2 The field-theoretical approach to GWs

In the previous chapter we investigated GWs using the geometric interpretation which is at the core of general relativity. This geometrical perspective emphasizes that $g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}$ is the metric of space-time, and therefore an incoming GW $h_{\mu\nu}$ induces perturbations in the space-time geometry. In this approach, the interaction of GWs with test masses is described by geometric tools such as the equation of the geodesic deviation, and the energy–momentum tensor of GWs is determined by examining how $h_{\mu\nu}$ contributes to the curvature of the background space-time.

General relativity can also be seen as a classical field theory, to which we can apply all standard field-theoretical methods. In this chapter we therefore go back to linearized gravity, writing $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, and we treat it as a classical field theory of the field $h_{\mu\nu}$ living in *flat* space-time with Minkowski metric $\eta_{\mu\nu}$. In this approach we are actually forgetting that $h_{\mu\nu}$ has an interpretation in terms of a space-time metric, and instead we treat it as any other field living in flat Minkowski space.

The fact that the beautiful geometric interpretation of $h_{\mu\nu}$ is hidden is compensated by the fact that we can put the conceptual issues discussed in Chapter 1 into the broader theoretical framework of classical field theory, and compare it to what happens in other field theories, such as classical electrodynamics. The geometric and the field-theoretical perspectives are indeed complementary; some aspects of GW physics can be better understood from the former perspective, some from the latter, and to study GWs from both vantage points results in a deeper overall understanding.

We will begin in Section 2.1.1 by recalling a basic tools of classical field theory, the Noether theorem, and we will see how we can reobtain the results of Chapter 1 using this standard field-theoretical tool. Besides providing a complementary understanding of various conceptual issues, the Noether theorem is also a very practical tool for explicit computations, and in particular we will see that it also provides the simplest way of deriving another important result, namely the general expression for the angular momentum carried by GWs.

In Section 2.2 we will pursue this field-theoretical approach further, discussing linearized gravity from the point of view of quantum field theory, and we will see how the notion of the graviton emerges. Actually, all astrophysical mechanisms of GW production, as well as the

interaction of GWs with a detector, are fully accounted for by *classical* general relativity. In actual calculations, the notion of the graviton will not surface until we examine some cosmological production mechanisms (in particular, mechanisms related to the amplification of vacuum fluctuations), in Vol. 2. Nevertheless, we will see that also at the quantum level, the field-theoretical approach is illuminating for many conceptual aspects.

We will conclude this chapter with a more advanced section, which investigates whether gravitons can have a small mass. This, from the field-theoretical point of view, seems to be one of the most natural generalizations of Einstein gravity. We will see however that a field theory describing massive gravitons can have problems of internal consistency which, to date, are not yet fully understood.

2.1 Linearized gravity as a classical field theory

2.1.1 Noether's theorem

We begin by recalling some basic facts of classical field theory and in particular Noether's theorem.[1] We consider a field theory living in flat space-time, with fields ϕ_i, labeled by an index i. The action S is the integral of the Lagrangian density $\mathcal{L}$,

$$S = \int dtd^3x \; \mathcal{L}(\phi_i, \partial\phi_i) \,. \tag{2.1}$$

In our case the fields ϕ_i will be the independent components of the metric $h_{\mu\nu}$, but it is useful to be more general, since we will also be interested in comparing with for instance classical electrodynamics, or any other classical field theory. We will denote collectively the fields ϕ_i simply by ϕ.

A transformation of the coordinates and of the fields is an operation that transforms the coordinates x^μ into new coordinates x'^μ, i.e. $x^\mu \to x'^\mu$, while at the same time the fields, denoted collectively by $\phi(x)$, are transformed into new functions of the new coordinates, $\phi(x) \to \phi'(x')$. To define the transformation means to state how x' is related to x and how $\phi'(x')$ is related to $\phi(x)$. For an infinitesimal transformation, we can write

$$x^\mu \to x'^\mu = x^\mu + \epsilon^a A^\mu_a(x) \,, \tag{2.2}$$
$$\phi_i(x) \to \phi'_i(x') = \phi_i(x) + \epsilon^a F_{i,a}(\phi, \partial\phi) \,, \tag{2.3}$$

and the transformation is specified assigning $A^\mu_a(x)$ and $F_{i,a}(\phi,\partial\phi)$. In general, the function $F_{i,a}$ depends on the collection of fields ϕ and on their derivatives, and will also mix different fields, so the transformation of a single field ϕ_i can depend also on all other fields ϕ_j with $j \neq i$. The above transformation is parametrized by a set of infinitesimal parameters ϵ^a, with $a = 1, \ldots, N$.

[1] We refer the reader to the textbook Maggiore (2005), Section 3.2, for further details and proofs. Note that in the present book, in order to follow the most common convention in general relativity, we are using the signature $\eta_{\mu\nu} = (-,+,+,+)$, while in Maggiore (2005) we use $\eta_{\mu\nu} = (+,-,-,-)$, which is the most common convention in quantum field theory. This is the origin of some sign differences between that textbook and the following equations.

Equations (2.2) and (2.3) define a symmetry transformation if they leave the action $S(\phi)$ invariant, for any ϕ. A symmetry transformation is called *global* if it leaves the action invariant when the parameters ϵ^a are constant, and *local* if it leaves invariant the action even when ϵ^a are allowed to be arbitrary functions of x.

Noether's theorem states that, for each generator of a global symmetry transformation, that is, for each of the parameters ϵ^a with $a = 1, \ldots, N$, there is a current j^μ_a (which is a functional of the fields ϕ) which, when evaluated on a classical solution of the equations of motion $\phi^{\rm cl}$, is conserved, i.e. satisfies

$$(\partial_\mu j^\mu_a)\,|_{\phi=\phi^{\rm cl}} = 0\,. \tag{2.4}$$

The corresponding charges Q_a are defined as the spatial integral of the $\mu = 0$ component of the currents

$$Q_a \equiv \int d^3x\, j^0_a(\mathbf{x}, t)\,. \tag{2.5}$$

Current conservation (in the sense of eq. (2.4)) implies that Q_a is conserved (in the sense that it is time-independent). In fact

$$\begin{aligned}\partial_0 Q_a &= \int d^3x\, \partial_0 j^0_a(\mathbf{x}, t)\\ &= -\int d^3x\, \partial_i j^i_a(\mathbf{x}, t)\,.\end{aligned} \tag{2.6}$$

This is the integral of a total divergence, and it vanishes if we assume a sufficiently fast decrease of the fields at infinity. More generally, in a finite volume the variation of the charge is given by a boundary term representing the incoming or outgoing flux.

The explicit form of the current can be written in full generality in terms of the Lagrangian density $\mathcal{L}$ of the theory and of the functions $A^\mu_a(x)$ and $F_{i,a}(\phi, \partial\phi)$ that define the symmetry transformation (2.2), (2.3), and is given by

$$j^\mu_a = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\left[A^\nu_a(x)\partial_\nu\phi_i - F_{i,a}(\phi, \partial\phi)\right] - A^\mu_a(x)\mathcal{L}\,. \tag{2.7}$$

The simplest application of this very general machinery is to the symmetry under space-time translations, and leads us to energy–momentum tensor. Under space-time translations we have $x^\mu \to x'^\mu = x^\mu + \epsilon^\mu$, and by definition all fields transform as $\phi'_i(x') = \phi_i(x)$, independently of their properties under Lorentz transformations; that is, a point P has the coordinate x in a frame and the coordinate x' in the translated frame, but the functional form of the fields changes so that the numerical values of the fields at the point P is invariant. So we write the transformation as

$$\begin{aligned}x^\mu \to x'^\mu &= x^\mu + \epsilon^\mu\\ &= x^\mu + \epsilon^\nu\delta^\mu_\nu\,,\end{aligned} \tag{2.8}$$

$$\phi_i(x) \to \phi'_i(x') = \phi_i(x)\,. \tag{2.9}$$

Observe that the index a appearing in ϵ^a in this case is a Lorentz index. Comparing with eqs. (2.2) and (2.3), we see that $A^\mu_\nu = \delta^\mu_\nu$ and $F_{i,a} = 0$.

The four conserved currents $\theta^\mu{}_\nu \equiv -j^\mu_{(\nu)}$ form a Lorentz tensor, known as the *energy–momentum* tensor. Using eq. (2.7) and raising the ν index, $\theta^{\mu\nu} = \eta^{\nu\rho}\theta^\mu{}_\rho$ we get[2]

$$\theta^{\mu\nu} = -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\partial^\nu\phi_i + \eta^{\mu\nu}\mathcal{L}\,. \tag{2.10}$$

Equation (2.4) becomes

$$\partial_\mu\theta^{\mu\nu} = 0\,, \tag{2.11}$$

on the solutions of the classical equations of motion. The conserved charge associated to space-time translations is, by definition, the four-momentum P^ν, and therefore[3]

$$cP^0 \equiv \int d^3x\,\theta^{00}\,, \tag{2.12}$$

$$cP^i \equiv \int d^3x\,\theta^{0i}\,. \tag{2.13}$$

This is the definition of four-momentum in classical field theory. A field configuration, solution of the equations of motion, carries an energy $E = cP^0$ and a spatial momentum P^i which can be calculated using eqs. (2.10), (2.12) and (2.13).

Observe that in general relativity the energy–momentum tensor of matter, defined by eq. (1.2), is automatically symmetric in the two indices μ, ν, since it is obtained by taking the functional derivative of the action with respect to the symmetric tensor $g_{\mu\nu}$. In contrast, the energy–momentum tensor defined from Noether's theorem, eq. (2.10), is not necessarily symmetric in the two indices μ, ν.

In fact, it is important to understand that the formal machinery of the Noether theorem, without some further physical input, is unable to uniquely fix the energy–momentum tensor, and more generally the Noether currents. For instance, consider what happens if we add a total four-divergence to the Lagrangian density,

$$\mathcal{L}' = \mathcal{L} + \partial_\mu K^\mu(\phi)\,. \tag{2.14}$$

A total divergence, when integrated over d^4x, gives a boundary term. The equations of motion of the classical theory are obtained from a variation of the action, holding fixed the value of the fields on the boundaries. Therefore the equations of motion obtained from the variation of the action $S' = \int d^4x\,\mathcal{L}'$ are the same as those obtained from the variation of $S = \int d^4x\,\mathcal{L}$, so these two Lagrangians define the same classical field theory. However, the currents obtained from eq. (2.7) using $\mathcal{L}'$ or using $\mathcal{L}$ are in general different, and their difference is such that j^0 changes by a total spatial divergence, so that the charge in eq. (2.5) changes by a boundary term. Therefore the Noether *currents* are not uniquely defined; however, the Noether *charges*, computed integrating over a spatial volume V, are well defined if, and only if, the fields inside V go to zero sufficiently fast when we approach the boundaries of V, so that

[2] In order to minimize the number of factors of c in the equations, we have defined the flat-space Lagrangian density $\mathcal{L}$ from

$$S = \int dt d^3x\,\mathcal{L}\,,$$

rather than from $S = \int d^4x\,\mathcal{L}$, as done, instead, in Landau and Lifshitz, Vol. II (1979). Recalling that, dimensionally, an action is (energy)×(time), in this way $\mathcal{L}$ has the same dimensions as $\theta^{\mu\nu}$, i.e. energy/volume. With our notation, for instance, the Lagrangian density of the electromagnetic field is $(-1/4)F_{\mu\nu}F^{\mu\nu}$, rather than $(-1/4c)F_{\mu\nu}F^{\mu\nu}$.

[3] The factor c provides the correct dimensions, recalling that $P^0 = E/c$, see the Notation.

all boundary terms can be neglected, and the ambiguity in the currents becomes irrelevant.

To illustrate this point, before computing the energy–momentum tensor of GWs it is instructive to recall what happens in the more familiar case of classical electrodynamics. In this case the Lagrangian density is

$$\mathcal{L}_{\rm em} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} , \tag{2.15}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. With our signature $\eta_{\mu\nu} = (-,+,+,+)$, the relation of $F_{\mu\nu}$ to the electric and magnetic fields is $F^{0i} = E^i$ and $F^{ij} = \epsilon^{ijk} B^k$. Equation (2.10) gives

$$\theta^{\mu\nu}_{\rm em} = -\frac{\partial \mathcal{L}_{\rm em}}{\partial(\partial_\mu A_\rho)} \partial^\nu A_\rho + \eta^{\mu\nu} \mathcal{L}_{\rm em} . \tag{2.16}$$

The functional derivative is easily computed,

$$\begin{aligned} \frac{\partial}{\partial(\partial_\mu A_\rho)} \left(-\frac{1}{4} F_{\alpha\beta} F^{\alpha\beta} \right) &= -\frac{1}{2} F^{\alpha\beta} \frac{\partial F_{\alpha\beta}}{\partial(\partial_\mu A_\rho)} \\ &= -F^{\alpha\beta} \frac{\partial(\partial_\alpha A_\beta)}{\partial(\partial_\mu A_\rho)} \\ &= -F^{\mu\rho} . \end{aligned} \tag{2.17}$$

Therefore

$$\theta^{\mu\nu}_{\rm em} = F^{\mu\rho} \partial^\nu A_\rho - \frac{1}{4} \eta^{\mu\nu} F^2 . \tag{2.18}$$

At first sight, this result is surprising. Recall in fact that classical electrodynamics is invariant under gauge transformations,

$$A_\mu \to A_\mu - \partial_\mu \theta , \tag{2.19}$$

as we can see observing that under eq. (2.19) $F_{\mu\nu}$ is invariant, and therefore also the Lagrangian (2.15) is invariant. However, the energy–momentum tensor (2.18) depends on A_μ not only through the gauge-invariant combination $F_{\mu\nu}$, but also through the term $\partial^\nu A_\rho$, which is not invariant, and under gauge transformations $\theta^{\mu\nu}_{\rm em}$ changes as

$$\theta^{\mu\nu}_{\rm em} \to \theta^{\mu\nu}_{\rm em} - F^{\mu\rho} \partial^\nu \partial_\rho \theta . \tag{2.20}$$

Apparently, we seem to be driven to the conclusion that the energy density θ^{00} of the electromagnetic field (as well as the momentum density θ^{0i}) is not gauge invariant. To deal with this problem, one first of all rewrites the energy–momentum tensor as follows,

$$\begin{aligned} \theta^{\mu\nu}_{\rm em} &= F^{\mu\rho} (\partial^\nu A_\rho - \partial_\rho A^\nu + \partial_\rho A^\nu) - \frac{1}{4} \eta^{\mu\nu} F^2 \\ &= (F^{\mu\rho} F^\nu{}_\rho - \frac{1}{4} \eta^{\mu\nu} F^2) + F^{\mu\rho} \partial_\rho A^\nu \\ &= (F^{\mu\rho} F^\nu{}_\rho - \frac{1}{4} \eta^{\mu\nu} F^2) + \partial_\rho (F^{\mu\rho} A^\nu) . \end{aligned} \tag{2.21}$$

(In the last line we used the equation of motion $\partial_\rho F^{\mu\rho} = 0$.) Therefore we have

$$\theta^{\mu\nu}_{\rm em} = T^{\mu\nu}_{\rm em} + \partial_\rho C^{\rho\mu\nu} \tag{2.22}$$

where

$$T^{\mu\nu}_{\rm em} = F^{\mu\rho} F^\nu{}_\rho - \frac{1}{4} \eta^{\mu\nu} F^2 , \tag{2.23}$$

while $C^{\rho\mu\nu} = F^{\mu\rho} A^\nu$ is a tensor antisymmetric in the indices (ρ, μ). Now, $T^{\mu\nu}_{\rm em}$ (which is sometimes called the "improved" energy–momentum tensor) is a gauge-invariant quantity, and its 00 component gives the usual form of the energy density,[4]

$$T^{00}_{\rm em}(x) = \frac{1}{2} (\mathbf{E}^2 + \mathbf{B}^2)(x) . \tag{2.24}$$

The term $\partial_\rho C^{\rho\mu\nu}$ instead is not gauge invariant, and we would like to get rid of it. The argument that is used to discard it is based on the following observations:

- If $\theta^{\mu\nu}_{\rm em}$ is conserved, also $T^{\mu\nu}_{\rm em}$ is conserved: in fact
$$\partial_\mu \partial_\rho C^{\rho\mu\nu} = 0 \tag{2.25}$$
automatically, whenever $C^{\rho\mu\nu}$ is antisymmetric under $\rho \leftrightarrow \mu$, as is our case.
- The difference between the charge cP^ν computed with $\theta^{\mu\nu}_{\rm em}$ and that computed with $T^{\mu\nu}_{\rm em}$ is given by
$$\int_V d^3x\, \partial_\rho C^{\rho 0 \nu} = \int_V d^3x\, \partial_i C^{i0\nu} . \tag{2.26}$$
where we used the fact that $C^{00\nu} = 0$, which follows again from the antisymmetry of $C^{\rho\mu\nu}$ under $\rho \leftrightarrow \mu$. This is the integral of a divergence, and vanishes if the fields go to zero sufficiently fast at the boundaries of the volume V. Therefore the four-momentum P^ν computed with $\theta^{\mu\nu}_{\rm em}$ is equal to that computed with $T^{\mu\nu}_{\rm em}$, and is gauge invariant.

What we learn from this example is that the expression for the energy–momentum tensor derived from the Noether theorem, eq. (2.10), is not necessarily a physical observable (in the case of electromagnetism it is not even gauge-invariant!).[5] Rather, it is just a mathematical expression that, when integrated over space, gives unambiguously the total energy and momentum of a classical field configuration, as long as this field configuration goes to zero sufficiently fast on the boundaries of the integration region.

Equivalently, instead of speaking of the total energy of a localized object, we can divide by the volume and say that the expression in eq. (2.10) is a quantity that can be used to compute the *average value* of the energy–momentum tensor over a region of space sufficiently large, so that all boundary terms vanish, and any ambiguity related to total divergences disappears. Then, for instance, from eq. (2.22) we have

$$\langle \theta^{00}_{\rm em} \rangle = \langle T^{00}_{\rm em} \rangle + \langle \partial_i C^{i00} \rangle , \tag{2.27}$$

[4] The use of the Lagrangian (2.15) implies that we are using Heaviside–Lorentz units (also called rationalized c.g.s. unit) for the electric charge; in unrationalized units the factor $(-1/4)$ in eq. (2.15) becomes $(-1/16\pi)$, and the factor 1/2 in eq. (2.24) becomes $1/(8\pi)$.

[5] Another way of understanding the existence of such an ambiguity in a gauge theory is the fact that, in principle, one can allow that the gauge field, under space-time translations, does not go simply into itself, as in eq. (2.9), but into a configuration related by an arbitrary gauge transformation.

where the bracket represents the average. On a volume such that boundary terms give zero, we have $\langle \partial_i C^{i00} \rangle = 0$ and therefore the average is unambiguously defined,

$$\begin{aligned} \langle \theta^{00}_{\rm em} \rangle &= \langle T^{00}_{\rm em} \rangle \\ &= \frac{1}{2} \langle \mathbf{E}^2 + \mathbf{B}^2 \rangle \, . \end{aligned} \tag{2.28}$$

Whether one of the many equivalent integrands, in the expression for cP^0, can be promoted to a physical observable, thereby providing a definition of a *local* energy density, is a physical question that cannot be answered using only the mathematics of the Noether theorem, without any additional physical input. We will discuss this issue at the end of the next section, and we will see that in fact the answer is in general different in electromagnetism and in general relativity.

