\circ}_{0a} = 0$. Then the De Donder condition implies that four more components of the homogeneous solution also vanish everywhere, specifically $\bar{h}^{\circ}_{00}$ and $\partial_b \bar{h}^{\circ ab}$. At this stage the gauge is completely fixed, and the only components left in the homogeneous solution are the gauge-invariant ones $\bar{h}^{\circ\mathrm{TT}}_{\mu\nu}$, which carry the independent degrees of freedom of gravitational waves. The general solution is

$$\bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4}\int d^4x' G(x,x')T_{\mu\nu}(x') + h^{\circ\mathrm{TT}}_{\mu\nu}. \qquad [1.106]$$

17. If $GM/c^2 \ll r$ in the static spherical coordinates, we can treat it as a perturbative solution with $h_{\mu\nu} = 2GM/(c^2 r)(\delta^t_\mu \delta^t_\nu + \delta^r_\mu \delta^r_\nu)$. We then see that these coordinates correspond to the Coulomb gauge fixing. Changing coordinates so to have temporal gauge with $h_{00} = 0$ would push the Newton potential in the B component, and introduce a dependence on the t coordinate, as mentioned earlier.

It is a complete gauge fixing that singles out a unique element of the De Donder family, and reduces to [1.77] for vacuum solutions. For generic sources, all components of the metric perturbation are non-zero in this gauge.[18]

A special situation occurs if the sources are static. In this case, it is possible to specialize the De Donder gauge so that

$$\partial_a h^a{}_0 = 0, \qquad \partial_a h^a{}_b + \frac{1}{3}\partial_b h_s = 0. \tag{1.107}$$

To prove this, we first observe that when rewritten in terms of the gauge-invariant quantities [1.97], the De Donder condition [1.74] implies

$$\Box W = -\frac{4}{c^3}\dot{\Psi}, \qquad \Box B = \frac{2}{c^2}(\Phi - \Psi), \qquad \Box B_a = \frac{4}{c^4}\dot{\Phi}_a. \tag{1.108}$$

For static sources $\Phi_a = \dot{\Phi} = 0$ and $\Psi = \Phi$, making it is possible to specialize the De Donder gauge requiring that W, B and B_a vanish everywhere, and then $h_{00} = -2c^{-2}\Phi$. This means achieving the Coulomb gauge described earlier, as opposed to the TT gauge, and again this is a complete gauge fixing singling out a member of the De Donder family. In other words, the De Donder gauge is compatible with the Coulomb gauge for static sources. Notice that this is what happens also in electromagnetism, where the Lorenz gauge is compatible with the Coulomb gauge for static sources.

18. More precisely, in the region causally connected to the sources, since the retarded Green function vanishes outside the light cone. So if the sources are present at all times, the metric perturbation is non-zero everywhere, whereas if the sources are ‘turned on’ at some initial time, then the perturbation vanishes outside the causal domain of the sources from that initial time. Alternatively, it is also possible to select a unique representative whose entire metric components $h_{00} = h_{0a}$ and their derivatives vanish on a given initial value surface, as considered in (Wald 1984). However these components will remain zero only in the region outside the causal domain of the sources from the initial value surface, so while interesting in principle, it is of less practical use. The procedure describe in the text describes instead a structure of the solutions valid at all times, and it is the analogue of the complete gauge-fixing used for instance in electro-magnetism by the Lienard-Wiechert potentials of a moving charge.

1.4. Appendix: Second order action for perturbations around any background solution

This appendix gives the details of the steps required to calculate the Einstein-Hilbert action of GR to second order in perturbations about a general background metric $\bar{g}_{\mu\nu}$:

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}. \quad [1.109]$$

The starting point is

$$S_{\text{EH}} \equiv \int d^4x\sqrt{-g}R = S^{(0)} + S^{(1)} + S^{(2)} + \dots \quad [1.110]$$

where $S^{(i)}$ is of $\mathcal{O}(h^i)$. The extremisation of the second order action $S^{(2)}$ with respect to $h_{\mu\nu}$ will give the linearized equations of motion for $h_{\mu\nu}$, namely those we want to calculate. The background metric $\bar{g}_{\mu\nu}$ satisfies the background Einstein equations which follow from $S^{(1)}$.

To find $S^{(2)}$ there are two main steps: calculating the Ricci scalar R to second order, and then the determinant of the metric and hence $\sqrt{-g}$ to second order.