2.1.2 The energy–momentum tensor of GWs

Now we can return to our original problem, which was the computation of the energy carried by GWs. We consider a wave-packet with reduced wavelengths centered around a value $\lambda\!\!\!^{-}$. In this case, according to the discussion above, the Noether theorem can give us an unambiguous answer for the energy density of the wave-packet, averaged over a box centered on the peak of the wave-packet, and with size $L \gg \lambda\!\!\!^{-}$. In this case the field is negligible on the boundaries and, using eq. (2.10),

$$t^{\mu\nu} = \langle - \frac{\partial \mathcal{L}}{\partial(\partial_\mu h_{\alpha\beta})} \partial^\nu h_{\alpha\beta} + \eta^{\mu\nu} \mathcal{L} \rangle \, , \tag{2.29}$$

where $\langle \ldots \rangle$ is a spatial average over several reduced wavelength (which, for plane waves, is the same as a temporal average over several periods), and $\mathcal{L}$ is the Lagrangian that governs the dynamics of $h_{\mu\nu}$.

As discussed in the previous subsection, the Noether theorem instead gives us an ambiguous answer if we ask what is the *local* energy and momentum density. Actually, we already saw in Section 1.4 that the final form of the energy–momentum tensor of GWs is indeed expressed as an average over several reduced wavelengths (or over several periods), and that this comes from a very fundamental reason, i.e. in order to discuss the back-reaction of GWs on the background, we need to perform a coarse-graining of the Einstein equations. Thus, we already know that it will not be possible to do better than this, and we cannot define a local expression for the energy and momentum density. Nevertheless, it is interesting to understand the reason also from a purely field-theoretical point of view; this will be discussed at the end of this section. First, we compute $t^{\mu\nu}$ from eq. (2.29), and we check that it agrees with the result that we derived in Section 1.4.

In order to use eq. (2.29) (and also to derive the angular momentum of GWs from Noether's theorem), we need the Lagrangian $\mathcal{L}$ or, equivalently, the action governing the dynamics of the field $h_{\mu\nu}$. To reproduce the Einstein equations to linear order in $h_{\mu\nu}$ we must expand the Einstein action to quadratic order in $h_{\mu\nu}$, while the linear term in the action vanishes, as always when we expand around a classical solution (in this case around the flat metric $\eta_{\mu\nu}$, since we consider Einstein equations in vacuum). We therefore start from the full Einstein action,

$$S_E = \frac{c^3}{16\pi G} \int d^4x \sqrt{-g}\, R \, , \tag{2.30}$$

and we expand $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. We observe that

$$R = g^{\mu\nu} R_{\mu\nu} = \left(\eta^{\mu\nu} - h^{\mu\nu} + O(h^2) \right) \left(R^{(1)}_{\mu\nu} + R^{(2)}_{\mu\nu} + O(h^3) \right) \, , \tag{2.31}$$

where $R^{(1)}_{\mu\nu}$ is linear in h and $R^{(2)}_{\mu\nu}$ is quadratic in h. $R^{(1)}_{\mu\nu}$ and $R^{(2)}_{\mu\nu}$ can be obtained by specializing eqs. (1.113) and (1.114) to a flat background metric $\eta_{\mu\nu}$. The expansion of $\sqrt{-g}$ can be computed by writing $g^\mu_\nu = \delta^\mu_\nu + h^\mu_\nu \equiv (I + H)^\mu_\nu$, where I is the identity matrix and H a matrix whose elements are h^μ_ν. Since $g_{\mu\nu} = \eta_{\mu\rho} g^\rho_\nu$ and $\det \eta_{\mu\rho} = -1$, we have $-g = \det(I + H)$. Using the identity $\log(\det A) = \mathrm{Tr}(\log A)$, valid for any non-degenerate matrix A, and expanding the logarithm,

$$\begin{aligned} \det(I + H) &= \exp\{\log \det(I + H)\} \\ &= \exp\{\mathrm{Tr} \log(I + H)\} \\ &= \exp\{\mathrm{Tr}[H + O(H^2)]\} \\ &= 1 + \mathrm{Tr} H + O(H^2) \\ &= 1 + h + O(h^2_{\mu\nu}) \, , \end{aligned} \tag{2.32}$$

where $h = h^\mu_\mu = \eta^{\mu\nu} h_{\mu\nu}$ is the trace of $h_{\mu\nu}$. Since we are expanding over flat space, the lowest non-zero term in R is already $O(h)$, see eq. (2.31), so the terms $O(h^2)$ in $\sqrt{-g}$ give a contribution to the action $O(h^3)$, and can be neglected. Performing straightforward algebra we then obtain, after some integration by parts,

$$\begin{aligned} S_E = -\frac{c^3}{64\pi G} \int d^4x \;\; & \left[\partial_\mu h_{\alpha\beta} \partial^\mu h^{\alpha\beta} - \partial_\mu h \partial^\mu h \right. \\ & \left. + 2 \partial_\mu h^{\mu\nu} \partial_\nu h - 2 \partial_\mu h^{\mu\nu} \partial_\rho h^\rho_\nu \right] \end{aligned} \tag{2.33}$$

and the corresponding Lagrangian density is (see Note 2)

$$\mathcal{L} = -\frac{c^4}{64\pi G} \left[\partial_\mu h_{\alpha\beta} \partial^\mu h^{\alpha\beta} - \partial_\mu h \partial^\mu h + 2 \partial_\mu h^{\mu\nu} \partial_\nu h - 2 \partial_\mu h^{\mu\nu} \partial_\rho h^\rho_\nu \right] . \tag{2.34}$$

We can now compute $t^{\mu\nu}$, using eq. (2.29). We evaluate the result directly in the gauge

$$\partial^\mu h_{\mu\nu} = 0, \qquad h = 0 \, . \tag{2.35}$$

Therefore, *after* computing the derivative $\delta \mathcal{L} / (\partial_\mu h_{\alpha\beta})$ in eq. (2.29), we impose the gauge condition (2.35). We observe that the second, third and fourth terms in brackets in eq. (2.33) are quadratic in quantities that will be set to zero by the gauge fixing and, after taking the functional derivative, they give contributions which are linear in h or in $\partial^\mu h_{\mu\nu}$, and therefore vanish when we impose eq. (2.35). So the only non-vanishing contribution comes from the term $\partial_\mu h_{\alpha\beta}\partial^\mu h^{\alpha\beta}$, and we get

$$\left.\frac{\partial\mathcal{L}}{\partial(\partial_\mu h_{\alpha\beta})}\right|_{\partial^\mu h_{\mu\nu}=h=0} = -\frac{c^4}{32\pi G}\,\partial^\mu h^{\alpha\beta}\,. \tag{2.36}$$

We next evaluate the term $\langle\mathcal{L}\rangle$ in eq. (2.29), in our gauge. We recall that inside the average we are free to perform integration by parts (compare with Note 23 on page 35). Then, since in our gauge $h_{\mu\nu}$ satisfies the equations of motion $\Box h_{\mu\nu} = 0$, even the term $\partial_\mu h_{\alpha\beta}\partial^\mu h^{\alpha\beta}$ gives zero because, after an integration by parts, it becomes $-h_{\alpha\beta}\Box h^{\alpha\beta}$ and vanishes upon use of the equations of motion, so $\langle\mathcal{L}\rangle = 0$.

In conclusion, we obtain

$$\boxed{t^{\mu\nu} = \frac{c^4}{32\pi G}\,\langle\partial^\mu h^{\alpha\beta}\partial^\nu h_{\alpha\beta}\rangle\,.} \tag{2.37}$$

As expected, the result agrees with eq. (1.133).

We can now come back to the problem of the localization of the energy of GWs. In the field-theoretical description, the issue is whether one of the many equivalent integrands, in the expression for P^0, can be promoted to a physical observable, thereby providing a definition of a *local* energy density. In electromagnetism there is a very natural candidate, the tensor $T^{\mu\nu}_{\rm em}$ of eq. (2.23), which is conserved and gauge invariant, so it can be sensible to identify $(1/2)(\mathbf{E}^2+\mathbf{B}^2)$ with the *local* energy density of the electromagnetic field. No such answer is possible for the gravitational field. The quantity $\partial^\mu h^{\alpha\beta}\partial^\nu h_{\alpha\beta}$ which appears inside the average in eq. (2.37) is not gauge-invariant. It is futile to search for another local expression, whose integral gives the energy, but which is already gauge invariant before integration: the equivalence principle tells us that, at a given point, we can always find a locally inertial frame (see Section 1.3.2), such that at the point in question the gravitational field vanishes. Therefore any candidate expression for a local energy density can always be set to zero at a given point with a coordinate transformation, so it cannot be gauge invariant. This is an important difference between gravity and electromagnetism. In electromagnetism, at least if we consider a slowly varying electromagnetic field, it makes sense to assign to each point a local energy density $(1/2)(\mathbf{E}^2+\mathbf{B}^2)$.

When we consider waves, however, concerning energy localization there is no real difference between gravity and electromagnetism. In both cases, all that can be really measured is the energy averaged over several wavelengths or periods. This can be understood even more easily at first by looking at the problem quantum-mechanically even if, as we will see below, the argument is really classical.

From the quantum point of view, a plane wave describes a collection of massless quanta (gravitons with helicity $h=\pm 2$ for gravitation, and

photons with helicity ± 1 in electrodynamics). Consider a collection of such free particles. To determine the energy of this system in a volume V we must know how many quanta of the field are within V at a given time, and the energy of each. If we take a volume with sides smaller than the reduced wavelength $\lambdabar$ of a photon or of a graviton, in order to know whether a given photon (or graviton) is inside the box we must measure its position with an error $\Delta x < \lambdabar$ and then, by the uncertainty principle, we have $\Delta p > \hbar/\lambdabar$, which is larger than the momentum $p = \hbar/\lambdabar$ of a quantum with reduced wavelength $\lambdabar$. Thus, we have completely lost information about the momentum $\mathbf{p}$ of the particle, and therefore (given that we are considering free particles described by simple plane waves), we lost information about its energy $E = c|\mathbf{p}|$, which means that we cannot localize the energy density better than to a few wavelengths. Alternatively, we can localize the energy in space if we delocalize it in time, according to $\Delta E\Delta t \gtrsim \hbar$. Clearly photons and gravitons do not show any difference in this respect.[6] In fact, even if we phrased it in a quantum language, this argument is really classical, and follows from the fact that the position x of the peak of a classical wave-packet, and its typical wave-vector k satisfy $\Delta x\Delta k \geqslant 1$, and in this form is simply a property of the Fourier transform. In quantum theory, one identifies the momentum $\mathbf{p}$ with $\hbar\mathbf{k}$, and $\Delta x\Delta k \geqslant 1$ becomes $\Delta x\Delta p \geqslant \hbar$.

In conclusion the field-theoretical approach, based on Noether's theorem, gives us an unambiguous recipe for computing a spatial (or a temporal) average of the energy–momentum tensor of GWs. The fact that, for GWs, we cannot do better than this, i.e. that the energy of GWs cannot be localized in space (or in time) with a precision better than a few wavelengths (or a few periods) can be understood as a consequence of the equivalence principle, but it is in fact also a general property of any wave governed by a massless wave equation, which at the quantum level translates into a limitation required by the uncertainty principle.

[6] The fact that, as a quantum field theory, general relativity is not renormalizable plays no role here. As long as we are at energies much smaller than the Planck mass, linearized theory can be promoted to a well defined effective quantum field theory, describing weakly interacting gravitons, as we discuss in more detail in Section 2.2.

2.1.3 The angular momentum of GWs

We next compute the angular momentum carried by GWs. Angular momentum is the conserved charge associated to invariance under spatial rotations. A symmetric tensor $h_{\mu\nu}$, from the point of view of spatial rotations, decomposes into h_{00} and the spatial trace h^i_i, which are both scalars under rotation and therefore are spin-0 fields, h_{0i} which is a spatial vector and therefore has spin 1, and a traceless symmetric tensor h_{ij}, which is a spin-2 field and has five degrees of freedom. To describe the GW we go to the TT gauge, so we have $h_{0\mu}=0$ and we are left only with the field $h^{\rm TT}_{ij}$, which satisfies $(h^{\rm TT})^i_i = 0$ and $\partial^i h^{\rm TT}_{ij} = 0$, compare with eq. (1.31). Observe that $h^{\rm TT}_{ij}$ has only two degrees of freedom, corresponding to a massless particle with helicity ± 2.[7] As before, the last three terms in eq. (2.33) give a contribution to the Noether current that vanishes when, after taking the functional derivatives, we impose the gauge fixing condition. Then, for the purpose of computing the Noether current in this gauge, we can use as Lagrangian the first term

[7] As we will recall in Section 2.2.2 and in Problem 2.1, the physical representation of the Poincaré group are of two types: massive representations, characterized by their spin j and having $2j+1$ states, and massless representations, which have a quantum number j but only two states, corresponding to helicities $h=\pm j$. We will use the name "spin-j field" to denote generically a field that can describe either a massive particle with spin j or a massless particle with helicities $h=\pm j$ (in which case, it will be subject to constraints that eliminate the extra states, see Section 2.2). The name "spin-j particle" will however be reserved to a massive particle with spin j and therefore $2j+1$ degrees of freedom. A massless particle with quantum number j will be referred to as "a massless particles with helicity $\pm j$".

in eq. (2.33), and we can keep only the physical degrees of freedom $h_{ij}^{\rm TT}$,

$$\mathcal{L} = -\frac{c^4}{64\pi G}\,\partial_\mu h_{ij}^{\rm TT}\partial^\mu h_{ij}^{\rm TT}\,. \tag{2.38}$$

A rotation of three-dimensional space is described by a 3×3 orthogonal matrix $\mathcal{R}$, which transforms the coordinates according to $x^i \to \mathcal{R}^{ij}x^j$ (observe that, since our signature is $(-,+,+,+)$, we do not need to be careful about raising or lowering spatial indices). An infinitesimal rotation can be written as

$$\mathcal{R}^{ij} = \delta^{ij} + \omega^{ij}\,, \tag{2.39}$$

and the condition that $\mathcal{R}$ is an orthogonal matrix gives $\omega^{ij} = -\omega^{ji}$. Therefore rotations can be parametrized by the three independent quantities ω^{12}, ω^{13} and ω^{23}, which play the role of the parameters ϵ^a in eqs. (2.2) and (2.3). So, for rotations eqs. (2.2) and (2.3) become

$$x^i \to x^i + \sum_{k<l}\omega^{kl}A^i_{kl}\,, \tag{2.40}$$

$$h_{ij}^{\rm TT} \to h_{ij}^{\rm TT} + \sum_{k<l}\omega^{kl}F_{ij,kl}\,, \tag{2.41}$$

and the sum goes over the independent parameters ω^{kl} with $k<l$. Observe that, in $F_{ij,kl}$, the first pair of indices, (i,j), is the label that identifies the field $h_{ij}^{\rm TT}$, and therefore plays the role that the index i had in eq. (2.3), while the second pair (k,l) plays the role of the index a in eq. (2.3). In particular, $F_{ij,kl}$ is symmetric with respect to the pair (i,j) and can be taken to be antisymmetric with respect to the pair (k,l).

The total angular momentum carried by the GWs is then given by

$$J^i = \frac{1}{2}\epsilon^{ijk}J_{kl}\,. \tag{2.42}$$

where J_{kl} is the conserved charge associated to rotations in the (k,l) plane,

$$J_{kl} = \frac{1}{c}\int d^3x\, j^0_{kl}\,, \tag{2.43}$$

(the factor $1/c$ provides the correct dimensions) and Noether's theorem gives

$$j^0_{kl} = \frac{\partial\mathcal{L}}{\partial(\partial_0 h_{ij}^{\rm TT})}\left[A^\nu_{kl}\partial_\nu h_{ij}^{\rm TT} - F_{ij,kl}\right] - A^0_{kl}\mathcal{L}\,. \tag{2.44}$$

To find the explicit form of A^ν_{kl} and $F_{ij,kl}$ we compare the generic formulas (2.40) and (2.41) with the actual transformation properties of $x^i$and $h_{ij}^{\rm TT}$ under infinitesimal rotations; A^i_{kl} is easily computed observing that

$$\begin{aligned} x^i \to \mathcal{R}^{ij}x^j &= x^i + \omega^{ij}x^j \\ &= x^i + \sum_{k<l}\omega^{kl}(\delta^{ik}x^l - \delta^{il}x^k)\,, \end{aligned} \tag{2.45}$$

so, comparing with eq. (2.40),

$$A^i_{kl} = \delta^{ik}x^l - \delta^{il}x^k\,. \tag{2.46}$$

As for A^0_{kl}, since time is unchanged under spatial rotations, we have $A^0_{kl} = 0$.

The quantity $F_{ij,kl}$ is determined by the properties of $h_{ij}^{\rm TT}$ under rotations. Since $h_{ij}^{\rm TT}$ is a spatial tensor, it transforms as

$$\begin{aligned} h_{ij}^{\rm TT} &\to R_i{}^k R_j{}^l h_{kl}^{\rm TT} \\ &= h_{ij}^{\rm TT} + \omega_j{}^l h_{il}^{\rm TT} + \omega_i{}^k h_{kj}^{\rm TT} \\ &= h_{ij}^{\rm TT} + \sum_{k<l}\omega^{kl}(\delta_{ik}h_{jl}^{\rm TT} - \delta_{il}h_{jk}^{\rm TT} + \delta_{jk}h_{il}^{\rm TT} - \delta_{jl}h_{ik}^{\rm TT})\,. \end{aligned} \tag{2.47}$$

Comparing eqs. (2.41) and (2.47), we see that

$$F_{ij,kl} = \delta_{ik}h_{jl}^{\rm TT} - \delta_{il}h_{jk}^{\rm TT} + \delta_{jk}h_{il}^{\rm TT} - \delta_{jl}h_{ik}^{\rm TT}\,. \tag{2.48}$$

We can now plug these results into eq. (2.44). We also observe that

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial(\partial_0 h_{ij}^{\rm TT})} &= -\frac{c^4}{32\pi G}\,\partial^0 h_{ij}^{\rm TT} \\ &= +\frac{c^3}{32\pi G}\,\dot{h}_{ij}^{\rm TT}\,, \end{aligned} \tag{2.49}$$

and (renaming the indices $(i,j)\to(a,b)$) we get

$$j^0_{kl} = \frac{c^3}{32\pi G}\left[-\dot{h}_{ab}^{\rm TT}(x^k\partial^l - x^l\partial^k)h_{ab}^{\rm TT} + 2\dot{h}_{ab}^{\rm TT}(\delta_{bl}h_{ak}^{\rm TT} + \delta_{al}h_{bk}^{\rm TT})\right]. \tag{2.50}$$

From eqs. (2.43), (2.44) and (2.42), the total angular momentum of the GW is therefore

$$\boxed{J^i = \frac{c^2}{32\pi G}\int d^3x\left[-\epsilon^{ikl}\dot{h}_{ab}^{\rm TT}x^k\partial^l h_{ab}^{\rm TT} + 2\epsilon^{ikl}h_{ak}^{\rm TT}\dot{h}_{al}^{\rm TT}\right].} \tag{2.51}$$

To understand the physical meaning of the two terms in bracket, it is useful to recall the analogous results for spin-0 and spin-1 fields.[8] For a real scalar field, applying Noether's theorem, one finds that the angular momentum carried by a field configuration ϕ is

$$J^i = -\epsilon^{ikl}\int d^3x\,(\partial_0\phi)\,x^k\partial^l\phi\,. \tag{2.52}$$

We see that this has the same structure as the first term in eq. (2.51) (after rescaling $h_{ab}^{\rm TT}$ by a factor $(32\pi G/c^3)^{1/2}$; as we will see in Section 2.2, this rescaling gives to $h_{ab}^{\rm TT}$ the standard field-theoretical normalization). The physical meaning of this term can be understood observing that, for a real field ϕ satisfying the massless Klein–Gordon equation $\Box\phi = 0$,

[8] For explicit proofs, see Maggiore (2005), Section 3.3.1 for spin-0 fields, and Section 4.3.1 for the electromagnetic field. Pay attention to the fact that this reference uses the opposite metric signature (following the most common convention used in field theory, while here we are following the most common convention used in general relativity), and units $\hbar = c = 1$.

one can define the scalar product between two field configurations ϕ and ϕ',

$$\langle\phi|\phi'\rangle = \frac{i}{2}\int d^3x\, \phi\, \overleftrightarrow{\partial_0}\, \phi' \,, \tag{2.53}$$

where, on any two functions f and g, we define $\overleftrightarrow{\partial_\mu}$ by $f\overleftrightarrow{\partial_\mu}g \equiv f\partial_\mu g - (\partial_\mu f)g$. Since ϕ and ϕ' are functions of t and $\mathbf{x}$, and in eq. (2.53) we integrate over d^3x, the result is a priori still a function of time. However, this scalar product is actually time-independent, if ϕ_1, ϕ_2 are solutions of the Klein–Gordon equation. It is then not so surprising that the conserved charges of the scalar field theory can be expressed as expectation values of suitable operators, with respect to this scalar product.