Perturbed Riemann tensor, Ricci tensor and scalar, Einstein tensor

• The inverse metric, or *contravariant metric tensor* corresponding to Eq. [1.109] is given at second order by

$$g^{\mu\nu} = \bar{g}^{\mu\nu} - h^{\mu\nu} + h^{\mu\rho}h_{\rho}^{\nu} \quad [1.111]$$

where indices of $h_{\mu\nu}$ are raised and lowered with $\bar{g}_{\mu\nu}$.

• The perturbed *Christoffel symbols* are given by

$$\Gamma^{\rho}_{\mu\nu} \equiv \frac{1}{2}g^{\rho\lambda}\left(\partial_\mu g_{\nu\lambda} + \partial_\nu g_{\mu\lambda} - \partial_\lambda g_{\mu\nu}\right) = \bar{\Gamma}^{\rho}_{\mu\nu} + \Gamma^{\rho(1)}_{\mu\nu} + \Gamma^{\rho(2)}_{\mu\nu}. \quad [1.112]$$

In the following we denote the covariant derivative with respect to $\bar{g}$ by $\bar{\nabla}$ (with of course $\bar{\nabla}_\mu \bar{g}_{\nu\alpha} = 0$). Substitution of Eqs. [1.109] and [1.111] gives

$$\bar{\Gamma}^\rho_{\mu\nu} = \frac{1}{2}\bar{g}^{\rho\lambda}\left(\partial_\mu \bar{g}_{\nu\lambda} + \partial_\nu \bar{g}_{\mu\lambda} - \partial_\lambda \bar{g}_{\mu\nu}\right)$$

$$\Gamma^{\rho(1)}_{\mu\nu} = \frac{1}{2}\bar{g}^{\rho\lambda}\left(\bar{\nabla}_\mu h_{\nu\lambda} + \bar{\nabla}_\nu h_{\mu\lambda} - \bar{\nabla}_\lambda h_{\mu\nu}\right) \qquad [1.113]$$

$$\Gamma^{\rho(2)}_{\mu\nu} = -h^\rho_\beta \Gamma^{\beta(1)}_{\mu\nu}. \qquad [1.114]$$

• Next we calculate the *Riemann tensor*. Let

$$\Gamma^\rho_{\mu\nu} = \bar{\Gamma}^\rho_{\mu\nu} + \delta\Gamma^\rho_{\mu\nu} \qquad [1.115]$$

where, from (1.112),

$$\delta\Gamma^\rho_{\mu\nu} = \Gamma^{\rho(1)}_{\mu\nu} + \Gamma^{\rho(2)}_{\mu\nu}. \qquad [1.116]$$

The definition of the Riemann tensor together with (1.115) gives

$$R^\mu{}_{\nu\rho\sigma} \equiv \partial_\rho \Gamma^\mu_{\sigma\nu} + \Gamma^\mu_{\rho\lambda}\Gamma^\lambda_{\sigma\nu} - (\rho \leftrightarrow \sigma) \qquad [1.117]$$

$$= \bar{R}^\mu{}_{\nu\rho\sigma} + \bar{\nabla}_\rho(\delta\Gamma^\mu_{\sigma\nu}) - \bar{\nabla}_\sigma(\delta\Gamma^\mu_{\rho\nu})$$

$$+(\delta\Gamma^\mu_{\rho\lambda})(\delta\Gamma^\lambda_{\sigma\nu}) - (\delta\Gamma^\mu_{\sigma\lambda})(\delta\Gamma^\lambda_{\rho\nu}) \qquad [1.118]$$

On writing

$$R^\mu{}_{\nu\rho\sigma} \equiv \bar{R}^\mu_{\nu\rho\sigma} + R^{(1)\mu}{}_{\nu\rho\sigma} + R^{(2)\mu}{}_{\nu\rho\sigma}$$

and using (1.116) one can read off the different orders of the Riemann tensor. To first order

$$R^{(1)\mu}{}_{\nu\rho\sigma} = \bar{\nabla}_\rho \Gamma^{\mu(1)}_{\sigma\nu} - \bar{\nabla}_\sigma \Gamma^{\mu(1)}_{\rho\nu} \qquad [1.119]$$

$$= \frac{1}{2}\left[\left(\bar{\nabla}_\rho\bar{\nabla}_\sigma - \bar{\nabla}_\sigma\bar{\nabla}_\rho\right)h^\mu_\nu + \left(\bar{\nabla}_\rho\bar{\nabla}_\nu h^\mu_\sigma - \bar{\nabla}_\sigma\bar{\nabla}_\nu h^\mu_\rho\right)\right.$$