Consider in particular $\hat{L}^i = -i\epsilon^{ikl}x^k\partial^l$, which is the orbital angular momentum operator (in units of $\hbar$). The expectation value of this operator, with respect to the scalar product (2.53), is

$$\begin{aligned}\langle\phi|\hat{L}^i|\phi\rangle &= \frac{i}{2}\int d^3x\, \left[\phi\hat{L}^i\partial_0\phi - (\partial_0\phi)\hat{L}^i\phi\right] \\ &= \frac{1}{2}\,\epsilon^{ikl}\int d^3x\, \left[\phi\, x^k\partial^l\partial_0\phi - (\partial_0\phi)x^k\partial^l\phi\right] \\ &= -\epsilon^{ikl}\int d^3x\,(\partial_0\phi)x^k\partial^l\phi\,, \end{aligned}\tag{2.54}$$

where, going from the second to the third line, we integrated by parts ∂^l in the first term. Comparing with eq. (2.52) we see that the total angular momentum carried by a scalar field configuration ϕ is equal to the expectation value of the orbital angular momentum operator $\hat{L}^i$, with respect to this scalar product,

$$J^i = \langle\phi|\hat{L}^i|\phi\rangle\,. \tag{2.55}$$

Observe also that, for a scalar field, there is no additional contribution, and the total angular momentum is given by the expectation value of the *orbital* angular momentum operator.

Consider now the spin-1 case. If one computes, using Noether's theorem, the angular momentum carried by the electromagnetic field in the radiation gauge, where the electromagnetic field is transverse, $\partial_i A^i = 0$, and satisfies $\Box A^i = 0$, one finds

$$J^i = \int d^3x\, \left[-\epsilon^{ikl}(\partial_0 A_j)\, x^k\partial^l A_j + \epsilon^{ikl}A_k\partial_0 A_l\right]. \tag{2.56}$$

Again, since the equation of motion is $\Box A_i = 0$, we can define a scalar product

$$\langle A|A'\rangle = \frac{i}{2}\int d^3x\, A_i\, \overleftrightarrow{\partial_0}\, A'_i\,, \tag{2.57}$$

which is conserved on the solutions of the equation of motion. Then, we see that the first term in eq. (2.56) is again the expectation value of $\hat{L}^i = -i\epsilon^{ikl}x^k\partial^l$ with respect to this scalar product, so it is the contribution from the orbital angular momentum. In the derivation based on the Noether theorem, this term is determined by the transformation of the coordinates x^i under spatial rotations, (i.e. it comes from the term proportional to A_a^ν in eq. (2.7)), so it is clear that it has the same structure for all fields, independently of their spin. The second term instead depends on the specific properties of the field under spatial rotations (i.e. it comes from the term proportional to $F_{i,a}$ in eq. (2.7)) and therefore it is the spin part. Defining

$$S^i = \epsilon^{ikl}\int d^3x\, A_k\partial_0 A_l, \tag{2.58}$$

the explicit computation[9] shows that, in second quantization, the circular polarizations states of the photons are eigenvectors of the helicity operator $\mathbf{S}\cdot\hat{\mathbf{p}}$ (where $\hat{\mathbf{p}}$ is the unit vector in the direction of propagation), with eigenvalues ±1, as required for a massless particle described by a vector field.

[9] See Maggiore (2005), pages 98–99.

We can now understand the meaning of the two terms in eq. (2.51). Since $h_{ab}^{\rm TT}$ satisfies $\Box h_{ab}^{\rm TT} = 0$, the conserved scalar product is defined as in the Klein–Gordon or in the electromagnetic case,

$$\langle h|h'\rangle = \frac{i}{2}\int d^3x\, h_{ab}^{\rm TT}\, \overleftrightarrow{\partial_0}\, h'^{\rm TT}_{ab}\,. \tag{2.59}$$

After rescaling $h_{ab}^{\rm TT} \to (32\pi G/c^3)^{1/2}h_{ab}^{\rm TT}$ (which, as we will see in Section 2.2.2, is the rescaling needed to give to $h_{\mu\nu}$ the canonical normalization to the kinetic term in the action) the first term in eq. (2.51) is the expectation value of the orbital angular momentum operator with respect to this scalar product, while the term $2\epsilon^{ikl}h_{ak}^{\rm TT}\dot{h}_{al}^{\rm TT}$ is the spin contribution. The factor of 2 in this term correctly reproduces the fact that the gravitational field has spin two, and therefore gravitons are eigenvectors of the helicity with eigenvalues ±2.

In the previous section we learned that the Noether currents cannot be localized better than a few wavelengths, so the physical density of angular momentum, j^i/c, is the integrand of eq. (2.51), averaged over a few wavelengths,

$$\frac{1}{c}j^i = \frac{c^2}{32\pi G}\langle -\epsilon^{ikl}\dot{h}_{ab}^{\rm TT}x^k\partial^l h_{ab}^{\rm TT} + 2\epsilon^{ikl}\dot{h}_{al}^{\rm TT}h_{ak}^{\rm TT}\rangle\,. \tag{2.60}$$

Consider now a GW propagating outward from a source. At time t we consider a portion of the wave front, at radial distance r from the source, and covering a solid angle $d\Omega$. At time $t+dt$ this portion of the wave front has swept the volume $d^3x = r^2dr d\Omega = r^2(c\,dt)d\Omega$. Since the angular momentum per unit volume is (j^i/c), the angular momentum carried away by the GW is $dJ^i = r^2(c\,dt)d\Omega\,(j^i/c)$. Therefore the rate of angular momentum emission due to GWs is

$$\boxed{\frac{dJ^i}{dt} = \frac{c^3}{32\pi G}\int r^2 d\Omega\,\langle -\epsilon^{ikl}\dot{h}_{ab}^{\rm TT}x^k\partial^l h_{ab}^{\rm TT} + 2\epsilon^{ikl}\dot{h}_{al}^{\rm TT}h_{ak}^{\rm TT}\rangle\,.} \tag{2.61}$$

This section is intended for readers with some knowledge of quantum field theory. In this section and in the next we use units

$$\hbar = c = 1\,.$$

It can be useful to recall how dimensional analysis works in these units. Since $c = 1$, velocities are pure numbers while, dimensionally,

$$\text{energy} = \text{momentum} = \text{mass}\,.$$

Recalling, from the uncertainty principle, that a length times a momentum has dimensions of $\hbar$, in units $\hbar = c = 1$ we also have

$$\text{length} = (\text{momentum})^{-1} = (\text{mass})^{-1}\,.$$

Then the dimensions of any quantity can be expressed as positive or negative powers of mass. For example, and energy density is energy/volume $= (\text{mass})^4$.

2.2 Gravitons

All astrophysical processes which generate GWs, as well as the interaction of GWs with detectors, are adequately described in the framework of a classical field theory of gravitation. Nevertheless, it can be instructive to discuss some conceptual issues using the vantage point provided by modern quantum field theory.

In this section we will see that, at the quantum level, gravity must be mediated by a massless particle with helicity ± 2, the graviton. Its free action, in flat space-time, is fixed by field-theoretical considerations, and reproduces the linearization of the Einstein action. Consistency with gauge-invariance then requires the introduction of non-linear couplings between the graviton, with three-graviton vertices, four-graviton vertices, etc., and a full non-linear structure emerges. Finally, we will briefly mention some of the issues involved in the quantization of gravity and we will discuss why, even in the absence of a full quantum theory of gravitation, general relativity makes sense as an effective field theory below the Planck scale, and the notion of graviton is well defined.

2.2.1 Why a spin-2 field?

In quantum field theory all interactions are mediated by the exchange of bosons, i.e. particles with integer spin. Since a flat space-time background is an excellent approximation in many situations, it make sense to look for a relativistic quantum field theory living in flat space-time, that in the non-relativistic limit reduces to Newtonian gravity. Such a theory should be mediated by a boson which propagates in this flat space-time.

Gravity is the interaction that in the non-relativistic limit couples to the mass. To obtain such a coupling from a *local* quantum field theory, we need to couple the field that mediates gravity to a local quantity (i.e. a quantity which is function of the space-time point x) whose spatial integral, in the non-relativistic limit, becomes the mass.

Mathematics alone is not sufficient to obtain uniquely a consistent theory that gives back Newton's law. We will also use as guiding principle the deep insight of Einstein that in the full theory of gravitation not only the mass is a source for gravitation but, more generally, all forms of energy. In a local field theory, we therefore look for a coupling between the field that mediates gravity and the energy density.

Since, from the point of view of Lorentz transformations, energy density is the (00) component of the energy–momentum tensor, consistency with special relativity requires that we couple our gravitational field to the energy–momentum tensor $T_{\mu\nu}(x)$. As long as we work at linearized level, we can neglect the contribution of the gravitational field itself to $T^{\mu\nu}$, so we can take as $T^{\mu\nu}$ the energy–momentum tensor of the matter fields only. Since we are assuming a flat background, at linearized level $T^{\mu\nu}$ obeys the flat-space conservation law

$$\partial_\mu T^{\mu\nu} = 0\,. \tag{2.62}$$

We will come back later to the inclusion of the non-linear terms.

Spin-0

The simplest possibility that one can examine is that gravity is mediated by a spin-0 boson, described by a scalar field ϕ. A scalar field ϕ carries no Lorentz index, so the only possibility is that it couples to the trace of the energy–momentum tensor, $T = T^\mu_\mu$. Therefore the Lagrangian density which describes the dynamics of this hypothetical scalar gravitational field and its coupling to matter, in linearized theory, must be of the form

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\phi\partial^\mu\phi + \mu^2\phi^2) + g\phi T\,, \tag{2.63}$$

where g is a coupling constant, and μ is the mass of the scalar field (more generally, we could also add a potential $V(\phi)$, that however does not influence the following discussion). In order to see if this theory reduces correctly to Newtonian gravity in the non-relativistic limit, we can compute the potential induced by the exchange of a ϕ boson between two static particles of masses m_1 and m_2. Let us recall that in quantum field theory the static interaction potential $V(\mathbf{x})$ is a derived concept, which makes sense only in the non-relativistic limit. To obtain it, one must compute the $2 \to 2$ scattering amplitude M_{fi} at tree level, taking the initial state equal to the final state, and then the potential is given by[10]

$$V(\mathbf{x}) = -\int \frac{d^3q}{(2\pi)^3} M_{fi}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{x}}\,. \tag{2.64}$$

The energy–momentum tensor $T^{\mu\nu}$ is quadratic in the matter fields, so the vertex $g\phi T$ involves two matter-field lines and one ϕ boson, and to compute M_{fi} we must evaluate a Feynman diagram of the type given in Fig. 2.1 (plus a possible exchange graph for identical particles), with $q = (q^0, \mathbf{q})$, setting $q^0 = 0$ because the potential is obtained in the static limit. In general, the details of the computation depend on the form of $T(x)$, i.e. on the specific type of matter field considered. However, if the matter field is massive, we can take the non-relativistic limit, and in this limit we can treat $T(x)$ as an external classical field. In this case, the interaction vertex of this theory becomes as shown in Fig. 2.2, where the cross denotes the insertion of $\tilde{T}(\mathbf{q})$. Then the diagram of Fig. 2.1 reduces to that shown in Fig. 2.3, and the amplitude M_{fi} is given simply by

$$iM_{fi}(\mathbf{q}) = (-ig)^2 \tilde{T}_1(\mathbf{q}) \tilde{D}(\mathbf{q}) \tilde{T}_2(-\mathbf{q})\,, \tag{2.65}$$

where

$$D(q) = \frac{-i}{q^2 + \mu^2}\,, \tag{2.66}$$

is the propagator of the scalar field ϕ (recall that we are using the signature $\eta_{\mu\nu} = (-,+,+,+)$), and T_1, T_2 are the traces of the energy–momentum tensor for particles of masses m_1 and m_2, respectively. In the static limit, $q^2 = -(q^0)^2 + \mathbf{q}^2 \to \mathbf{q}^2$, and the propagator becomes

$$D(\mathbf{q}) = \frac{-i}{\mathbf{q}^2 + \mu^2}\,. \tag{2.67}$$

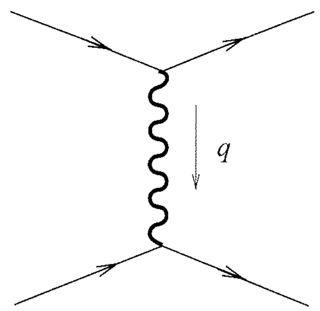

Fig. 2.1 The Feynman diagram that gives the scattering amplitude at tree level.

[10] See e.g. Maggiore (2005) (Section 6.6 and eqs. (7.56)–(7.59)), for the derivation of this result and for explicit computations of the potential induced by the exchange of scalar or vector particles.

Fig. 2.2 The Feynman diagram that gives the vertex, when $T(x)$ is treated as an external field.

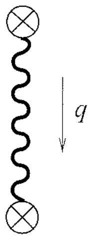

Fig. 2.3 The Feynman diagram that gives the scattering amplitude, when $T(x)$ is treated as an external field.

Equation (2.64) then gives

$$V(\mathbf{x}) = -ig^2 \int \frac{d^3q}{(2\pi)^3} \tilde{T}_1(\mathbf{q})\tilde{D}(\mathbf{q})\tilde{T}_2(-\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{x}} . \tag{2.68}$$

For a relativistic classical particle moving on the trajectory $\mathbf{x}_0(t)$, the energy–momentum tensor is given by (see, e.g. Landau and Lifshitz, Vol. II 1979)

$$T^{\mu\nu}(\mathbf{x},t) = \frac{p^\mu p^\nu}{p^0}\,\delta^{(3)}(\mathbf{x}-\mathbf{x}_0(t)) , \tag{2.69}$$

where p^μ is the four-momentum. Using $p^\mu p_\mu = -m^2$ and, for a particle at rest, $p^0 = m$, the trace of the energy–momentum tensor of a heavy source becomes $T(\mathbf{x}) = -m\,\delta^{(3)}(\mathbf{x})$, which in turn gives $\tilde{T}(\mathbf{q}) = -m$.[11] Therefore

$$V(\mathbf{x}) = -ig^2 m_1 m_2 D(\mathbf{x}) . \tag{2.70}$$

[11] More precisely, for a particle in $\mathbf{x}_1$, $T(\mathbf{x}) = -m\,\delta^{(3)}(\mathbf{x}-\mathbf{x}_1)$, so $\tilde{T}(\mathbf{q}) = -m\exp\{-i\mathbf{q}\cdot\mathbf{x}_1\}$. In the field theoretical language, the factors $\exp\{-i\mathbf{q}\cdot\mathbf{x}_1\}$ from $\tilde{T}_1(\mathbf{q})$ and $\exp\{+i\mathbf{q}\cdot\mathbf{x}_2\}$ from $\tilde{T}_2(-\mathbf{q})$ comes from the wavefunctions of the external legs, and in eq. (2.68) they have already been taken into account by the term $\exp\{i\mathbf{q}\cdot(\mathbf{x}_2-\mathbf{x}_1)\}$, where $\mathbf{x}_2 - \mathbf{x}_2 \equiv \mathbf{x}$.

If the mass μ vanishes,

$$D(\mathbf{x}) = \int \frac{d^3q}{(2\pi)^3}\frac{-i}{\mathbf{q}^2}\,e^{i\mathbf{q}\cdot\mathbf{x}}$$
$$= \frac{-i}{4\pi r} , \tag{2.71}$$

where $r = |\mathbf{x}|$, and therefore we get the correct Newtonian potential

$$V(r) = -\frac{Gm_1m_2}{r} , \tag{2.72}$$

once we make the identification $g^2/(4\pi) = G$. If instead the mass μ is non-vanishing, we get a Yukawa potential

$$V(\mathbf{x}) = -\frac{Gm_1m_2}{r}\,e^{-\mu r} . \tag{2.73}$$

This result shows that, as far as the non-relativistic Newtonian limit is concerned, a spin-0 massless scalar field is a viable possibility.

However, a spin-0 field fails when we come to the new predictions of this theory of gravitation in the relativistic regime. In particular, we see from eq. (2.23) that the energy–momentum tensor of the electromagnetic field is traceless,

$$T_{\rm em} = F^{\mu\rho}F_{\mu\rho} - \frac{1}{4}\delta^\mu_\mu F^2 = 0 . \tag{2.74}$$

Therefore, in this theory, photons do not couple to gravity. Experimentally, the gravitational bending of light rays from massive objects is very well established (it was in fact the first experimental confirmation of general relativity and nowadays, in the form of gravitational lensing, it is a beautiful routine tool in astrophysics). Therefore a spin-0 theory of gravitation is ruled out.[12]

[12] Obviously, this rules out the possibility that the main contribution to gravity comes from spin-0 fields. It says nothing about the possibility that there are additional gravitationally interacting scalar fields, which give small corrections to general relativity either because they are massive or because their gravitational coupling is suppressed with respect to Newton constant G. Indeed, additional scalar fields enter in most extensions of general relativity.

Spin-1

The next possibility is a spin-1 field, just like in electrodynamics. In order to get a long-range potential, we need again a massless field,

but a massless vector field A_μ can be coupled consistently only if we respect gauge invariance. In electrodynamics this can be obtained with a coupling $A_\mu j^\mu$, imposing that j^μ is a conserved current. In fact, under a gauge transformation, $A_\mu \to A_\mu - \partial_\mu\theta$ and, after an integration by parts, the term $-(\partial_\mu\theta)j^\mu \to \theta\partial_\mu j^\mu = 0$, so the action is invariant.

Therefore, a coupling between such a vector field A_μ and the energy–momentum tensor of the form $A_\mu A_\nu T^{\mu\nu}$ is immediately ruled out, because it is not gauge invariant (and furthermore, the simultaneous exchange of two gauge bosons gives a potential $1/r^3$ rather than $1/r$.) A derivative coupling $(\partial_\mu A_\nu)T^{\mu\nu}$ is also not viable, since after integration by parts it gives zero, because of energy–momentum conservation.