$$\left. - \left(\bar{\nabla}_\rho\bar{\nabla}^\mu h_{\sigma\nu} - \bar{\nabla}_\sigma\bar{\nabla}^\mu h_{\rho\nu}\right)\right] \qquad [1.120]$$

where we have used Eq. [1.113]. The second order term follows from (1.118) and (1.116) and reads

$$R^{(2)\mu}{}_{\nu\rho\sigma} = \left(\bar{\nabla}_\rho \Gamma^{\mu(2)}_{\sigma\nu} - \bar{\nabla}_\sigma \Gamma^{\mu(2)}_{\rho\nu}\right) + \left(\Gamma^{\mu(1)}_{\rho\lambda}\Gamma^{\lambda(1)}_{\sigma\nu} - \Gamma^{\mu(1)}_{\sigma\lambda}\Gamma^{\lambda(1)}_{\rho\nu}\right). \qquad [1.121]$$

• The *Ricci tensor* is then obtained by contraction:

$$R_{\nu\sigma} \equiv R^\mu{}_{\nu\mu\sigma} = \bar{R}_{\nu\sigma} + R^{(1)}_{\nu\sigma} + R^{(2)}_{\nu\sigma}. \qquad [1.122]$$

From (1.120) it follows that

$$R^{(1)}_{\nu\sigma} = \frac{1}{2}\left[\bar{\nabla}_\mu \bar{\nabla}_\sigma h^\mu_\nu + \bar{\nabla}_\mu \bar{\nabla}_\nu h^\mu_\sigma - \bar{\Box} h_{\nu\sigma} - \bar{\nabla}_\sigma \bar{\nabla}_\nu h\right], \qquad [1.123]$$

whereas from (1.121)

$$R^{(2)}_{\nu\sigma} = \bar{\nabla}_\rho \Gamma^{\rho(2)}_{\sigma\nu} - \bar{\nabla}_\sigma \Gamma^{\rho(2)}_{\rho\nu} + \Gamma^{\rho(1)}_{\rho\lambda}\Gamma^{\lambda(1)}_{\sigma\nu} - \Gamma^{\rho(1)}_{\sigma\lambda}\Gamma^{\lambda(1)}_{\rho\nu} \qquad [1.124]$$

Its explicit form is not required below, but for completeness we give it here:

$$\begin{aligned}\bar{R}^{(2)}_{\nu\sigma} &= \frac{1}{4}\bar{\nabla}_\nu h_{\alpha\beta}\bar{\nabla}_\sigma h^{\alpha\beta} + \frac{1}{2}\bar{\nabla}^\beta h^\beta_\sigma(\bar{\nabla}_\beta h_{\alpha\nu} - \bar{\nabla}_\alpha h_{\beta\nu}) \\ &+ \frac{1}{2}h^{\alpha\beta}\left(\bar{\nabla}_\nu \bar{\nabla}_\sigma h_{\alpha\beta} + \bar{\nabla}_\beta \bar{\nabla}_\alpha h_{\nu\sigma} - \bar{\nabla}_\beta \nabla_\sigma h_{\alpha\nu} - \bar{\nabla}_\beta \bar{\nabla}_\nu h_{\alpha\sigma}\right) \\ &- \frac{1}{2}(\bar{\nabla}_\beta h^{\alpha\beta} - \frac{1}{2}\bar{\nabla}^\alpha h)(\bar{\nabla}_\sigma h_{\nu\alpha} + \bar{\nabla}_\nu h_{\alpha\sigma} - \bar{\nabla}_\alpha h_{\nu\sigma}).\end{aligned}$$

• The *Ricci scalar* is obtained from Eqs. [1.111] and [1.122] and is given by

$$\begin{aligned}R &\equiv g^{\nu\sigma}R_{\nu\sigma} = (\bar{g}^{\nu\sigma} - h^{\nu\sigma} + h^{\nu\rho}h^\sigma_\rho)(\bar{R}_{\nu\sigma} + R^{(1)}_{\nu\sigma} + R^{(2)}_{\nu\sigma}) \\ &= \bar{R} + R^{(1)} + R^{(2)}, \end{aligned} \qquad [1.125]$$

where

$$R^{(1)} = \bar{g}^{\nu\sigma}R^{(1)}_{\nu\sigma} - \bar{R}^{\nu\sigma}h_{\nu\sigma} = \bar{\nabla}_\mu \bar{\nabla}_\nu h^{\mu\nu} - \Box h - \bar{R}^{\nu\sigma}h_{\nu\sigma} \qquad [1.126]$$