If we limit ourselves to the level of first quantization, we could write down a coupling of a vector field to a point-like particles in the form

$$\int d^4x\, A_\mu(x)\, m\frac{dx_0^\mu}{d\tau}\,\delta^{(3)}(\mathbf{x}-\mathbf{x}_0(t)) , \tag{2.75}$$

where $x_0^\mu(\tau)$ is the particle world-line. However, in quantum field theory the four-vector $j^\mu(x)$ that, in the limit of point-like particle, reduces to $m(dx_0^\mu/d\tau)\delta^{(3)}(\mathbf{x}-\mathbf{x}_0(t))$, is a $U(1)$ current. Then $Q \equiv \int d^3x\, j^0$ is equal to the mass m times (number of particles *minus* number of antiparticles), so it is not positive definite, and $\int d^3x\, j^0$ cannot be interpreted as the mass (unless we assign a negative gravitational mass to antiparticles). Furthermore, even if we ignore this problem and interpret j^0 has a mass density, still this attempt to construct a spin-1 theory of gravity fails because, as we know from classical electromagnetism, the interaction mediated by the photon between two particles of the same charge is repulsive. Technically, this comes out because the term $\tilde{T}_1(\mathbf{q})\tilde{D}(\mathbf{q})\tilde{T}_2(-\mathbf{q})$ in eq. (2.68) is now replaced by

$$\tilde{j}^\mu(\mathbf{q})\tilde{D}_{\mu\nu}(\mathbf{q})\tilde{j}^\nu(-\mathbf{q}) , \tag{2.76}$$

where $\tilde{D}_{\mu\nu}(q)$ is the propagator of the massless vector field A_μ. In momentum space,

$$\tilde{D}_{\mu\nu}(q) = \frac{-i}{q^2}\,\eta_{\mu\nu} . \tag{2.77}$$

In the static limit $q^2 = -(q^0)^2 + \mathbf{q}^2 \to \mathbf{q}^2$, and the propagator becomes

$$\tilde{D}_{\mu\nu}(\mathbf{q}) = \frac{-i}{\mathbf{q}^2}\,\eta_{\mu\nu} . \tag{2.78}$$

Because of the factor $\eta_{\mu\nu}$, the propagator of the spatial components A_i is equal to that of a scalar field, but the propagator of A_0 has the opposite sign. In the non-relativistic limit $j^\mu \to (j^0, 0)$. Then in eq. (2.76) only the component D_{00} contributes, so we get the opposite sign compared to the scalar case, i.e. a repulsive potential between positive masses. In conclusion, also spin-1 is ruled out.

Spin $j \geqslant 3$

Values of the spin $j \geqslant 3$ are also ruled out: the need for a long-range force requires again a massless field, which can be coupled consistently only to a conserved tensor. Except possibly for total derivative terms, there is no conserved tensor with three or more indices, so massless particles with $j \geqslant 3$ cannot produce long-range forces, neither gravitational nor of any other type.

The only possibility which is left is $j = 2$, and we examine it in the next section.

2.2.2 The Pauli–Fierz action

The considerations of the previous section lead us to study the action for a spin-2 massless field. To identify the field that describes a massless particle with helicity ± 2, let us first recall some elementary facts about the representations of the Lorentz and Poincaré groups, and their decomposition under representations of the rotation group.

Massless particles in field theory

The irreducible tensor representations of the Lorentz group are given by tensors that, with respect to any pair of indices, are either symmetric and traceless, or antisymmetric. An irreducible representation of the Lorentz group provides of course also a representation of the rotation subgroup $SO(3)$. However, a representation that is irreducible with respect to the full Lorentz group, will be reducible if we limit ourselves to the rotation subgroup (except for the trivial case of the scalar representation), so it decomposes into the direct sum of irreducible representations of the rotation group. For instance a four-vector A_μ is an irreducible representation of the Lorentz group, but from the point of view of rotations it decomposes into a scalar A_0 and a vector $\mathbf{A}$, or

$$A_\mu \in \mathbf{0} \oplus \mathbf{1}\,, \tag{2.79}$$

where we denote by $\oplus$ the direct sum of representations, and by $\mathbf{s}$ the representation of the rotation group corresponding to spin s (so $\mathbf{0}$ is the scalar and $\mathbf{1}$ is the vector representation). The representation $\mathbf{s}$ has dimension $2s+1$, so in particular the scalar is one-dimensional and the vector is a three-dimensional representation. When we consider tensors with two indices, an antisymmetric tensor $A^{\mu\nu}$ decomposes as

$$A_{\mu\nu} \in \mathbf{1} \oplus \mathbf{1}\,, \tag{2.80}$$

while a traceless symmetric tensor $S^{\mu\nu}$ decomposes as

$$S_{\mu\nu} \in \mathbf{0} \oplus \mathbf{1} \oplus \mathbf{2}\,. \tag{2.81}$$

Therefore, the simplest tensor that contains a spin-2 is the traceless symmetric tensor, and a spin-2 can be described using $S_{\mu\nu}$ and imposing conditions that eliminate the extra degrees of freedom. Equation (2.81) states that the nine independent components of a traceless symmetric tensor with two indices decompose into a scalar, the three components of a spin-1, and the five ($2s+1$ with $s = 2$) components of a spin-2.

A further complication arises if we want to describe a *massless* particle. As we recall in more detail in Problem 2.1, particles are representations of the Poincaré group, and the physically interesting representations of the Poincaré group are of two type. (1) Massive representations, with $-P_\mu P^\mu = m^2 > 0$, labeled by the mass m and by the spin j, which takes integer and half-integer values, $j = 0, 1/2, 1, \ldots$. The dimension of these representations is $2j+1$. (2) Massless representations, $P_\mu P^\mu = 0$, which are two-dimensional (actually, one-dimensional, but become two-dimensional if we also require that they are representations of parity) and are characterized by two helicity states $h = \pm j$.

In particular, a massive spin-1 particle has three degrees of freedom, and a massive spin-2 particle has five degrees of freedom. In contrast, a massless particle with $j = 1$ and a massless particle with $j = 2$ both have only two degrees of freedom, the former with helicities $h = \pm 1$, and the latter with $h = \pm 2$. This means that, in the case of a massless spin-1 field, a four-vector field A_μ contains two redundant degrees of freedom, while for a massless spin-2 field the nine components of $S^{\mu\nu}$ contain seven redundant degrees of freedom. The way to eliminate these spurious degrees of freedom is to introduce a gauge-invariance. For electromagnetism, one imposes that the theory is invariant under

$$A_\mu \to A_\mu - \partial_\mu \theta\,. \tag{2.82}$$

It is a standard textbook exercise to show that we can choose $\theta(x)$ so to set $A_0 = 0$. A residual gauge-invariance remains, due to the possibility of performing a further transformation with a function $\theta(\mathbf{x})$ independent of time and (making use of the Maxwell equation in vacuum $\boldsymbol{\nabla}\cdot\mathbf{E} = 0$) it can be used to set $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$, so we eliminate also the longitudinal component of a plane wave solution, and we are left with only two degrees of freedom, at least at the classical level. Then, one can quantize the free theory and verify that one obtains a massless particle with two helicity states, the photon.

The graviton and its action

We want to do the same for a massless particle with $j = 2$, that we call the graviton. The strategy is therefore to start from a traceless symmetric tensor, to impose a local invariance, and to write down a Lagrangian that respects this local symmetry. Then (using the equations of motion derived from this Lagrangian) we can see how many degrees of freedom remain in the theory. In fact, it is even technically simpler to start from a tensor $h_{\mu\nu}$ which is symmetric, but not traceless, so from the point of view of Lorentz symmetry it decomposes into the trace and the symmetric traceless part, and from the point of view of rotations $h_{\mu\nu} \in \mathbf{0} \oplus (\mathbf{0} \oplus \mathbf{1} \oplus \mathbf{2})$. We therefore want to impose an invariance that eliminates eight spurious degrees of freedom.

In order to generalize eq. (2.82) to the case of a field $h_{\mu\nu}$ with two indices, we must assign a Lorentz index to the function that parametrizes the gauge transformation. Let us call $\xi_\mu(x)$ this function. Since we must respect the fact that $h_{\mu\nu}$ is symmetric in (μ,ν), the natural generalization of eq. (2.82) is

$$h_{\mu\nu}(x) \to h_{\mu\nu}(x) - (\partial_\mu \xi_\nu + \partial_\nu \xi_\mu)\,, \tag{2.83}$$

that we still call a gauge transformation. Observe that this is nothing but the symmetry (1.8) of linearized Einstein gravity, so we are on the good track for recovering the linearization of general relativity.

Next, we want to construct a gauge-invariant action, for the free theory. Remarkably, the condition of gauge-invariance fixes this action uniquely. In fact, by inspection we see that the possible terms that one can write down, quadratic in $h_{\mu\nu}$ and with two derivatives, are

$$\partial_\rho h_{\mu\nu}\partial^\rho h^{\mu\nu}\,, \quad \partial_\rho h_{\mu\nu}\partial^\nu h^{\mu\rho}\,, \quad \partial_\nu h^{\mu\nu}\partial^\rho h_{\mu\rho}\,, \quad \partial_\nu h^{\mu\nu}\partial_\mu h\,, \quad \partial^\mu h \partial_\mu h\,, \tag{2.84}$$

where $h = h^\mu_\mu$ is the trace of $h_{\mu\nu}$. Terms of the schematic form $h\partial\partial h$ are related to those written above by a single integration by parts. Furthermore, the second and third term in eq. (2.84) are related by two integrations by parts, that swaps the two derivatives. Therefore, the most general form of the free action is

$$S_2 = \int d^4x\, [a_1 \partial_\rho h_{\mu\nu}\partial^\rho h^{\mu\nu} + a_2 \partial_\rho h_{\mu\nu}\partial^\nu h^{\mu\rho} + a_3 \partial_\nu h^{\mu\nu}\partial_\mu h + a_4 \partial^\mu h \partial_\mu h]\,, \tag{2.85}$$

where the label in S_2 stresses that this quantity is quadratic in $h_{\mu\nu}$. We now impose invariance under the gauge transformation (2.83). This fixes all the coefficients $a_1, \ldots, a_4$, except of course for an overall normalization. We then obtain (choosing the normalization $a_1 = -1/2$; the sign is fixed requiring that the energy is positive definite),

$$S_2 = \frac{1}{2}\int d^4x\, [-\partial_\rho h_{\mu\nu}\partial^\rho h^{\mu\nu} + 2\partial_\rho h_{\mu\nu}\partial^\nu h^{\mu\rho} - 2\partial_\nu h^{\mu\nu}\partial_\mu h + \partial^\mu h \partial_\mu h]\,. \tag{2.86}$$

This is the Pauli–Fierz action. Comparison with eq. (2.33) shows that we have indeed recovered the Einstein action of linearized theory, after a rescaling

$$h_{\mu\nu} \to (32\pi G)^{-1/2} h_{\mu\nu}\,, \tag{2.87}$$

and taking into account that the last term in eq. (2.33) is equal to the second in eq. (2.86) after swapping the two derivatives with integrations by parts. We have therefore found that the linearized Einstein action is the *unique* action that describes a free massless particle with helicities ± 2, propagating in flat space.

We can now repeat the considerations already made in Sections 1.1 and 1.2, see in particular the discussion around eqs. (1.26) and (1.27): we can use the gauge-invariance (2.83) to choose the Lorentz gauge (1.18). This eliminates four of the 10 degrees of freedom in $h_{\mu\nu}$, and still leaves a residual gauge invariance, i.e. the transformations (2.83) with functions ξ_μ that satisfy $\Box \xi_\mu = 0$. In the vacuum, where $T^{\mu\nu} = 0$, using the Lorentz gauge, the equations of motion derived from the linearized Einstein action are $\Box \bar{h}_{\mu\nu} = 0$, and therefore four functions ξ_μ that satisfy $\Box \xi_\mu = 0$ can be used to eliminate four components of $\bar{h}_{\mu\nu}$, so we remain with two degrees of freedom, and we arrive at the TT gauge, eq. (1.31). The requirements of gauge invariance therefore fixes uniquely the linearized action, and leaves us with a massless spin-2 field with two transverse degrees of freedom, the graviton.

For the interaction term, we can write

$$S_{\rm int} = \frac{\kappa}{2}\int d^4x\, h_{\mu\nu}T^{\mu\nu}\,. \tag{2.88}$$

The coupling constant κ will be fixed below. Observe that this interaction term is invariant under the gauge transformation (2.83) because, after an integration by parts, the term $(\partial_\mu\xi_\nu + \partial_\nu\xi_\mu)T^{\mu\nu}$ vanishes, using $\partial_\mu T^{\mu\nu} = 0$.

Just as in electrodynamics, to find the graviton propagator we must add a gauge-fixing term to eq. (2.86), since otherwise the quadratic form is not invertible. The Lorentz gauge can be obtained adding the gauge-fixing term

$$\begin{aligned} S_{\rm gf} &= -\int d^4x\, (\partial^\nu \bar{h}_{\mu\nu})^2 \\ &= -\int d^4x \left(\partial^\nu h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\partial^\nu h\right)^2 \\ &= \int d^4x \left(-\partial_\rho h_{\mu\nu}\partial^\nu h^{\mu\rho} + \partial_\nu h^{\mu\nu}\partial_\mu h - \frac{1}{4}\partial^\mu h\partial_\mu h\right)\,, \end{aligned} \tag{2.89}$$

where in the first term we swapped the derivatives integrating by parts. The overall numerical coefficient in eq. (2.89) has been chosen so that the terms $\partial_\rho h_{\mu\nu}\partial^\nu h^{\mu\rho}$ and $\partial_\nu h^{\mu\nu}\partial_\mu h$ cancel between S_2 and $S_{\rm gf}$, so this choice corresponds to the Feynman gauge in electrodynamics. Putting everything together, we find

$$\boxed{\begin{aligned} S &= S_2 + S_{\rm gf} + S_{\rm int} \\ &= \int d^4x \left[-\frac{1}{2}\partial_\rho h_{\mu\nu}\partial^\rho h^{\mu\nu} + \frac{1}{4}\partial^\mu h\partial_\mu h + \frac{\kappa}{2}h_{\mu\nu}T^{\mu\nu}\right]\,. \end{aligned}} \tag{2.90}$$

The equations of motion obtained performing the variation of this action are

$$\Box \bar{h}_{\mu\nu} = -\frac{\kappa}{2}T_{\mu\nu}\,, \tag{2.91}$$

or, equivalently

$$\Box h_{\mu\nu} = -\frac{\kappa}{2}\left(T_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}T\right)\,. \tag{2.92}$$

Comparing with eq. (1.24) and taking into account the rescaling (2.87), we get (in the units $c=1$ that we are using in this section),

$$\kappa = (32\pi G)^{1/2}\,. \tag{2.93}$$

The graviton propagator

We now find the graviton propagator. Integrating by parts, we have

$$\int d^4x \left[-\frac{1}{2}\partial_\rho h_{\mu\nu}\partial^\rho h^{\mu\nu} + \frac{1}{4}\partial^\mu h \partial_\mu h \right] = \frac{1}{2}\int d^4x\, h^{\mu\nu} A_{\mu\nu\rho\sigma}\partial^2 h^{\rho\sigma}, \tag{2.94}$$

where

$$A_{\mu\nu\rho\sigma} = \frac{1}{2}\left(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho} - \eta_{\mu\nu}\eta_{\rho\sigma}\right). \tag{2.95}$$

The graviton propagator in this gauge is obtained inverting this matrix. Observe that

$$A_{\mu\nu\alpha\beta}A^{\alpha\beta}{}_{\rho\sigma} = \frac{1}{2}(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho}). \tag{2.96}$$

The right-hand side is nothing but the identity matrix, in the space of tensors symmetric in (μ,ν) and in (ρ,σ). Therefore in this space the inverse of A is A itself,[13] so the propagator is given by

$$\boxed{\tilde{D}_{\mu\nu\rho\sigma}(k) = \frac{1}{2}\left(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho} - \eta_{\mu\nu}\eta_{\rho\sigma}\right)\left(\frac{-i}{k^2 - i\epsilon}\right),} \tag{2.97}$$

where, as usual, the $i\epsilon$ prescription selects the Feynman propagator. In particular, $\tilde{D}_{0000}(k) = -i/(2k^2)$ and $D_{0000}(r) = -i/(8\pi r)$.[14] Comparing with eqs. (2.66) and (2.78) we see that the propagator of h_{00} has the same sign as the propagator of the scalar field (and the opposite sign of the propagator of A_0), since $(\eta_{00})^2 = +1$. Therefore, in the static limit h_{00} mediates an *attractive* gravitational potential. Using the interaction terms (2.88) and repeating the same steps performed for the scalar field, we get

$$\begin{aligned} V(\mathbf{x}) &= -i\frac{\kappa^2}{4}\int \frac{d^3q}{(2\pi)^3}\tilde{T}_1^{00}(\mathbf{q})\tilde{D}_{0000}(\mathbf{q})\tilde{T}_2^{00}(-\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{x}} \\ &= -i\frac{\kappa^2}{4}m_1 m_2 D_{0000}(\mathbf{x}) \\ &= -\frac{\kappa^2}{32\pi}\frac{m_1 m_2}{r}. \end{aligned} \tag{2.98}$$

This gives again $\kappa = (32\pi G)^{1/2}$, in agreement with eq. (2.93) (of course, this numerical value depends on the choice of normalization made for $h_{\mu\nu}$), and we have recovered the Newtonian limit.

[13] Alternatively, we can define the 10 fields ϕ_i, $i = 1,\ldots,10$, by $\phi_1 = h_{00}, \phi_2 = h_{01}, \ldots, \phi_{10} = h_{33}$. We then write the kinetic term in the form $A_{ij}\partial_\mu\phi_i\partial^\mu\phi_j$, and we invert the 10×10 matrix A_{ij}.

[14] The overall factor depends of course on the overall normalization of the action, i.e. on the normalization of $h_{\mu\nu}$. With our choice the field h_{00} is not canonically normalized, since its propagator has an extra factor 1/2 with respect to the usual normalization. In contrast, writing $\phi_{\mu\nu} = h_{\mu\nu} - (1/4)\eta_{\mu\nu}h$, the field $\phi_{\mu\nu}$ is canonically normalized. Observe that $\phi_{\mu\nu}$ is a pure spin-2 field since it is traceless, while h is a spin-0 field.

2.2.3 From gravitons to gravity

We can now understand, from a field-theoretical point of view, that the simple action (2.90) cannot be the whole story, and that the correct field theory of gravitation must develop a full non-linear structure. The reason is that, as we have seen, the theory of a massless particle with

helicities ± 2 must necessarily be gauge invariant. We have shown that we can use gauge-invariance to impose the Lorentz condition

$$\partial^\mu \bar{h}_{\mu\nu} = 0, \tag{2.99}$$

and that in this gauge the equations of motion read

$$\Box \bar{h}_{\mu\nu} = -\frac{\kappa}{2}T_{\mu\nu}. \tag{2.100}$$

Equations (2.99) and (2.100) together imply that the energy–momentum tensor of matter satisfies the flat-space conservation law $\partial_\mu T^{\mu\nu} = 0$. In integrated form, this conservation laws reads

$$\frac{d}{dt}\int_V d^3x\, T^{00} = -\int_V d^3x\, \partial_i T^{0i}, \tag{2.101}$$

and states that any change in the energy of the matter field in a volume V is due uniquely to the flux of matter field flowing inside or outside this volume. However, this can be true only as long as we consider $T^{\mu\nu}$ as a given classical external source. As soon as we replace it with the energy–momentum tensor of dynamical matter fields, there will necessarily be an exchange of energy and of momentum between matter and the gravitational field. For instance, there will be gravitational radiation emitted by the matter and going off to infinity, draining energy from the matter sources. Therefore the equation $\partial_\mu T^{\mu\nu} = 0$ is untenable in a full dynamical theory, and eq. (2.100) cannot be exact.

In the full theory what should be conserved is not the energy and momentum of the matter alone, but that of matter plus gravitational field. To remedy this we could try to perform, on the right-hand side of eq. (2.100), the replacement

$$T_{\mu\nu} \to T_{\mu\nu} + t^{(2)}_{\mu\nu}, \tag{2.102}$$

where $t^{(2)}_{\mu\nu}$ is the energy–momentum tensor of the gravitons, obtained from the Pauli–Fierz action using the Noether theorem. Using eq. (2.10), we see that a Lagrangian quadratic in $h_{\mu\nu}$, such as the Pauli–Fierz Lagrangian, produces an energy–momentum tensor quadratic in $h_{\mu\nu}$, of the type, symbolically, $\partial h \partial h$. In eq. (2.102) we added to $t_{\mu\nu}$ the superscript (2) to emphasize that it is quadratic in $h_{\mu\nu}$. Then, we are led to try

$$\Box \bar{h}_{\mu\nu} = -\frac{\kappa}{2}(T_{\mu\nu} + t^{(2)}_{\mu\nu}), \tag{2.103}$$

and the gauge condition $\partial^\mu \bar{h}_{\mu\nu}$ is consistent with the conservation of the graviton plus matter energy–momentum tensor,

$$\partial^\mu(T_{\mu\nu} + t^{(2)}_{\mu\nu}) = 0. \tag{2.104}$$

However, this cannot yet be the end of the story. The equation of motion (2.103) has a term, $(-1/2)\kappa t^{(2)}_{\mu\nu}$, which is quadratic in $h_{\mu\nu}$ (and linear in κ) and to derive it from an action principle, we must add to the action a term *cubic* in $h_{\mu\nu}$ and proportional to κ. Symbolically (i.e. omitting Lorentz indices), $t^{(2)}_{\mu\nu} \sim \partial h \partial h$ and a term of this type in the equations of motion can be obtained adding to the action a term of the form $h\partial h \partial h$, or, restoring Lorentz indices,

$$S_3 = \frac{\kappa}{2} \int d^4x \, h_{\mu\nu} \mathcal{S}^{\mu\nu}(\partial h) \,, \tag{2.105}$$

where $\mathcal{S}^{\mu\nu}(\partial h)$ should be of the general form

$$\mathcal{S}^{\mu\nu}(\partial h) = A^{\mu\nu\rho\sigma\alpha\beta\gamma\delta} \partial_\rho h_{\alpha\beta} \partial_\sigma h_{\gamma\delta} \,, \tag{2.106}$$

and the tensor $A^{\mu\nu\rho\sigma\alpha\beta\gamma\delta}$ is a product of flat metric factors. At this stage, the action of the matter plus gravitational field must therefore be of the general form

$$\begin{aligned} S &= S_2 + S_{\rm gf} + S_{\rm int} + S_3 \\ &= \int d^4x \left[-\frac{1}{2} \partial_\rho h_{\mu\nu} \partial^\rho h^{\mu\nu} + \frac{1}{4} \partial^\mu h \partial_\mu h + \frac{\kappa}{2} h_{\mu\nu} T^{\mu\nu} + \frac{\kappa}{2} h_{\mu\nu} \mathcal{S}^{\mu\nu}(\partial h) \right] . \end{aligned} \tag{2.107}$$

Observe that $\mathcal{S}^{\mu\nu}(\partial h)$ is *not* equal to the energy–momentum tensor $t^{(2)}_{\mu\nu}$. In fact, the variation of a term $h^{\mu\nu} t^{(2)}_{\mu\nu}(h)$ in the action would produce, in the equations of motion, a term

$$\frac{\delta}{\delta h^{\mu\nu}} \left[h^{\alpha\beta} t^{(2)}_{\alpha\beta}(h) \right] = t^{(2)}_{\mu\nu}(h) + h^{\alpha\beta} \frac{\delta}{\delta h^{\mu\nu}} \left[t^{(2)}_{\alpha\beta}(h) \right] , \tag{2.108}$$

so there would be an extra term due to the variation of $t^{(2)}_{\alpha\beta}(h)$. Therefore, in the action $h_{\mu\nu}$ couples to matter through the energy–momentum tensor of matter, but the coupling to itself is through a different tensor. In other words, the equivalence principle means that the energy–momentum tensor of GWs enters on the same footing as the energy–momentum tensor of matter in the equations of motion, not in the action.

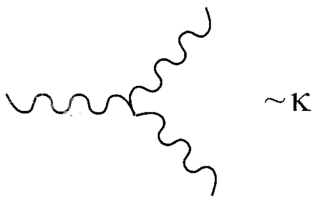

Fig. 2.4 The three-graviton vertex. Two lines carry a contribution proportional to their momentum.

The cubic term in the action means that gravitons have a non-linear coupling to themselves and, in the language of Feynman graphs, corresponds to a vertex as shown in Fig. 2.4. It is instructive to compare this situation with electrodynamics and with Yang–Mills theories. In electrodynamics, the photon mediates the interaction, but carries no electric charge. Therefore it does not couple directly to itself, and it does not contribute to the electric current. In Yang–Mills theories, instead, the gauge bosons are charged with respect to the gauge group, and therefore they couple non-linearly to themselves, and in the theory there are three-boson vertices and four-boson vertices. The situation for gravitons is analogous. Here the role of the current is played by the energy–momentum tensor, and the gravitons couple to their own energy–momentum tensor (in the equations of motion; or to the corresponding functional $\mathcal{S}^{\mu\nu}(\partial h)$ in the action). We see that, with the inclusion of the cubic term in the action, our gauge theory of a massless particle with helicities ± 2 begins to resemble to a non-Abelian gauge theory.

This suspicion is confirmed observing that the action (2.107) is no longer invariant under the linear gauge transformation (2.83) that was our starting point, since the term $h_{\mu\nu} T^{\mu\nu}$ is invariant only if $\partial_\mu T^{\mu\nu} = 0$, which is no longer true. Rather, under the transformation (2.83),

$$\begin{aligned} \delta \int d^4x \, \frac{\kappa}{2} h_{\mu\nu} T^{\mu\nu} &= \kappa \int d^4x \, (\partial_\mu \xi_\nu) T^{\mu\nu} \\ &= -\kappa \int d^4x \, \xi_\nu \, \partial_\mu T^{\mu\nu} \\ &= +\kappa \int d^4x \, \xi_\nu \, \partial_\mu t^{\mu\nu(2)} \\ &= -\kappa \int d^4x \, (\partial_\mu \xi_\nu) \, t^{\mu\nu(2)} \,. \end{aligned} \tag{2.109}$$

However, we have seen that the presence of a local gauge invariance is crucial to eliminate the spurious degrees of freedom in $h_{\mu\nu}$, and we cannot afford to lose it at higher orders in κ. To remedy this, one can observe that this extra term can be canceled, at $O(\kappa)$, promoting the linear gauge transformation (2.83) to a non-linear transformation of the generic form

$$h_{\mu\nu} \to h_{\mu\nu} - (\partial_\mu \xi_\nu + \partial_\nu \xi_\mu) + \kappa O(h \partial \xi) \,. \tag{2.110}$$

Then the transformation of the graviton kinetic term produces a term $O(\kappa)$, and it is possible to choose the tensorial structure of the $O(h\partial\xi)$ term so that it cancels the extra term coming from eq. (2.109). Thus, eq. (2.83) is the gauge symmetry of the theory only at the infinitesimal level, and at the finite level a non-linear gauge transformation emerges, just as in non-Abelian gauge theories. With the hindsight coming from the fact that we know already that the full theory of gravitation is general relativity, we recognize in eq. (2.110) the expansion up to next-to-leading order of a finite diffeomorphism.

It is clear that the iteration procedure does not stop here, neither in the action nor in the gauge transformation. Once we add to the action the term S_3, which is cubic in $h_{\mu\nu}$ and proportional to κ, this produces (through Noether's theorem) a contribution to the energy–momentum tensor of the graviton again cubic in $h_{\mu\nu}$ and proportional to κ. We find useful to display the powers of κ explicitly, so we denote it by $\kappa t^{(3)}_{\mu\nu}$. Noether theorem, applied to the action (2.107), now gives

$$\partial^\mu (T_{\mu\nu} + t^{(2)}_{\mu\nu} + \kappa t^{(3)}_{\mu\nu}) = 0 \,. \tag{2.111}$$

Therefore, consistency with $\partial^\mu \bar{h}_{\mu\nu} = 0$ now requires to replace eq. (2.103) by

$$\Box \bar{h}_{\mu\nu} = -\frac{\kappa}{2} (T_{\mu\nu} + t^{(2)}_{\mu\nu} + \kappa \, t^{(3)}_{\mu\nu}) \,. \tag{2.112}$$

In turn, the action which generates this equation of motion, that has a term cubic in $h_{\mu\nu}$ and proportional to κ^2, must have an additional term quartic in $h_{\mu\nu}$ and proportional to κ^2, so also its associated energy–

momentum tensor has a further term quartic in the fields and proportional to κ^2, that we denote by $\kappa^2 t^{(4)}_{\mu\nu}$, and then

$$\Box \bar{h}_{\mu\nu} = -\frac{\kappa}{2}(T_{\mu\nu} + t^{(2)}_{\mu\nu} + \kappa\, t^{(3)}_{\mu\nu} + \kappa^2 t^{(4)}_{\mu\nu} + \ldots)\,, \tag{2.113}$$

where the dots indicate that the iteration continues to all orders. We recognize the full non-linear structure typical of Einstein gravity, with arbitrarily large powers of $h_{\mu\nu}$, and a non-linear gauge invariance. Indeed, there exists a simple and explicit resummation algorithm, due to Deser, which gives back the Einstein equations. This algorithm uses the first order Palatini formalism, where the action becomes a cubic polynomial in the variable $\sqrt{-g}\, g_{\mu\nu}$ and in the Christoffel symbol (which in a first order formalism are varied independently), and the iteration stops at finite order, see the Further Reading section. So, general relativity can be inferred "bottom up", i.e. starting from gravitons. However, some aspects of this reconstruction procedure are not unique. In particular the full Einstein action includes a boundary term, since it can be written as

$$\begin{aligned} S_E &= \frac{1}{16\pi G}\int d^4x\, \sqrt{-g}\, R \\ &= \frac{1}{16\pi G}\int d^4x\, \left[\sqrt{-g}\, \mathcal{L}_2 - \partial_\mu K^\mu\right]\,, \end{aligned} \tag{2.114}$$

where

$$\mathcal{L}_2 = g^{\mu\nu}\left(\Gamma^\alpha_{\beta\mu}\Gamma^\beta_{\alpha\nu} - \Gamma^\alpha_{\mu\nu}\Gamma^\beta_{\alpha\beta}\right) \tag{2.115}$$

is quadratic in the first derivative of the metric (it is usually called the 'ΓΓ' Lagrangian), and

$$K^\mu = \sqrt{-g}\,(g^{\mu\nu}\Gamma^\alpha_{\alpha\nu} - g^{\alpha\beta}\Gamma^\mu_{\alpha\beta})\,. \tag{2.116}$$

The latter term is a total divergence and therefore does not affect the equations of motion.[15] However, boundary terms are relevant in the quantum theory, and in particular the boundary term in eq. (2.114) becomes physically relevant in semiclassical quantum gravity, in connection with black hole thermodynamics. It is clear that boundary terms are beyond the reach of the iterative procedure that starts from the Pauli–Fierz action, since at each stage of the procedure we have ambiguities connected with the possibility of dropping boundary terms. For instance, from the very beginning, we might have chosen to retain the third term in eq. (2.84) rather than the second term. More formally, expanding $g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}$, $\mathcal{L}_2$ has the structure (we drop all Lorentz indices, since we are only interested in the dependence on κ)

$$\begin{aligned} \frac{1}{32\pi G}\int d^4x\, \sqrt{-g}\, \mathcal{L}_2 &\sim \frac{1}{\kappa^2}\int d^4x\, \left[\kappa^2(\partial h)^2 + \kappa^3 h\partial h\partial h + \ldots\right] \\ &\sim \int d^4x\, \left[(\partial h)^2 + \kappa h\partial h\partial h + O(\kappa^2)\right]\,, \end{aligned} \tag{2.117}$$

[15] Nor the general covariance of the theory. Since boundary terms in the action do not affect classical physics, a classical theory is invariant under a given transformation even if its action changes by a boundary term. This typically happens, for instance, in supersymmetric theories, where under a supersymmetry transformation the action changes by a boundary term.

and therefore is analytic in κ. On the contrary,

$$\begin{aligned} \frac{1}{32\pi G}\int d^4x\, \partial_\mu K^\mu &\sim \frac{1}{\kappa^2}\int d^4x\, \partial^2(\eta + \kappa h) \\ &\sim \frac{1}{\kappa}\int d^4x\, \partial^2 h\,. \end{aligned} \tag{2.118}$$

Therefore this term is non-analytic in κ and cannot be obtained unambiguously from the resummation of an expansion in powers of κ, without further physical input.

Alternatively we can proceed top-down, starting from Einstein action and expanding it in powers of $h_{\mu\nu}$, obtaining all non-linear interaction terms. Observe that in this section we have given to $h_{\mu\nu}$ the dimensions of mass that are canonical in field theory, and the dimensionless metric $h_{\mu\nu}$ is recovered with the rescaling (2.87). Therefore, the expansion that we are performing in this section can be written as

$$\begin{aligned} g_{\mu\nu} &= \eta_{\mu\nu} + (32\pi G)^{1/2} h_{\mu\nu} \\ &= \eta_{\mu\nu} + \kappa h_{\mu\nu}\,, \end{aligned} \tag{2.119}$$

where $h_{\mu\nu}$ has the canonical dimension of mass and κ, dimensionally, is the inverse of a mass. Given the cubic, quartic, and higher terms in the action, one can read the corresponding vertices and compute scattering amplitudes using the Feynman rules. For instance, the $2 \to 2$ graviton scattering amplitude is obtained from the s-channel graph of Fig. 2.5 (together with the corresponding u-channel and t-channel graphs), and from the four-graviton vertex of Fig. 2.6, so this amplitude is $O(\kappa^2)$, i.e. $O(G)$.[16]

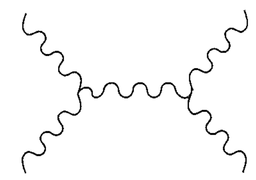

Fig. 2.5 A diagram contributing to the $2 \to 2$ graviton scattering amplitude.

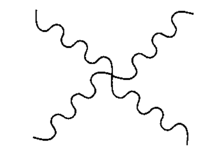

Fig. 2.6 The four-graviton vertex.

[16] The explicit expression for the three-graviton vertex, and the (discouragingly long) four-graviton vertex can be found in eqs. (2.6) and (2.7) of DeWitt (1967). To compute a $2 \to 2$ graviton scattering amplitude in the Born approximation it is however sufficient to know the three-graviton vertex with two gravitons on-shell and only one off-shell, while in the four-graviton vertex all lines are on-shell. Then in the on-shell lines we can impose $\partial^\nu h_{\mu\nu} = 0$ and $h = 0$, and the vertices becomes more manageable, see Grisaru, van Nieuwenhuizen and Wu (1975).

2.2.4 Effective field theories and the Planck scale

In units $\hbar = c = 1$, the Newton constant has the dimensions of the inverse of a mass squared, and therefore $\kappa \sim G^{1/2}$ is the inverse of a mass. The fact that the coupling constant κ has dimensions of the inverse of mass means that the quantization of gravity should give rise to a non-renormalizable theory. This is expected first of all on the basis of simple dimensional considerations. For instance, we have seen that the three-graviton vertex has the generic form $\kappa h\partial h\partial h$. The fact that there are two derivatives is dictated by dimensional analysis. In fact, the dimensions of any term in the Lagrangian density $\mathcal{L}$ must be $(\text{mass})^4$, so that the action $\int d^4x\, \mathcal{L}$ is dimensionless (recall that in this section we are using units $\hbar = c = 1$, and then in particular a length has the dimensions of $(\text{mass})^{-1}$, see Note 2). The kinetic term is, symbolically, $(\partial h)^2$, and since $\partial_\mu = 1/\text{length} = \text{mass}$, we see that $h_{\mu\nu}$ must have dimensions of mass.

The three-graviton vertex therefore must necessarily carry two powers of the momentum: in fact $\kappa hhh \sim (\text{mass})^2$, and to get something trilinear in h and proportional to κ we must add two derivatives, so we must have something of the form $\kappa h\partial h\partial h$. In momentum space each derivative gives a power of momentum, and we can see, writing explicitly some Feynman graphs, that with more and more insertions of this vertex we get Feynman graphs which (barring cancellations) are more and more divergent. Consider for instance the insertion of an internal graviton line in a Feynman graph, obtained adding two three-graviton vertices on two external lines, and connecting them by a graviton propagator, as in Fig. 2.7. Denote by q^μ the momentum of this internal line. Adding this internal line we have created a new closed loop, which gives an integration over d^4q. There are three propagators in the loop, and at large q each of them is $\sim 1/q^2$. Finally, each of the two vertices gives a factor $q_\mu q_\nu$. Then, barring cancellations, the insertion of two vertices brings a further factor that, in the ultraviolet, is of order

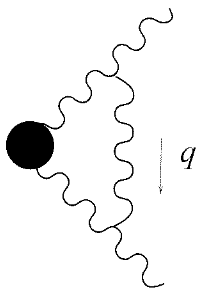

Fig. 2.7 The insertion of an internal graviton line among two external lines. The blob denotes all the remaining (unspecified) part of the Feynman graph.

$$\kappa^2 \int d^4q \, \frac{1}{(q^2)^3} \, q_\mu q_\nu q_\rho q_\sigma \,, \tag{2.120}$$

and so is quadratically divergent. Therefore, adding two vertices we have introduced a new divergence in the graph. Inserting more vertices worsens the situation further and we get stronger and stronger divergencies. The same happens with insertions of the four-graviton vertex, five-graviton vertex, etc.

This lack of renormalizability, however, is not at all a problem as long as we study processes taking place at sufficiently small energies. The problem with non-renormalizable theories is a matter of predictivity, not of mathematical consistency: in the language of counterterms, the divergences are canceled adding to the Lagrangian terms that have a different functional form, compared to the only term $\sqrt{-g}R$ present in Einstein gravity and, as we increase the number of loops, more and more counterterms are required. For instance, to one-loop order one must add to the Lagrangian a term proportional to $R_{\mu\nu}R^{\mu\nu}$ and a term proportional to R^2, to two-loop order we have for instance terms proportional to $R_{\mu\nu\rho\sigma}R^{\rho\sigma\alpha\beta}R_{\alpha\beta}{}^{\mu\nu}$, etc.[17] Therefore the divergences cannot be reabsorbed into a renormalization of the Newton constant and of the fields. Rather, the coefficients of all these terms have a divergent part which is chosen so that it cancels the divergencies coming from the loops, and a finite part that must be fixed by comparison with experiments. Thus, we end up with a theory which, by definition, is finite, but apparently is left with very little predictive power, since we must fix an infinite number of amplitudes by comparison with experiments.