(in the last line we have used (1.123)), and

$$R^{(2)} = h^{\nu\alpha}h^\sigma_\alpha \bar{R}_{\nu\sigma} - h^{\nu\sigma}R^{(1)}_{\nu\sigma} + \bar{g}^{\nu\sigma}R^{(2)}_{\nu\sigma}. \qquad [1.127]$$

• For completeness we also give the first order perturbed *Einstein tensor*

$$G^{(1)}_{\mu\nu} = R^{(1)}_{\mu\nu} - \frac{1}{2}\bar{R}h_{\mu\nu} - \frac{1}{2}\bar{g}_{\mu\nu}R^{(1)} \quad [1.128]$$

$$= -\frac{1}{2}\bar{\Box}h_{\mu\nu} + \bar{\text{NewA}}_{(\mu}\bar{\text{NewA}}_{\rho}h^{\rho}{}_{\nu)} - \frac{1}{2}\bar{\text{NewA}}_{\mu}\bar{\text{NewA}}_{\nu}h - \frac{1}{2}\bar{g}_{\mu\nu}(\bar{\text{NewA}}_{\mu}\bar{\text{NewA}}_{\nu}h^{\mu\nu} - \bar{\Box}h)$$

$$+ \bar{G}_{\rho(\mu}h_{\nu)}{}^{\rho} - (\bar{R}_{\mu\rho\nu\sigma} - \frac{1}{2}\bar{g}_{\mu\nu}\bar{R}_{\rho\sigma})h^{\rho\sigma}. \quad [1.129]$$

Perturbed metric determinant

To expand $\sqrt{-g}$ to second order, we write Eq. [1.109] $g_{\alpha\beta} = \bar{g}_{\alpha\mu}(\delta^{\mu}_{\beta} + M^{\mu}_{\beta})$ where $M^{\mu}_{\beta} = \bar{g}^{\mu\lambda}h_{\lambda\beta}$. Thus

$$\det(g) = \det(\bar{g})\det(\mathbf{1} + \mathbf{M}) \quad [1.130]$$

where the matrix $\mathbf{M}$ has components M^{μ}_{β}. To quadratic order

$$\det(\mathbf{1} + \tau\mathbf{M}) = 1 + \text{tr}\mathbf{M} + \frac{1}{2}\left(2(\text{tr}\mathbf{M})^2 - \text{tr}(\mathbf{M}^2)\right) + \ldots \quad [1.131]$$

Replacing $\text{tr}\mathbf{M} = h$ and $\text{tr}(\mathbf{M}^2) = h_{\mu\nu}h^{\mu\nu}$ in Eq. [1.130], and then taking the square root, gives

$$\sqrt{-\det(g_{\mu\nu})} = \sqrt{-\det(\bar{g}_{\mu\nu})}\left[1 + \frac{1}{2}h + \frac{1}{8}(h^2 - 2h^2_{\mu\nu})\right]. \quad [1.132]$$

EH action to second order

Substituting (1.125) and (1.132) into the perturbed Einstein Hilbert action Eq. [1.110] gives

$$S_{(0)} = \int d^4x\sqrt{-\bar{g}}\bar{R}$$

$$S_{(1)} = \int d^4x\sqrt{-\bar{g}}\left(R^{(1)} + \frac{h}{2}\bar{R}\right)$$

$$S_{(2)} = \int d^4x\sqrt{-\bar{g}}\left(R^{(2)} + \frac{h}{2}R^{(1)} + \frac{\bar{R}}{8}(h^2 - 2h_{\mu\nu}h^{\mu\nu})\right).$$

• The *first order action* is

$$S_{(1)} = \int d^4x\sqrt{-\bar{g}}\left(\bar{\nabla}_\mu\bar{\nabla}_\nu h^{\mu\nu} - \Box h - \bar{R}^{\nu\sigma}h_{\nu\sigma} + \frac{h}{2}\bar{R}\right)$$

The first two terms are total derivatives. After integration by parts and dropping the boundary terms

$$S_{(1)} = \int d^4x\sqrt{-\bar{g}}\left(-\bar{G}^{\nu\sigma}h_{\nu\sigma}\right). \qquad [1.133]$$

On including matter through the stress tensor, the variation of this gives the background Einstein equation.