However, this loss of predictivity is not important at low energies. In quantum gravity, we have seen that the coupling constant is $\kappa = (32\pi G)^{1/2}$. Defining the Planck mass $M_{\rm Pl}$ (in our units $\hbar = c = 1$) by

$$G = \frac{1}{M_{\rm Pl}^2} \,, \tag{2.121}$$

we have

$$\kappa = \frac{(32\pi)^{1/2}}{M_{\rm Pl}} \,. \tag{2.122}$$

The perturbative expansion in quantum gravity, which is an expansion in κ^2, is therefore an expansion in powers of $1/M_{\rm Pl}^2$. After renormalizing

[17] More precisely, in the background field method one expands $g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}$, where $\bar{g}_{\mu\nu}$ is a classical solution of the equations of motion and $h_{\mu\nu}$ is a quantum field which, in the path integral formulation, is integrated over. After integrating over the quantum fluctuations $h_{\mu\nu}$, we are left with the counterterms evaluated on $\bar{g}_{\mu\nu}$. For pure gravity, i.e. when we neglect matter fields, $R_{\mu\nu}$ evaluated on any classical solution $\bar{g}_{\mu\nu}$ gives zero, and therefore the one-loop divergences disappear. This miracle does not repeat at higher orders, and does not take place even at one loop if we also have matter fields (in pure supergravity the divergences cancel even at two-loop order, but not to higher orders).

the theory, each scattering amplitude with N external legs, computed up to order $(\kappa^2)^n \sim 1/M_{\rm Pl}^{2n}$, has the generic form

$$A_N(E) = A_N^0(E)\left(1 + c_1\frac{E^2}{M_{\rm Pl}^2} + c_2\frac{E^4}{M_{\rm Pl}^4} + \ldots + c_n\frac{E^{2n}}{M_{\rm Pl}^{2n}}\right) \,, \tag{2.123}$$

For simplicity we assumed that there is only one relevant energy scale E in the amplitude (which in principle depends on all the Lorentz-invariant quantities that one can make with the external momenta). The non-renormalizability of the theory means that, whatever the value of N, we can always find a sufficiently large order n in perturbation theory, where a genuinely new divergence appears, which is not automatically cured by the renormalization of Green's functions with a smaller number of external legs. Therefore, the coefficients $c_1, c_2, \ldots, c_{n-1}$ are finite and calculable once we have renormalized the Green's functions A_M with $M < N$, but the coefficient c_n must be fixed by comparison with experiment, and this is the origin of the lack of predictivity.

However, as long as $E \ll M_{\rm Pl}$, terms suppressed by powers of $E/M_{\rm Pl}$ are completely irrelevant, and the lack of predictivity due to the fact that c_n should be fixed by comparison with experiment is more apparent than real. As long as we study, say, a graviton-graviton scattering process at a center-of-mass energy $E \ll M_{\rm Pl}$, all higher-loop corrections are totally insignificant. Given the huge value of the Planck mass,

$$M_{\rm Pl} \simeq 1.2 \times 10^{19}\,{\rm GeV} \,, \tag{2.124}$$

this means that the lowest-order effective action, i.e. classical general relativity, is completely adequate in all "normal" situations. To see the effect of higher-order terms in the perturbative expansion, we should for instance perform a scattering experiment at center-of-mass energies of order $M_{\rm Pl}$, or, equivalently, we should probe the structure of space-time at length-scale of the order of the Planck length, $l_{\rm Pl} = 1/M_{\rm Pl} \sim 10^{-33}$ cm. In accelerator physics, this is beyond any conceivable future development. Possibly, our best chances of exploring this extremely energetic region is through the observation of some relic of the Big Bang, and in particular stochastic backgrounds of relic GWs, as we will discuss in Vol. 2. However, as long as we consider "normal" astrophysical situations, the lack of predictivity at the Planck scale becomes irrelevant, and general relativity is a totally adequate low-energy effective field theory. The fact that the expansion (2.123) blows up as E approaches $M_{\rm Pl}$ signals that at this point a new theory must set in.

2.3 Massive gravitons

This section lies outside the main theme of the book, and can be omitted in a first reading. As in the previous section, we use units $\hbar = c = 1$.

In Section 2.2 we introduced the notion of graviton, and we saw that it is described by a massless spin-2 field. From a particle physicist's point of view, one of the most natural extensions of Einstein gravity consists in adding to the graviton a small mass term. However, as we will see in this section, the introduction of a mass term for the graviton turns out to be quite subtle from a field-theoretical point of view. In particular the limit $m_g \to 0$ is very delicate, up to the point that one is led to discuss whether the graviton mass should be *identically* zero. Before entering into these considerations, however, we discuss at a simpler phenomenological level the bounds on the graviton mass.

2.3.1 Phenomenological bounds

In general, we expect that a boson with a mass m_g should mediate a short-range force which, compared to the massless case, is suppressed by a factor proportional to $\exp\{-m_g r\}$. In the case of the graviton, such an exponential would cut off the gravitational interaction at a distance r larger than the reduced Compton wavelength $\bar{\lambda}_g = 1/m_g$ (or, reinstating $\hbar$ and c, at distances larger than $\bar{\lambda}_g = \hbar/(m_g c)$).

However, we know experimentally that our Galaxy is held together by gravitation, which means that, at least up to a scale $r_{\rm gal} \sim 10$ kpc, there is no sensible attenuation of the Newton's law, so λ_g cannot be much smaller that 10 kpc. Taking, for definiteness, $\lambda_g > 2$ kpc, this already gives a bound $m_g = 1/\lambda_g < 3 \times 10^{-27}$ eV.

Actually, the experimental bound on the mass of the graviton is even stronger, since we know that the gravitational interaction is not exponentially suppressed even at the intergalactic scale. Our Galaxy has a number of small satellite galaxies, bound by the gravitational force, at distances up to 260 kpc.[18] The Andromeda galaxy, at a distance of 730 kpc, is the nearest giant spiral galaxy, and is approaching the center of mass of our galaxy with a speed $v = -119$ km/s. The most natural explanation is that the relative Hubble expansion between our Galaxy and Andromeda has been halted and reversed by the mutual gravitational attraction.[19] This tells us that $\bar{\lambda}_g$ cannot be much smaller than, say, $O(100)$ kpc. On the scale of hundreds of kpc to 1 Mpc, galaxies are seen to be distorted gravitationally by their reciprocal interaction, creating bridges and tails in their shapes. On the scale (1–10) h_0 Mpc, where h_0 is the Hubble expansion rate in units of 100 km s^{-1}Mpc^{-1} (the current values is $h_0 = 0.73 \pm 0.03$), clusters of galaxies are held together by the gravitational attraction, so we can infer that $\bar{\lambda}_g$ cannot be much smaller than a few Mpc. Taking, conservatively,

$$\lambda_g > 300\, h_0 \text{ kpc}\,, \tag{2.125}$$

results in a bound

$$m_g < 2 \times 10^{-29}\, h_0^{-1}\, \text{eV}. \tag{2.126}$$

This bound, of course, refers only to the lightest particle that mediates the gravitational interaction. In some extensions of general relativity, and in particular in theories with extra dimensions, the graviton is accompanied by a whole family of massive excitations (the Kaluza–Klein modes). The bound (2.126) only refers to the lowest lying state, and says nothing about the possibility of further massive gravitons, whose effect vanishes on much shorter length-scales.

[18]See Binney and Tremaine (1994), Table 10.1.

[19]See Binney and Tremaine (1994), Section 10.2.1.

There is however a potential loophole in the above arguments. We saw in eq. (2.98) that the static gravitational potential is determined by D_{0000}, i.e. by the propagator of the component h_{00} of $h_{\mu\nu}$, which is a scalar under spatial rotation. Gravitational waves, instead, are described by $h_{ij}^{\rm TT}$, which is a spin-2 tensor under rotations. It is possible to construct consistent models where Lorentz invariance is broken and the masses of scalar, vector and tensor perturbations are different. In particular, $h_{ij}^{\rm TT}$ can be massive while scalar perturbations (obtained from gauge-invariant combinations of h_{00} and of the trace h) remains massless, see the Further Reading. The bounds that we discussed above really refer to the mass of the scalar perturbations, and is the same as the mass of $h_{ij}^{\rm TT}$ only if Lorentz invariance holds.

A direct bound on the mass of the tensor mode $h_{ij}^{\rm TT}$ can be obtained from binary pulsars. A binary neutron star system looses energy because it radiates GWs, and this changes its orbital period. The remarkable agreement between the prediction of general relativity for the orbital change, and the measured value for the binary pulsar PSR B1913+16, is in fact one of the great experimental triumphs of general relativity, and also constitutes the first experimental confirmation of the emission of GWs, and will be discuss in great detail in Chapter 6.

We can understand qualitatively how a bound on the graviton mass emerges from the study of binary pulsars, as follows. The system emits GWs at frequencies of order of its orbital frequency ω_s.[20] Then, first of all we must have $m_g c^2 < O(\hbar\omega_s)$ or, in the units $\hbar = c = 1$ that we are using in this chapter, $m_g < O(\omega_s)$, otherwise such massive gravitons could not even be emitted. For m/ω_s small, one finds that the correction to the energy radiated in GWs by the source are of order $(m_g/\omega_s)^2$, with respect to the massless case.[21] Since the agreement between theory and experiment is of order of 0.1% (see Chapter 6) we must actually have, in order of magnitude, $(m_g/\omega_s)^2 < O(10^{-3})$. Given that the orbital period of PSR B1913+16 is about 8 hr, we immediately get an order-of-magnitude estimate of the bound, $m_g < O(10^{-20})$ eV. A more quantitative analysis of the orbital decay rate of PSR B1913+16 (and also of another binary system that will be discussed in Chapter 6, PSR B1534+12) including a mass term for the graviton gives[22]

$$m_g < 7.6 \times 10^{-20}\,\text{eV}, \tag{2.127}$$

corresponding to a value of λ_g of the order of the size of the solar system. Similar bounds on m_g would also come from the direct observation of the waveform of inspiraling compact binaries with interferometric detectors.[23]

The issue of the graviton mass has potentially important cosmological implications, e.g. in connection with attempts of modifying Einstein gravity at cosmological distances (which is largely motivated by the problem of dark matter and dark energy). Furthermore, the whole subject is quite interesting from a field-theoretical point of view. We therefore discuss, in the rest of this section, the field-theoretical problems that arise when we attempt to give a mass to the graviton.

[20]As we will compute in Section 4.1.2, for PSR B1913+16 the spectrum of the radiation emitted is actually peaked toward high harmonics of the orbital frequency, because of the large eccentricity of this binary system.

[21]This comes out repeating the same steps that we will do in Section 3.1, using a massive rather than a massless wave equation for $h_{\mu\nu}$. Since a massive Klein–Gordon type equation depends on m_g^2, the first correction to the radiated energy will also be proportional to the second power of m_g.

[22]This analysis has been done by Finn and Sutton (2002) using a mass term for the graviton which actually implies the existence of six degrees of freedom, the five associated to a massive spin-2 graviton plus an additional scalar field, which however is a ghost, i.e. it has the wrong sign of the kinetic energy (see the discussion in Section 2.3.2 below). Different mass terms, and in particular Lorentz-violating mass terms, should however give similar results.

[23]The basic idea is that, for a massive graviton, the dispersion relation is $\hbar\omega = \gamma mc^2$, where $\gamma = (1 - v^2/c^2)^{-1/2}$ is the usual relativistic factor. Inverting this for v, and using units $\hbar = c = 1$, we get

$$v(\omega) = \sqrt{1 - \frac{m^2}{\omega^2}}\,. \tag{2.128}$$

In the inspiral amplitude (that will be studied in Section 4.1) the radiation emitted earlier in the inspiral phase is at lower frequencies, and therefore travels slightly slower than the radiation emitted at later times, resulting in a potentially observable distortion of the waveform observed at the detector, see Will (1998).

2.3.2 Field theory of massive gravitons

A warm-up: massive photons

We want to understand how to construct a consistent field theory for a massive graviton. As a warm-up exercise, let us first recall what happens in electrodynamics if we add by hand a mass term to the photon. The Lagrangian of a massive photon interacting with a conserved current j_μ is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}m_\gamma^2 A_\mu A^\mu - j_\mu A^\mu . \tag{2.129}$$

This is known as the Proca Lagrangian and is not gauge-invariant, because of the mass term.[24] This is as it should be, because in the massless case gauge invariance reduces the four degrees of freedom of A_μ to only two degrees of freedom, the two helicity states of the photon, while we want to describe a massive particle, and a massive particle with spin $j = 1$ has $2j + 1 = 3$ degrees of freedom. Still, in order to describe the three physical degrees of freedom of a massive photon with the four components of A_μ, we need to eliminate one degree of freedom. Of course, we cannot do it by imposing the condition $\partial_\mu A^\mu = 0$ in the above Lagrangian with a gauge-fixing procedure, since there is no gauge symmetry to be fixed. Rather, the condition $\partial_\mu A^\mu = 0$ is recovered as follows. The equations of motion obtained from (2.129) are

$$\partial_\mu F^{\mu\nu} - m_\gamma^2 A^\nu = j^\nu . \tag{2.130}$$

Acting with ∂_ν on both sides we find $\partial_\nu\partial_\mu F^{\mu\nu} = 0$, because $\partial_\nu\partial_\mu$ is symmetric while $F^{\mu\nu}$ is antisymmetric, and $\partial_\nu j^\nu = 0$ because we have coupled A^μ to a conserved current. Then, eq. (2.130) implies

$$m_\gamma^2\, \partial_\nu A^\nu = 0 , \tag{2.131}$$

and, if $m_\gamma \neq 0$, we get the Lorentz condition $\partial_\nu A^\nu = 0$ dynamically, and we have eliminated one degree of freedom. In momentum space this gives $k_\mu \tilde{A}^\mu(k) = 0$. Since for a massive particle the rest frame exists, we can choose the frame where $k^\mu = (m_\gamma, 0, 0, 0)$, so we have eliminated the component A_0, and we remain with the three components of the vector $\mathbf{A}$ that describe a massive spin-1 particle, as it should be.[25]

Still, it can be disturbing to observe that, apparently, the zero mass limit is discontinuous, because the number of physical degrees of freedom seems to change abruptly from three, for $m_\gamma \neq 0$, to two for $m_\gamma = 0$. To understand this point, let us see what are the polarization states of a massive photon and their coupling to the current j^μ. Consider first the propagation of a free massive photon, i.e. eq. (2.130) with $j^\nu = 0$,

$$\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) - m_\gamma^2 A^\nu = 0 . \tag{2.132}$$

Using $\partial_\mu A^\mu = 0$, which follows from eq. (2.131), this becomes a massive wave equation,

$$(\Box - m_\gamma^2)A^\mu = 0 , \tag{2.133}$$

[24] For the purpose of our discussion, it will not be important whether this mass term has been added by hand, or if it results from a Higgs mechanism.

[25] Actually, even if $\partial_\nu j^\nu \neq 0$, we find that A^0 is fixed in terms of $\partial_\nu j^\nu$, so we have anyway eliminated one degree of freedom. Current conservation is however crucial to ensure a smooth limit as $m_\gamma \to 0$, as we will see below.

whose solution is a superposition of plane waves of the form $\epsilon^\mu(k)e^{ikx}$ with $k^2 = -m_\gamma^2$, and of their complex conjugates. The condition (2.131) implies

$$\epsilon_\mu(k)k^\mu = 0 . \tag{2.134}$$

We choose a frame where $k^\mu = (\omega, 0, 0, k_3)$, with $\omega^2 = k_3^2 + m_\gamma^2$. In this frame two solution of eq. (2.134) are given by the transverse vectors

$$\epsilon_\mu^{(1)}(k) = (0,1,0,0) , \quad \epsilon_\mu^{(2)}(k) = (0,0,1,0) , \tag{2.135}$$

which are the same as the usual transverse degrees of freedom of a massless photons. For a massive photon there is however also a third physical solution

$$\epsilon_\mu^{(3)}(k) = \frac{1}{m_\gamma}(-k_3, 0, 0, \omega) . \tag{2.136}$$

All three polarization vectors are normalized so that $\epsilon_\mu\epsilon^\mu = 1$. To understand what happens to the polarization state (2.136) in the limit $m_\gamma \to 0$, we observe that we can rewrite it in terms of $k_\mu = (-\omega, 0, 0, k_3)$, as

$$\epsilon_\mu^{(3)}(k) = \frac{1}{m_\gamma}k_\mu + \frac{\omega - k_3}{m_\gamma}(1,0,0,1) . \tag{2.137}$$

The interaction of this state with the current j^μ is proportional to $\epsilon_\mu^{(3)}(k)\tilde{j}^\mu(k)$. If the current is conserved, $k_\mu\tilde{j}^\mu(k) = 0$ and the first term in eq. (2.137) does not contribute. In the limit $m_\gamma \to 0$, we expand $\omega = (k_3^2 + m_\gamma^2)^{1/2} \simeq k_3 + m_\gamma^2/(2k_3)$, and therefore

$$\begin{aligned}\epsilon_\mu^{(3)}(k)\tilde{j}^\mu(k) &= \frac{\omega - k_3}{m_\gamma}\left[\tilde{j}^0(k) + \tilde{j}^3(k)\right] \\ &\simeq \frac{m_\gamma}{2k_3}\left[\tilde{j}^0(k) + \tilde{j}^3(k)\right] .\end{aligned} \tag{2.138}$$

Therefore, in the massless limit, the longitudinal mode of a massive photon decouples, and the limit $m_\gamma \to 0$ is continuous.[26] Observe that current conservation has been crucial to show the decoupling of the longitudinal state.

The continuity of the limit $m_\gamma \neq 0$, as far as physical observables are concerned, can also be seen from the propagator, as follows. The propagator of the massive photon is found from the quadratic term in the action which, after an integration by parts, reads

$$S = \frac{1}{2}\int d^4x\, A_\mu\left[\eta^{\mu\nu}(\partial^2 - m_\gamma^2) - \partial^\mu\partial^\nu\right]A_\nu . \tag{2.139}$$

In momentum space, $\partial_\mu \to ik_\mu$ and we must therefore invert the matrix

$$M^{\mu\nu}(k) = -\eta^{\mu\nu}(k^2 + m_\gamma^2) + k^\mu k^\nu . \tag{2.140}$$

The inversion is easily performed writing M^{-1} in the general form

$$(M^{-1})_{\mu\nu} = a(k)\eta_{\mu\nu} + b(k)k_\mu k_\nu , \tag{2.141}$$

and fixing $a(k)$ and $b(k)$ from $M^{\mu\rho}(M^{-1})_{\rho\nu} = \delta^\mu_\nu$. This gives $a =$

[26] To illustrate the continuity of the physics as $m_\gamma \to 0$, consider the following example. In a thermal ensemble, each degree of freedom contributes to the thermodynamical properties, such as the internal energy or specific heat. For instance, in a relativistic Bose gas at equilibrium, the average energy density is $\rho = (\pi^2/30)gT^4$, where g is the number of polarization degrees of freedom. Therefore there is a finite difference between the energy density for massless photons ($g = 2$) and massive photons ($g = 3$). Apparently, from a measure of thermal properties of such a gas we should be able to decide whether $m_\gamma = 0$ exactly, or $m_\gamma \neq 0$. However, in the limit $m_\gamma \to 0$, the longitudinal mode takes an infinite time to reach equilibrium with the thermal bath, because its interactions goes to zero. Therefore any experiment lasting a finite time will only be able either to discover a positive mass or to put an upper limit on it, but not to state that $m_\gamma = 0$ *exactly*.

$-1/(k^2+m_\gamma^2)$, $b = a/m_\gamma^2$. The propagator is defined by $\tilde{D}_{\mu\nu}(k) = i(M^{-1})_{\mu\nu}$. Then[27]

$$\tilde{D}_{\mu\nu}(k) = \frac{-i}{k^2+m_\gamma^2}\left(\eta_{\mu\nu} + \frac{k_\mu k_\nu}{m_\gamma^2}\right). \tag{2.142}$$

[27] We do not write explicitly the $i\epsilon$ prescription. In all cases, with our choice of signature, it is obtained replacing $k^2 \to k^2 - i\epsilon$ in the denominators.