• The *second order* part becomes, on substituting [1.127],

$$S_{(2)} = \int d^4x\sqrt{-\bar{g}}\left(\bar{g}^{\nu\sigma}R^{(2)}_{\nu\sigma} + \frac{h}{2}R^{(1)} - h^{\nu\sigma}R^{(1)}_{\nu\sigma}\right.$$
$$\left. + \left[h^{\nu\alpha}h^{\sigma}_{\alpha}\bar{R}_{\nu\sigma} + \frac{\bar{R}}{8}(h^2 - 2h_{\mu\nu}h^{\mu\nu})\right]\right) \qquad [1.134]$$

where $R^{(1)}$ and $R^{(1)}_{\nu\sigma}$ are given in [1.126] and [1.123] respectively.

The first term $\int d^4x\sqrt{-\bar{g}}\bar{g}^{\nu\sigma}R^{(2)}_{\nu\sigma}$ splits into four parts on using (1.124). The first two parts are total derivatives and do not contribute. The last two parts give

$$\int d^4x\sqrt{-\bar{g}}\left(\bar{g}^{\nu\sigma}R^{(2)}_{\nu\sigma}\right) = +\int d^4x\sqrt{-\bar{g}}\bar{g}^{\nu\sigma}\left(\Gamma^{\mu(1)}_{\rho\lambda}\Gamma^{\lambda(1)}_{\sigma\nu} - \Gamma^{\mu(1)}_{\sigma\lambda}\Gamma^{\lambda(1)}_{\rho\nu}\right)$$
$$= \int d^4x\sqrt{-\bar{g}}\frac{1}{2}\left[(\bar{\nabla}^\lambda h)\left(\bar{\nabla}^\nu h_{\lambda\nu} - \frac{1}{2}\bar{\nabla}_\lambda h\right)\right.$$
$$\left. - (\bar{\nabla}^\nu h^{\mu\lambda})\left(\bar{\nabla}_\mu h_{\lambda\nu} - \frac{1}{2}\bar{\nabla}_\nu h_{\mu\lambda}\right)\right].$$

Collecting the expressions together into [1.134] gives

$$\begin{aligned} S_{(2)} = \frac{1}{2}\int d^4x\sqrt{-\bar{g}}\Bigg\{ & \left[(\bar{\nabla}^\lambda h)\left(\bar{\nabla}^\nu h_{\lambda\nu} - \frac{1}{2}\bar{\nabla}_\lambda h\right)\right. \\ & \left. - (\bar{\nabla}^\nu h^{\mu\lambda})\left(\bar{\nabla}_\mu h_{\lambda\nu} - \frac{1}{2}\bar{\nabla}_\nu h_{\mu\lambda}\right)\right] \\ & + \left[h\bar{\nabla}_\mu\bar{\nabla}_\nu h^{\mu\nu} - h\Box h\right] \\ & + \left[-2h^{\nu\sigma}\bar{\nabla}^\mu\bar{\nabla}_\sigma h_{\mu\nu} + h^{\mu\nu}\Box h_{\mu\nu} + (\bar{\nabla}_\sigma\bar{\nabla}_\nu h)h^{\sigma\nu}\right] \\ & + \left[-\bar{R}^{\mu\nu}hh_{\mu\nu} + 2h^{\nu\alpha}h^\sigma_\alpha\bar{R}_{\nu\sigma} + \frac{\bar{R}}{4}(h^2 - 2h_{\mu\nu}h^{\mu\nu})\right]\Bigg\}. \end{aligned}$$

Finally, after integration by parts,

$$\begin{aligned} S_{(2)} = \frac{1}{2}\int d^4x\sqrt{-\bar{g}}\Bigg\{ & -(\bar{\nabla}^\lambda h)(\bar{\nabla}^\nu h_{\lambda\nu}) + \frac{1}{2}(\bar{\nabla}_\lambda h)(\bar{\nabla}^\lambda h) \\ & -\frac{1}{2}(\bar{\nabla}^\nu h^{\lambda\mu})(\bar{\nabla}_\nu h_{\lambda\mu}) + (\bar{\nabla}^\nu h^{\lambda\mu})(\bar{\nabla}_\mu h_{\lambda\nu}) \\ & + \left[-\bar{R}^{\mu\nu}hh_{\mu\nu} + 2h^{\nu\alpha}h^\sigma_\alpha\bar{R}_{\nu\sigma} + \frac{\bar{R}}{4}(h^2 - 2h_{\mu\nu}h^{\mu\nu})\right]\Bigg\}. \end{aligned}$$

In flat space $\bar{g}_{\mu\nu} = \eta_{\mu\nu}$, the terms in square brackets all vanish and this reduces to the usual perturbed equations around Minkowski space.