The propagator of a massive photon is therefore singular in the limit $m_\gamma \to 0$. However, the singularity disappears from physical observables, because the amplitudes are proportional to

$$\tilde{j}_\mu^*(k)\tilde{D}^{\mu\nu}(k)\tilde{j}_\nu(k). \tag{2.143}$$

Since the current is conserved, $k^\mu \tilde{j}_\mu(k) = 0$, and therefore the terms $k_\mu k_\nu / m_\gamma^2$ in the propagator give zero. We can therefore simply take as propagator of the massive photon the expression

$$\tilde{D}_{\mu\nu}(k) = \frac{-i}{k^2+m_\gamma^2}\eta_{\mu\nu}, \tag{2.144}$$

and in the the limit $m_\gamma \to 0$ we recover smoothly the massless photon propagator (2.77). The whole procedure required that the massive photon is coupled to a *conserved* current j^μ.

Massive gravitons

Now that we have understood how to describe a massive photon, we can come back to our original problem, the construction of a field theory for massive gravitons. We start from the Pauli–Fierz action (2.86), and we add a mass term. The most general Lorentz-invariant mass term that one can add to the Pauli–Fierz action is a combination of the two scalars $h_{\mu\nu}h^{\mu\nu}$ and h^2. Of course both terms break the gauge invariance (2.83) of the massless theory, as it should, since we want to describe the $2j+1 = 5$ degrees of freedom of a massive spin-2 particle, while we have seen that gauge invariance reduces them to only two. Nevertheless, a generic symmetric tensor $h_{\mu\nu}$ has 10 degrees of freedom so we expect that the appropriate number of conditions to reduce it to five emerges dynamically, similarly to what we have found for massive photons.

Of course, for massive gravitons we must start from the Pauli–Fierz action (2.86) *before* gauge-fixing, since now there is no gauge symmetry to fix. Adding the mass term, and writing also the source term, we have

$$S_2 = \frac{1}{2}\int d^4x\, [\, -\partial_\rho h_{\mu\nu}\partial^\rho h^{\mu\nu} + 2\partial_\rho h_{\mu\nu}\partial^\nu h^{\mu\rho} - 2\partial_\nu h^{\mu\nu}\partial_\mu h + \partial^\mu h \partial_\mu h + c_1 h^2 + c_2 h_{\mu\nu}h^{\mu\nu} + \tilde{\kappa}\, h_{\mu\nu}T^{\mu\nu}]. \tag{2.145}$$

We denote the coupling of the massive theory by $\tilde{\kappa}$ since, as we will see below, it need not be the same as the coupling κ introduced in the massless case. We can now obtain the propagator of the massive graviton repeating the same steps performed above for the photon, and we discover that, if we add to the Pauli–Fierz Lagrangian a mass term with

coefficients c_1 and c_2 arbitrary, the resulting propagator in general has ghosts, i.e. poles with the "wrong" sign of the residue, corresponding to degrees of freedom with negative kinetic energy, which generates an instability of the vacuum. The only combination that does not introduces ghosts turns out to be $h_{\mu\nu}h^{\mu\nu} - h^2$, which is called the Pauli–Fierz mass term. Then, in linearized theory we are led to consider, as the action describing massive gravitons,

$$\boxed{S_2 = \frac{1}{2}\int d^4x\, [\, -\partial_\rho h_{\mu\nu}\partial^\rho h^{\mu\nu} + 2\partial_\rho h_{\mu\nu}\partial^\nu h^{\mu\rho} - 2\partial_\nu h^{\mu\nu}\partial_\mu h + \partial^\mu h \partial_\mu h + m_g^2(h^2 - h_{\mu\nu}h^{\mu\nu}) + \tilde{\kappa} h_{\mu\nu}T^{\mu\nu}].} \tag{2.146}$$

Before discussing the propagator obtained from this action, it is instructive to count the number of dynamical degrees of freedom of the theory, to see if they match with the five degrees of freedom of a massive spin-2 particle. We proceed in parallel to what we have done for a massive photon, and we write the equations of motions derived from the action (2.146),

$$\Box h^{\mu\nu} - (\partial^\nu\partial_\rho h^{\mu\rho} + \partial^\mu\partial_\rho h^{\nu\rho}) + \eta^{\mu\nu}\partial_\rho\partial_\sigma h^{\rho\sigma} + \partial^\mu\partial^\nu h - \eta^{\mu\nu}\Box h = -\frac{\tilde{\kappa}}{2}T^{\mu\nu} + m_g^2(h^{\mu\nu} - \eta^{\mu\nu}h). \tag{2.147}$$

Introducing on the left-hand side $\bar{h}_{\mu\nu} = h_{\mu\nu} - (1/2)\eta_{\mu\nu}h$, we can also write this as

$$\Box\bar{h}^{\mu\nu} + \eta^{\mu\nu}\partial^\rho\partial^\sigma\bar{h}_{\rho\sigma} - \partial_\rho\partial^\nu\bar{h}^{\mu\rho} - \partial_\rho\partial^\mu\bar{h}^{\nu\rho} = -\frac{\tilde{\kappa}}{2}T^{\mu\nu} + m_g^2(h^{\mu\nu} - \eta^{\mu\nu}h). \tag{2.148}$$

In this form we immediately see that this equation reduces to eq. (1.17) for $m_g = 0$. We can now apply ∂_μ to both sides. The left-hand side gives zero identically, while on the right-hand side $\partial_\mu T^{\mu\nu} = 0$, since we are working at linear order and then, as we saw in the previous section, $T^{\mu\nu}$ is an external conserved source. Then we get

$$\boxed{m_g^2\,\partial_\mu(h^{\mu\nu} - \eta^{\mu\nu}h) = 0,} \tag{2.149}$$

which is analogous to eq. (2.131). Furthermore, taking the trace of both sides of eq. (2.147), we get

$$2\partial_\nu\partial_\mu(h^{\mu\nu} - \eta^{\mu\nu}h) = -\frac{\tilde{\kappa}}{2}T - 3m_g^2 h. \tag{2.150}$$

When $m_g \neq 0$, the left-hand side vanishes because of eq. (2.149), and we get

$$\boxed{-3m_g^2\, h = \frac{\tilde{\kappa}}{2}T.} \tag{2.151}$$

In particular in the vacuum, where $T_{\mu\nu} = 0$, eq. (2.151) gives $h = 0$ and then eq. (2.149) gives $\partial^\mu h_{\mu\nu} = 0$ or, equivalently (since $h = 0$), $\partial^\mu \bar{h}_{\mu\nu} = 0$. So we have five conditions that reduce the 10 components of the symmetric matrix $h_{\mu\nu}$ to the five components expected for a massive spin-2 particle. The Pauli–Fierz mass term is actually the only combination for which there is no term proportional to $\Box h$ on the left-hand side of eq. (2.151). For all other choices, such a term does appear, so h becomes a propagating degree of freedom, which furthermore turns out to be ghost-like.

The limit $m_g \to 0$

As soon as we switch on the interaction, i.e. a non-vanishing $T_{\mu\nu}$, the limit $m_g \to 0$ becomes quite peculiar. In fact, in the limit $m_g \to 0$, for generic matter field with a non-vanishing trace of the energy–momentum tensor, eq. (2.151) gives $h \to \infty$ rather than $h = 0$ as in the free theory. To understand the physical consequences of this result we observe that, inserting eqs. (2.149) and (2.151) (i.e. $\partial_\mu h^{\mu\nu} = \partial^\nu h$ and $h = -\tilde{\kappa}T/(6m_g^2)$) into eq. (2.147) we get

$$(\Box - m_g^2)h_{\mu\nu} = -\frac{\tilde{\kappa}}{2}\left(T_{\mu\nu} - \frac{1}{3}\eta_{\mu\nu}T + \frac{1}{3m_g^2}\partial_\mu\partial_\nu T\right) \equiv S_{\mu\nu}\,. \tag{2.152}$$

Comparing with eq. (2.92), we see that in the limit $m_g \to 0$ the left-hand side goes smoothly into the massless limit, but on the right-hand side there are two differences: the coefficient of $\eta^{\mu\nu}T$ is $-1/3$ rather than $-1/2$, and there is an apparently singular term $\sim (1/m_g^2)\partial^\mu\partial^\nu T$. Both find their origin in the fact that, according to eq. (2.151), $m_g^2 h$ stays finite as $m_g \to 0$.

To see the effect produced by these differences, we can consider the effective matter–matter interaction, which is given by

$$S_{\text{eff}} = \frac{\tilde{\kappa}}{2}\int d^4x\, h_{\mu\nu}(x)T^{\mu\nu}(x)\,, \tag{2.153}$$

where $h_{\mu\nu}(x)$ is the solution of eq. (2.152),

$$h_{\mu\nu}(x) = \int d^4x'\, \Delta(x-x')S_{\mu\nu}(x')\,, \tag{2.154}$$

and $\Delta(x-x')$ is a Green's function of the massive Klein–Gordon equation, defined by

$$(\Box_x - m_g^2)\Delta(x-x') = \delta^{(4)}(x-x')\,. \tag{2.155}$$

This gives

$$S_{\text{eff}} = \frac{\tilde{\kappa}}{2}\int d^4x\, d^4x'\, T^{\mu\nu}(x)\Delta(x-x')S_{\mu\nu}(x')\,. \tag{2.156}$$

Using the conservation of the energy–momentum tensor we can show that the singular term in $S_{\mu\nu}$ drops from this expression. In fact, integrating twice by parts,

$$\begin{aligned}
&\int d^4x\, d^4x'\, T^{\mu\nu}(x)\Delta(x-x')\frac{\partial}{\partial x'^\mu}\frac{\partial}{\partial x'^\nu}T(x')\\
&= -\int d^4x\, d^4x'\, T^{\mu\nu}(x)\left[\frac{\partial}{\partial x'^\mu}\Delta(x-x')\right]\frac{\partial}{\partial x'^\nu}T(x')\\
&= +\int d^4x\, d^4x'\, T^{\mu\nu}(x)\left[\frac{\partial}{\partial x^\mu}\Delta(x-x')\right]\frac{\partial}{\partial x'^\nu}T(x')\\
&= -\int d^4x\, d^4x'\left[\frac{\partial}{\partial x^\mu}T^{\mu\nu}(x)\right]\Delta(x-x')\frac{\partial}{\partial x'^\nu}T(x')\\
&= 0\,.
\end{aligned} \tag{2.157}$$

However, the fact that the coefficient of $\eta^{\mu\nu}T$ is $-1/3$ rather than $-1/2$ gives a genuine difference in the physical amplitude,

$$\begin{aligned}
S_{\text{eff}} &= \frac{\tilde{\kappa}}{2}\int d^4x d^4x'\, T^{\mu\nu}(x)\Delta(x-x')[T_{\mu\nu}(x') - \frac{1}{3}\eta_{\mu\nu}T(x')]\\
&= \frac{\tilde{\kappa}}{2}\int d^4x d^4x'\, T^{\mu\nu}(x)\Delta_{\mu\nu\rho\sigma}(x-x';m_g)T^{\rho\sigma}(x')
\end{aligned} \tag{2.158}$$

where

$$\Delta_{\mu\nu\rho\sigma}(x-x';m_g) = \Delta(x-x')\left[\frac{1}{2}(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho}) - \frac{1}{3}\eta_{\mu\nu}\eta_{\rho\sigma}\right]. \tag{2.159}$$

The same result, of course, could have been obtained directly inverting the quadratic part of the action to find the propagator, just as we did for the massless graviton and for the massive photons. Repeating steps analogous to those leading to eq. (2.142), we get

$$\tilde{D}_{\mu\nu\rho\sigma}(k;m_g) = \left[\frac{1}{2}(\Pi_{\mu\rho}\Pi_{\nu\sigma} + \Pi_{\mu\sigma}\Pi_{\nu\rho}) - \frac{1}{3}\Pi_{\mu\nu}\Pi_{\rho\sigma}\right]\left(\frac{-i}{k^2 + m_g^2 - i\epsilon}\right), \tag{2.160}$$

where

$$\Pi_{\mu\nu} \equiv \eta_{\mu\nu} + \frac{k_\mu k_\nu}{m_g^2}\,. \tag{2.161}$$

Since the energy–momentum tensor is conserved, $k_\mu \tilde{T}^{\mu\nu}(k) = 0$ and, when contracted with $T^{\mu\nu}(-k)T^{\rho\sigma}(k)$, the above propagator is equivalent to

$$\tilde{D}_{\mu\nu\rho\sigma}(k;m_g) = \left[\frac{1}{2}(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho}) - \frac{1}{3}\eta_{\mu\nu}\eta_{\rho\sigma}\right]\left(\frac{-i}{k^2 + m_g^2 - i\epsilon}\right). \tag{2.162}$$

Comparing eqs. (2.160) or (2.162) with the massless graviton propagator found in eq. (2.97) we see the following. The terms $k_\mu k_\nu/m_g^2$, that are singular in the massless limit, give a vanishing contribution to the physical amplitudes, because $h_{\mu\nu}$ is coupled to a conserved energy–momentum tensor. This is completely analogous to the situation for a massive photon, where the terms $k_\mu k_\nu/m_\gamma^2$ disappeared because the photon is coupled to a conserved current. However, there is difference in the numerical coefficient in front of $\eta_{\mu\nu}\eta_{\rho\sigma}$ and, unlikely the case of the photon, the propagator (2.162) does not reduces to the massless propagator as $m_g \to 0$. Writing

$$\frac{1}{2}(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho}) - \frac{1}{3}\eta_{\mu\nu}\eta_{\rho\sigma} = \frac{1}{2}(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho} - \eta_{\mu\nu}\eta_{\rho\sigma}) + \frac{1}{6}\eta_{\mu\nu}\eta_{\rho\sigma}\,, \tag{2.163}$$

we see that

$$\lim_{m_g\to 0} T^{\mu\nu}(-k)\tilde{D}_{\mu\nu\rho\sigma}(k;m_g)T^{\rho\sigma}(k) = T^{\mu\nu}(-k)\tilde{D}_{\mu\nu\rho\sigma}(k)T^{\rho\sigma}(k) + \frac{1}{6}T(-k)\,\frac{-i}{k^2}\,T(k)\,, \tag{2.164}$$

where $\tilde{D}_{\mu\nu\rho\sigma}(k)$ is the propagator computed directly in the massless theory, eq. (2.97). The second term on the right-hand side corresponds to the exchange of a massless scalar particle, coupled to the trace of the energy–momentum tensor, and it is because of this additional term that the limit $m_g \to 0$ of the massive theory does not reproduce the result obtained in the massless theory. This is usually referred to as the van Dam–Veltman–Zakharov (vDVZ) discontinuity.[28]

[28] Actually, it was discovered independently by Iwasaki (1970), van Dam and Veltman (1970) and Zakharov (1970).

Having identified the problem in the existence of an unexpected scalar degree of freedom, we can now understand what happened. Recall that a massive spin-j particles has $2j+1$ degrees of freedom, while a massless particle with quantum number j has only two degrees of freedom if $j > 0$ (and if we demand that it is also a representation of parity, beside of Poincaré transformations), and one degree of freedom if $j = 0$. For a massive photon, we saw that the three degrees of freedom, in the massless limit, decompose into the two transverse degrees of freedom of a massless photon, which have helicities ± 1, plus the longitudinal component. The latter, in the limit $m_\gamma \to 0$, becomes a scalar particle and decouples, since its coupling becomes proportional to $\partial_\mu j^\mu$, which vanishes. The fact that the coupling is to $\partial_\mu j^\mu$ is dictated by the fact that this is the only scalar that we can make with j^μ and with derivatives, which is linear in the current. The decoupling of the spurious scalar mode is therefore ensured by current conservation.

Similarly, in the limit $m_g \to 0$, the five degrees of freedom of the massive gravition decompose into two states with helicities ± 2, which make up a massless graviton, two states with helicities ± 1, often termed the graviphoton, and a scalar, called the graviscalar. The graviphoton, being a vector, must be coupled to a four-vector made with $T_{\mu\nu}$ and possibly with derivatives, and linear in $T_{\mu\nu}$. The only possibility is $\partial_\mu T^{\mu\nu}$. However, this quantity vanishes and therefore the graviphoton decouples. The graviscalar, on the other hand, can couple to the trace of the energy–momentum tensor. This is in general non-zero, and therefore the graviscalar does not decouple. It is, in fact, responsible for the additional term in eq. (2.164) and therefore for the vDVZ discontinuity.

We can now compare some predictions of the massless and massive theory. Consider first the Newtonian potential in the non-relativistic limit. We found in eq. (2.98) that

$$V(\mathbf{x}) = -i\frac{\kappa^2}{4}\, m_1 m_2 D_{0000}(\mathbf{x})\,. \tag{2.165}$$

In the massless theory, using eq. (2.97), we saw that

$$\tilde{D}_{0000}(k) = \frac{1}{2}\,\frac{-i}{k^2}\,. \tag{2.166}$$

This gives $D_{0000}(r) = -i/(8\pi r)$, so

$$V(\mathbf{x}) = -\frac{\kappa^2}{32\pi}\,\frac{m_1 m_2}{r}\,, \tag{2.167}$$

and we recovered the Newtonian potential setting $\kappa^2 = 32\pi G$. In the massive theory, instead, we see from eq. (2.162) that

$$\tilde{D}_{0000}(k) = \frac{2}{3}\,\frac{-i}{k^2 + m_g^2}\,, \tag{2.168}$$

and we get

$$V(\mathbf{x}) = -\frac{4}{3}\left(\frac{\tilde{\kappa}^2}{32\pi}\right)\frac{m_1 m_2}{r}\,\exp\{-m_g r\}\,. \tag{2.169}$$

The Yukawa potential was of course expected because a massive particle mediates a short-range force. The result however differs also by an overall factor 4/3 from the massless result, and the difference is due to the additional attractive contribution of the graviscalar.

As far as the Newtonian limit is concerned, one can simply reabsorb this difference setting, in the massive theory,

$$\frac{4}{3}\left(\frac{\tilde{\kappa}^2}{32\pi}\right) = G\,. \tag{2.170}$$

Thus, at $m_g r \ll 1$, the correct Newtonian potential is obtained, at the price that the coupling $\tilde{\kappa}^2$ of the massive theory is smaller than the coupling κ^2 of the massless theory by a factor 3/4. The problem however comes when we consider the predictions of the massive theory in the relativistic regime, in particular when we compute the deflection of light by a massive object. In the massless theory, the deflection angle of light skimming the surface of the Sun (and first detected, during a total eclipse, by Eddington in 1919), is

$$\Delta\theta = \frac{4GM_\odot}{R_\odot}\,, \tag{2.171}$$

where $M_\odot$ and $R_\odot$ are the mass and radius of the Sun. Since the energy–momentum tensor of the electromagnetic field is traceless, when we repeat the computation in the massive theory, the additional term proportional to $T(-k)\,(-i/k^2)\,T(k)$ in eq. (2.164) vanishes. This means that the result for the deflection of light in the massive theory, in the limit $m_g \to 0$, is the same as in the massless theory except that, instead of κ^2, we have $\tilde{\kappa}^2$. Since $\tilde{\kappa}^2 = (3/4)\kappa^2$, the prediction for light bending in the limit $m_g \to 0$ of the massive theory is smaller by a factor 3/4 than the prediction of the massless theory, i.e. is $\Delta\theta = 3GM_\odot/R_\odot$. Nowadays the experimental precision on the measurement of the bending of light by the Sun is better than one part in 10^4, and confirms the prediction of the massless theory. Apparently, one then arrives at the amazing conclusion that the mass of the graviton must be *exactly* zero.