In terms of the *trace reversed perturbation*

$$h_{\alpha\beta} = \bar{h}_{\alpha\beta} - \frac{1}{2}\bar{g}_{\alpha\beta}\bar{h} \qquad [1.135]$$

this becomes

$$\begin{aligned} S_{(2)} = \frac{1}{2}\int d^4x\sqrt{-\bar{g}}\Bigg\{ & -\frac{1}{2}(\bar{\nabla}^\nu\bar{h}^{\lambda\mu})(\bar{\nabla}_\nu\bar{h}_{\lambda\mu}) + \frac{1}{4}(\bar{\nabla}^\lambda\bar{h})(\bar{\nabla}^\lambda\bar{h}) \\ & +(\bar{\nabla}^\nu\bar{h}^{\lambda\mu})(\bar{\nabla}_\mu\bar{h}_{\lambda\nu}) + \\ & + \left[-\bar{R}^{\mu\nu}\bar{h}\bar{h}_{\mu\nu} + 2\bar{h}^{\nu\alpha}\bar{h}^\sigma_\alpha\bar{R}_{\nu\sigma} - \frac{\bar{R}}{2}\bar{h}_{\mu\nu}\bar{h}^{\mu\nu} + \frac{1}{4}\bar{R}\bar{h}^2\right]\Bigg\} \end{aligned}$$

Linearised equations of motion

- *Equations of motion.* Variation of action [1.134] with respect to $h^{\alpha\beta}$ gives

$$0 = \int d^4x\sqrt{-\bar{g}}\delta h^{\alpha\beta}\left(-G^{(1)}_{\alpha\beta} - \frac{1}{2}h\bar{G}_{\alpha\beta} + h^{\mu\nu}(\bar{g}_{\nu\beta}\bar{G}_{\alpha\nu} + \bar{g}_{\nu\alpha}\bar{G}_{\beta\nu})\right)$$

leading to the equations of motion

$$G^{(1)}_{\alpha\beta} = -\frac{1}{2}h\bar{G}_{\alpha\beta} + h^{\mu\nu}(\bar{g}_{\nu\beta}\bar{G}_{\alpha\nu} + \bar{g}_{\nu\alpha}\bar{G}_{\beta\nu}). \qquad [1.136]$$

In terms of the trace reversed perturbation, after commuting covariant derivatives, for example,

$$\bar{\nabla}_\mu\bar{\nabla}_\beta\bar{h}^\mu_\alpha = \bar{\nabla}_\beta(\bar{\nabla}_\mu\bar{h}^\mu_\alpha) + \bar{R}_{\lambda\beta}h^\lambda_\alpha - \bar{R}^\lambda{}_{\alpha\mu\beta}h^\mu_\lambda \qquad [1.137]$$

and then imposing the Lorenz gauge $\bar{\nabla}_\mu\bar{h}^{\mu\nu} = 0$ these read in empty space

$$\bar{\Box}\bar{h}_{\mu\nu} + 2\bar{R}_{\mu\rho\nu\sigma}\bar{h}^{\rho\sigma} - 2\bar{G}_{\rho(\mu}\bar{h}_{\nu)}{}^\rho - \bar{g}_{\mu\nu}(\bar{R}_{\rho\sigma}\bar{h}^{\rho\sigma}) = 0 \qquad [1.138]$$

which can be rewritten as

$$\bar{\Box}\bar{h}_{\alpha\beta} + 2\bar{R}^\mu{}_{\alpha\nu\beta}\bar{h}^\nu_\mu + S_{\mu\alpha\nu\beta}\bar{h}^{\mu\nu} = 0, \qquad [1.139]$$

where

$$S_{\mu\alpha\nu\beta} = [\bar{G}_{\mu\alpha}g_{\beta\nu} + \bar{G}_{\mu\beta}g_{\alpha\nu}] - \bar{R}_{\mu\nu}g_{\alpha\beta}. \qquad [1.140]$$

1.5. Appendix: Green's functions

Given a linear differential operator Δ acting on functions on $\mathbb{R}^n$, its corresponding Green's function is a function on $\mathbb{R}^n \times \mathbb{R}^n$ satisfying

$$\Delta G(x, x') = \delta^{(n)}(x, x'). \qquad [1.141]$$

Green's functions are useful to solve differential equations in the presence of sources, since they allow one to write the solutions of $\Delta\Phi = J$ as

$$\Phi(x) = \int d^nx' G(x, x')J(x') + \Phi^\circ(x), \qquad [1.142]$$

where Φ° is a solution of the homogeneous problem with $J = 0$. For a given Δ the Green function is typically not unique, but a unique one can be selected via boundary conditions or other physical requirements.

For the Laplace equation, $n = 3$ and $\Delta = \vec{\partial}^2$, there is a unique solution of [1.141] with vanishing boundary conditions at infinity, given by

$$G(\vec{x}, \vec{x}') = -\frac{1}{4\pi|\vec{x} - \vec{x}'|}. \tag{1.143}$$

Therefore

$$\Phi(\vec{x}) = \frac{1}{4\pi} \int d^3x' \frac{J(\vec{x}')}{|\vec{x} - \vec{x}'|} + \Phi^\circ(\vec{x}). \tag{1.144}$$

For the d'Alembert equation, $n = 4$ and $\Delta = \Box$, fixing vanishing boundary conditions at spatial infinity is not enough to have a single solution: there are infinitely many solutions of the homogeneous equation that can be added to any given G and still satisfy the defining equation. Two notable examples are the retarded and advanced ones, which are uniquely characterised by vanishing for x respectively in the past or the future of x', and given by

$$\begin{aligned} G_\pm(x, x') &= -\frac{\delta(t - t' \mp |\vec{x} - \vec{x}'|)}{4\pi|\vec{x} - \vec{x}'|} \\ &= -\frac{1}{2\pi}\Theta\big(\pm (t - t')\big)\delta\big((t - t')^2 - |\vec{x} - \vec{x}'|^2\big), \end{aligned} \tag{1.145}$$

where Θ is Heaviside's step function. The retarded solution imposes no-incoming radiation boundary conditions, and it is the one relevant to study the emission of waves from a source.

1.6. Bibliography

Abbott, B. P. *et al.* (2016), Observation of Gravitational Waves from a Binary Black Hole Merger, *Phys. Rev. Lett.*, 116(6), 061102.

Agazie, G. *et al.* (2023), The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background, *Astrophys. J. Lett.*, 951(1), L8.

Amaro-Seoane, P. *et al.* (2017), Laser Interferometer Space Antenna, *arXiv e-prints*, p. arXiv:1702.00786.

Andersson, N. (2019), *Gravitational-Wave Astronomy*, Oxford Graduate Texts, Oxford University Press.

Antoniadis, J. *et al.* (2023), The second data release from the European Pulsar Timing Array - III. Search for gravitational wave signals, *Astron. Astrophys.*, 678, A50.

Ashtekar, A., Khera, N., Kolanowski, M., Lewandowski, J. (2022), Charges and fluxes on (perturbed) non-expanding horizons, *JHEP*, 02, 066.

Ashtekar, A., Streubel, M. (1981), Symplectic Geometry of Radiative Modes and Conserved Quantities at Null Infinity, *Proc. Roy. Soc. Lond. A*, 376, 585–607.

Barack, L., Pound, A. (2019), Self-force and radiation reaction in general relativity, *Rept. Prog. Phys.*, 82(1), 016904.

Blanchet, L. (2006), Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries, *Living Rev. Rel.*, 9(1), 1–114.

Bondi, H. (1960), Gravitational Waves in General Relativity, *Nature*, 186(4724), 535–535.

Bondi, H., van der Burg, M. G. J., Metzner, A. W. K. (1962), Gravitational waves in general relativity. 7. Waves from axisymmetric isolated systems, *Proc. Roy. Soc. Lond. A*, 269, 21–52.

Buonanno, A., Damour, T. (1999), Effective one-body approach to general relativistic two-body dynamics, *Phys. Rev. D*, 59, 084006.

Caprini, C., Figueroa, D. G. (2018), Cosmological Backgrounds of Gravitational Waves, *Class. Quant. Grav.*, 35(16), 163001.

Chandrasekaran, V., Flanagan, E. E., Prabhu, K. (2018), Symmetries and charges of general relativity at null boundaries, *JHEP*, 11, 125.

Choquet-Bruhat, Y. (1952), Théorème d'existence pour certains systèmes d'équations aux dérivées partielles non linéaires, *Acta Mathematica*, 88, 141–225.

Colpi, M. *et al.* (2024), LISA Definition Study Report, , .

Deruelle, N., Uzan, J.-P. (2018), *Relativity in Modern Physics*, Oxford University Press.
URL: *https://global.oup.com/academic/product/relativity-in-modern-physics-9780198822156*

Dray, T., Streubel, M. (1984), Angular momentum at null infinity, *Class. Quant. Grav.*, 1(1), 15–26.

E.Poisson, An advanced course on General Relativity (n.d.).

Goldberger, W. D., Rothstein, I. Z. (2006), An Effective field theory of gravity for extended objects, *Phys. Rev. D*, 73, 104029.

GraceDB (n.d.).

Holz, D. E., Hughes, S. A. (2005), Using gravitational-wave standard sirens, *Astrophys. J.*, 629, 15–22.

Iacovelli, F., Mancarella, M., Foffa, S., Maggiore, M. (2022), Forecasting the Detection Capabilities of Third-generation Gravitational-wave Detectors Using GWFAST, *Astrophys. J.*, 941(2), 208.

Iyer, V., Wald, R. M. (1994), Some properties of Noether charge and a proposal for dynamical black hole entropy, *Phys. Rev. D*, 50, 846–864.

Kokkotas, K. D., Schmidt, B. G. (1999), Quasinormal modes of stars and black holes, *Living Rev. Rel.*, 2, 2.

Maggiore, M. (2007), *Gravitational Waves. Vol. 1: Theory and Experiments*, Oxford University Press.

Maggiore, M. (2018), *Gravitational Waves. Vol. 2: Astrophysics and Cosmology*, Oxford University Press.

Maggiore, M., Iacovelli, F., Belgacem, E., Mancarella, M., Muttoni, N. (2024), Comparison of global networks of third-generation gravitational-wave detectors, , .

Newman, E. T., Penrose, R. (1968), New conservation laws for zero rest-mass fields in asymptotically flat space-time, *Proc. Roy. Soc. Lond. A*, 305, 175–204.

Poisson, E., Will, C. M. (2014), *Gravity: Newtonian, Post-Newtonian, Relativistic*, Cambridge University Press.

Reardon, D. J. *et al.* (2023), Search for an Isotropic Gravitational-wave Background with the Parkes Pulsar Timing Array, *Astrophys. J. Lett.*, 951(1), L6.

Reitze, D. *et al.* (2019), Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO, *Bull. Am. Astron. Soc.*, 51(7), 035.

Sachs, R. K. (1962), Gravitational waves in general relativity. 8. Waves in asymptotically flat space-times, *Proc. Roy. Soc. Lond. A*, 270, 103–126.

Samajdar, A., Janquart, J., Van Den Broeck, C., Dietrich, T. (2021), Biases in parameter estimation from overlapping gravitational-wave signals in the third-generation detector era, *Phys. Rev. D*, 104(4), 044003.

Santoni, L. (2024), 'Aspects of black hole perturbation theory'.

Sathyaprakash, B. *et al.* (2012), Scientific Objectives of Einstein Telescope, *Class. Quant. Grav.*, 29, 124013. [Erratum: Class.Quant.Grav. 30, 079501 (2013)].

Schutz, B. F. (1986), Determining the Hubble Constant from Gravitational Wave Observations, *Nature*, 323, 310–311.

Thorne, K. S. (1980), Multipole Expansions of Gravitational Radiation, *Rev. Mod. Phys.*, 52, 299–339.

Wald, R. M. (1984), *General relativity*, Chicago Univ. Press, Chicago, IL.
URL: *https://cds.cern.ch/record/106274*

Wald, R. M., Zoupas, A. (2000), A General definition of 'conserved quantities' in general relativity and other theories of gravity, *Phys. Rev. D*, 61, 084027.

Weber, J. (1960), Detection and generation of gravitational waves, *Physical Review*, 117(1), 306–313.

Will, C. M. (2014), The Confrontation between General Relativity and Experiment, *Living Rev. Rel.*, 17, 4.

Xu, H. *et al.* (2023), Searching for the Nano-Hertz Stochastic Gravitational Wave Background with the Chinese Pulsar Timing Array Data Release I, *Res. Astron. Astrophys.*, 23(7), 075024.