A loophole in the above argument was found by Vainshtein, considering the validity of the linearized approximation in the massive theory. To study the expansion in $h_{\mu\nu}$ systematically, we should start from the full Einstein action plus the Pauli–Fierz mass term,

$$S = \frac{1}{16\pi G}\int d^4x\,\sqrt{-g}\left[R + \frac{1}{4}m_g^2\left(h^2 - h_{\mu\nu}h^{\mu\nu} + O(h_{\mu\nu}^3)\right)\right]\,, \quad (2.172)$$

We then expand $g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}$ where $\bar{g}_{\mu\nu}$ is the appropriate background metric and, to have the canonical normalization for $h_{\mu\nu}$, we finally rescale $h_{\mu\nu} \to (32\pi G)^{1/2}h_{\mu\nu} = \kappa h_{\mu\nu}$. Bosonic matter fields, as usual, are coupled replacing in the matter action all ordinary derivative with covariant derivatives. Actually, one can use any form for the mass term that, in the linearized limit, reduces to the Pauli–Fierz mass term (in particular, we can decide to raise and lower the indices in the mass term with $\eta_{\mu\nu}$ or with $g_{\mu\nu}$. The difference shows up only at cubic and higher-order terms, that are not fixed anyway), so we also included the possibility of non-linear corrections to the Pauli–Fierz mass term.

To compute the gravitational scattering by a fixed heavy mass M, we expand around a metric $\bar{g}_{\mu\nu}$ which is a generalization of the Schwarzschild metric generated by the heavy mass M, so it is computed from the action (2.172), which includes the graviton mass. Naively one would expect that, if $m_g \to 0$, this metric goes smoothly into the usual Schwarzschild metric computed in standard general relativity, with massless graviton, for all values of r. However, the explicit computation shows that this is not true. We can search for the metric generated by an heavy mass in the theory (2.172), writing[29]

[29] We follow here the treatment of Deffayet, Dvali, Gabadadze and Vainshtein (2002), to whom we refer for more details.

$$ds^2 = -e^{\nu(\rho)}dt^2 + e^{\sigma(\rho)}d\rho^2 + e^{\mu(\rho)}\rho^2(\delta\theta^2 + \sin^2\theta d\phi^2)\,. \quad (2.173)$$

When the graviton is massless, the function $\mu(\rho)$ can be set to zero using coordinate invariance. In the theory (2.172), however, the reparametrization invariance is broken, and $\mu(\rho)$ must be kept. One then performs the substitutions

$$r \equiv \rho\, e^{\mu/2}\,, \qquad e^{\lambda} \equiv \left(1 + \frac{\rho}{2}\frac{d\mu}{d\rho}\right)^{-2} e^{\sigma-\mu}\,. \quad (2.174)$$

The standard Schwarzschild solution of the massless theory corresponds to

$$\nu(r) = -\lambda(r) = \log\left(1 - \frac{R_S}{r}\right) = -\frac{R_S}{r} - \frac{1}{2}\left(\frac{R_S}{r}\right)^2 + \ldots\,, \quad (2.175)$$

$$\mu(r) = 0\,. \quad (2.176)$$

In the theory with massive graviton one rather finds, up to next-to-leading order in G (therefore in R_S),

$$\nu(r) = -\frac{R_S}{r}\left[1 + O\left(\frac{R_S}{m_g^4 r^5}\right) + \ldots\right]\,, \quad (2.177)$$

$$\lambda(r) = \frac{1}{2}\frac{R_S}{r}\left[1 + O\left(\frac{R_S}{m_g^4 r^5}\right) + \ldots\right]\,, \quad (2.178)$$

$$\mu(r) = -\frac{1}{2}\frac{R_S}{m_g^2 r^3}\left[1 + O\left(\frac{R_S}{m_g^4 r^5}\right) + \ldots\right]\,. \quad (2.179)$$

If we limit ourselves to leading order, we observe that $\lambda(r)$ is smaller by a factor 1/2 compared to the result in the massless theory. This is the origin of the vDVZ discontinuity. However, the surprise comes looking at the corrections, since they blow up if $m_g \to 0$! In other words, in the massive theory the linearized expansion becomes invalid if we send $m_g \to 0$ at fixed r. This does not mean that linearized theory is completely useless: if we define the *Vainshtein radius* R_V,

$$R_V = \left(R_S\lambda_g^4\right)^{1/5}\,, \quad (2.180)$$

where $\lambda_g = 1/m_g$, we see that the corrections are proportional to $(R_V/r)^5$. Therefore linearized theory is valid at $r \gg R_V$. We take $\lambda_g > 200$ kpc, in agreement with eq. (2.125) (with $h_0 \simeq 0.7$), and we consider the scattering of light from the Sun, which has $R_S \sim 3$ km. Then we get $R_V > 40$ pc, i.e. R_V is at least 10^7 times larger than the Earth–Sun distance. Therefore, the Newtonian potential found in eq. (2.169), and the result for the light deflection in the massive theory discussed below eq. (2.169), are simply not applicable at the solar system scale.

On the other hand, in the opposite limit $r \ll R_V$, it is possible to find a consistent expansion of the Schwarzschild solution in powers of m_g that, to lowest order, reproduces the Schwarzschild solution of the massless theory, of the form[30]

[30] The existence of such a solution depends on the specific form of the non-linear corrections to the Pauli–Fierz mass term, see Damour, Kogan and Papazoglou (2003).

$$\nu(r) = -\frac{R_S}{r} + O\left(m_g^2\sqrt{R_S r^3}\right)\,, \quad (2.181)$$

$$\lambda(r) = +\frac{R_S}{r} + O\left(m_g^2\sqrt{R_S r^3}\right)\,, \quad (2.182)$$

$$\mu(r) = \sqrt{\frac{8R_S}{13r}} + O(m_g^2 r^2)\,. \quad (2.183)$$

Therefore, at $r \ll R_V$, there is no mass discontinuity. To prove that there is no mass discontinuity altogether, one should be able to resum the whole perturbative expansion (2.177–2.179), which is valid at $r \gg R_V$, and it is in principle possible that, in the resummed expression, there is no singularity as $m_g \to 0$.[31] If one can show that such a resummed solution, as we approach $r \sim R_V$, matches smoothly the solution (2.181)–(2.183), which is valid at $r \ll R_V$, we have constructed a solution with a smooth limit $m_g \to 0$, valid for all r. Observe also that, since the linearized theory does not apply at the solar system scale, there is no need to require $\tilde{\kappa}^2 = (3/4)\kappa^2$. On the contrary, since at $r \ll R_V$ the expansion in m_g reproduces smoothly the massless limit, we must choose $\tilde{\kappa} = \kappa$, and all the results of the massive theory, from the Newtonian potential to the light deflection, go smoothly into those of the massless theory. At $r \gg R_V$, instead, where linearized theory can be trusted, we get a gravitational potential

$$V(r) = -\frac{4}{3}\frac{Gm_1m_2}{r}\exp\{-m_g r\}\,, \qquad (2.184)$$

since we have fixed $\tilde{\kappa} = \kappa$. In conclusion, in this scenario, inside the Vainsthein radius a tiny graviton mass has negligible effects, while at $r \gg R_V$ the graviscalar becomes effective, and gives a further attractive contribution to the gravitational potential.

To prove that this what actually happens requires to show that the solutions found in the regimes $r \ll R_V$ and $r \gg R_V$ do match. This is a non-trivial problem, because the solution (2.181)–(2.183) is obtained from an expansion in m_g while the solution (2.177)–(2.179) is non-analytic in m_g, since we have seen that it diverges as $m_g \to 0$. Conversely, the solution (2.177)–(2.179) is obtained performing an expansion in G, while the solution (2.181)–(2.183) is non-analytic in G, since $\mu(r) \sim \sqrt{R_S} \sim \sqrt{G}$. Put it differently, the difficulty of the problem is that, as we approach R_V from the large distance region, the graviscalar becomes strongly coupled, and perturbation theory breaks down.

Numerical studies of the inward continuation of asymptotically flat solutions indicate that, for small m_g, they end up in a singularity at finite radius, rather then matching a continuous solution inside the Vainshtein radius. It is possible however that the matching takes place not for asymptotically flat solution but for asymptotically De Sitter solution. This would be physically acceptable, given the experimental evidence for a cosmological constant, and also because for m_g sufficiently small the value of R_V can be larger than the Hubble radius, in which case the form of the solution at $r \gg R_V$ is not physically relevant.

A further complication is that, in curved space, the trace degree of freedom h, which in linearized theory is eliminated through eq. (2.151), becomes dynamical again, so we have six degrees of freedom rather than five, and furthermore it is a ghost. In fact, in the theory with action (2.172), i.e. full Einstein gravity supplemented by a mass term, the linearized equation of motion (2.147) is replaced by

$$G_{\mu\nu} = \frac{\tilde{\kappa}}{4}T_{\mu\nu} - \frac{1}{2}m_g^2[ah_{\mu\nu} + bh\eta_{\mu\nu} + O(h_{\mu\nu}^2)]\,, \qquad (2.185)$$

[31] For example, consider the function $f(\epsilon, x) = e^{-\epsilon/x}$. If we expand it in powers of ϵ at fixed x we get $f(\epsilon, x) = 1-(\epsilon/x)+(1/2)(\epsilon/x)^2+\ldots$. Of course, this expansion is not suitable for studying the limit $x \to 0$, since the various terms are more and more singular. However, if one knows the full resummed expression $e^{-\epsilon/x}$, one realizes that, for $\epsilon > 0$, the limit $x \to 0^+$ exists, and is in this case zero. In our case the role of ϵ is played by R_S and the role of x by m_g.

where $G_{\mu\nu}$ is the full Einstein tensor, rather than its linearization that appears in eq. (2.147), and we also allowed for a generic mass term, including higher-order corrections (the Pauli–Fierz mass term corresponds to $a = -b = 1$). Using the Bianchi identity $D^\mu G_{\mu\nu} = 0$ as well as the covariant conservation of the energy–momentum tensor, we get the four conditions

$$m_g^2 D^\mu[ah_{\mu\nu} + bh\eta_{\mu\nu} + O(h_{\mu\nu}^2)] = 0\,, \qquad (2.186)$$

which replace their linearized version (2.149) and again allow us to eliminate four degrees of freedom. The elimination of these degrees of freedom is therefore a consequence of the Bianchi identity, or, equivalently, of the invariance of the Einstein–Hilbert action under diffeomorfisms. On the other hand, the elimination of h in linearized theory, eq. (2.151), is not ensured by any symmetry, but is a consequence of the fine-tuning in the mass term that leads to the Pauli–Fierz combination. Then, this constraint does not survive in curved space, and h becomes dynamical again and it can be shown to be ghost-like, see the Further Reading section for discussions.

Thus, presently the issue of the consistency of a field theory of massive gravitons is not settled. If, in some form, a continuous solution indeed exists, then the limit $m_g \to 0$ is smooth and it make sense to deform Einstein gravity by adding a small mass term for the graviton, with a value bounded experimentally by (2.126). Otherwise, one should accept the (rather odd) conclusion that the graviton mass must be identically zero. We finally observe that it is possible to add mass terms that break Lorentz invariance (which indeed emerge quite naturally if one breaks spontaneously the diffeomorfism invariance of general relativity), and in this case there are neither ghosts nor the vDVZ discontinuity. Again, we refer the reader to the Further Reading section.

2.4 Solved problems

Problem 2.1. The helicity of gravitons

We have seen that, for a given propagation direction $\hat{\mathbf{n}}$, a GW is described by a 2×2 matrix in the plane orthogonal to $\hat{\mathbf{n}}$, with matrix elements given in terms of the amplitudes h_+ and $h_\times$ of the two polarization. The Lorentz transformations that leave invariant the propagation direction $\hat{\mathbf{n}}$ are the rotations around the $\hat{\mathbf{n}}$ axis and the boosts in the $\hat{\mathbf{n}}$ direction. Under these operations h_+ and $h_\times$ will therefore transform between themselves. In this problem, we compute explicitly the transformation of h_+ and $h_\times$ under rotations around the $\hat{\mathbf{n}}$ axis and under boosts in the $\hat{\mathbf{n}}$ direction.

In general, under a Lorentz transformation $x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu$, a tensor $h_{\mu\nu}$ transforms as

$$h_{\mu\nu}(x) \to h'_{\mu\nu}(x') = \Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma h_{\rho\sigma}(x)\,. \qquad (2.187)$$

Choosing $\hat{\mathbf{n}} = \hat{\mathbf{z}}$, a rotation around the z axis and a boost along z are written respectively as

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\psi & -\sin\psi & 0 \\ 0 & \sin\psi & \cos\psi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{(rotation)}\,, \tag{2.188}$$

and

$$\Lambda = \begin{pmatrix} \cosh\eta & 0 & 0 & -\sinh\eta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\eta & 0 & 0 & \cosh\eta \end{pmatrix} \quad \text{(boost)}\,, \tag{2.189}$$

where Λ denote the matrix whole elements are $\Lambda_\mu{}^\nu$; ψ is the rotation angle and η is related to the velocity v of the boost by $v = \tanh\eta$. Writing eq. (2.187) in the TT gauge we have

$$(h_{ab}^{\rm TT})'(x') = \begin{pmatrix} h'_+ & h'_\times \\ h'_\times & -h'_+ \end{pmatrix}_{ab} e^{ikx}\,, \tag{2.190}$$

where

$$\begin{pmatrix} h'_+ & h'_\times \\ h'_\times & -h'_+ \end{pmatrix}_{ab} = \Lambda_a{}^c\Lambda_b{}^d \begin{pmatrix} h_+ & h_\times \\ h_\times & -h_+ \end{pmatrix}_{cd}\,, \tag{2.191}$$

and a, b take the value 1, 2. Since $kx = k'x'$, where k' is the four-momentum in the new frame, we can also rewrite eq. (2.190) as

$$(h_{ab}^{\rm TT})'(x') = \begin{pmatrix} h'_+ & h'_\times \\ h'_\times & -h'_+ \end{pmatrix}_{ab} e^{ik'x'} \tag{2.192}$$

or, since x' is generic,

$$(h_{ab}^{\rm TT})'(x) = \begin{pmatrix} h'_+ & h'_\times \\ h'_\times & -h'_+ \end{pmatrix}_{ab} e^{ik'x}\,. \tag{2.193}$$

Using eq. (2.191), with $\Lambda_a{}^c$ given by the 2×2 submatrix made by the second and third rows and columns of eq. (2.188), and performing the matrix multiplication, we get

$$\boxed{\begin{aligned} h'_+ &= h_+\cos 2\psi - h_\times \sin 2\psi\,, \\ h'_\times &= h_+ \sin 2\psi + h_\times \cos 2\psi\,. \end{aligned}} \tag{2.194}$$

Under boosts, the matrix $\Lambda_a{}^c$ (i.e. the 2×2 submatrix made by the second and third rows and columns of eq. (2.189)) is just the 2×2 identity matrix, so $h'_+ = h_+$ and $h_\times = h'_\times$. The GW amplitudes h_+ and $h_\times$ are therefore invariant under boosts.

From eq. (2.194) we see that, under rotations around the z axis, the combinations $h_\times \pm ih_+$ transform as

$$(h_\times \pm ih_+) \to e^{\mp 2i\psi}(h_\times \pm ih_+)\,. \tag{2.195}$$

To understand the meaning of this transformation law, we recall some basic results from the theory of representations of the Poincaré group (see, e.g. Maggiore 2005, Chapter 2). The Poincaré group has two types of physically interesting representations:

- massive representation, which are labeled by the mass m, with $-P_\mu P^\mu = m^2 > 0$ (where P^μ is the four-momentum) and by the spin j, which (in units of $\hbar$) can take integer or half-integer values, $j = 0, 1/2, 1, \ldots$. The representation with spin j has dimension $2j+1$. Physically, this follows from the fact that for a massive particle exists the rest frame, and in the rest frame the component along the z axis of the spin, for a particle with spin j, can take the $2j+1$ possible values $j_z = -j, -j+1, \ldots, j$. So, in particular, a massive spin-1 particle has three degrees of freedom and a massive spin-2 particle has five degrees of freedom.
- massless representation, which are characterized by $P_\mu P^\mu = 0$ and by a quantum number j, which again can be integer or half-integer. For massless particles the rest frame does not exist and the previous argument about the existence of $2j+1$ states does not go through. Rather, these representations are one-dimensional, and are characterized by a definite value of the *helicity*, which is defined as the projection of the total angular momentum on the direction of motion,[32]

$$h = \mathbf{J}\cdot\hat{\mathbf{n}}\,. \tag{2.196}$$

The total angular momentum is the sum of orbital and spin angular momenta, $\mathbf{J} = \mathbf{L} + \mathbf{S}$. Of course $\mathbf{L}\cdot\hat{\mathbf{n}} = (\mathbf{x}\times\mathbf{p})\cdot\hat{\mathbf{n}} = 0$ since $\mathbf{p} = |\mathbf{p}|\,\hat{\mathbf{n}}$, and therefore the helicity is equal to the projection of the spin on the direction of motion, $h = \mathbf{S}\cdot\hat{\mathbf{n}}$.

Under a rotation by an angle ψ around the direction of motion a helicity eigenstate $|h\rangle$ transforms as

$$|h\rangle \to e^{ih\psi}|h\rangle\,. \tag{2.197}$$

Massless representations can have either $h = +j$ or $h = -j$. Each of these possibilities, separately, provides a one-dimensional representation of the Poincaré group: the states with $h = \pm j$ do not mix between them under (proper) spatial rotations, boosts nor translations. However, under a parity transformation $\hat{\mathbf{n}}$ changes sign, while the angular momentum is a pseudovector and is unchanged. Therefore the helicity is a pseudoscalar, i.e. it changes sign under parity, $h \to -h$. For this reason, in a theory which conserves parity (like gravity or electromagnetism) it is more convenient to define particles as representations of the Poincaré group *and* of parity, that is, to assemble the two Poincaré representations with $h = \pm j$ and to consider them as two polarization states of the same particle. The photon is then defined as a massless particle with two helicity states $h = \pm 1$, while the graviton is defined as a massless particle with two helicity states $h = \pm 2$. On the contrary, since weak interactions violate parity, in the limit in which the neutrinos can be taken as massless (there is nowadays evidence for a small neutrino mass from oscillation experiments), we reserve the name neutrino to the one-dimensional massless Poincaré representation with $h = -1/2$, while the antineutrino is defined as the representation with $h = +1/2$.[33]

In the light of these results, we see that eq. (2.195), together with the fact that $h_{ij}^{\rm TT}$ satisfies the massless Klein–Gordon equation $\Box h_{ij}^{\rm TT} = 0$, means that the quanta of the gravitational field are massless particles, and the combinations $h_\times \mp ih_+$ are the helicity eigenstates and have helicities ± 2, respectively.

[32] The symbol h is traditionally used for the helicity, and of course it should not be confused with the GW amplitudes h_+ and $h_\times$.

[33] The fact that, for each massless representation with a given helicity, there is a corresponding representation with the opposite helicity is a consequence of the CPT symmetry, which is present in any Lorentz-invariant quantum field theory with a hermitian Hamiltonian.

Problem 2.2. Angular momentum and parity of graviton states

In this problem we examine the possible angular momentum states of the graviton. Let us first recall how such an analysis works for the photon, following Landau and Lifshitz, Vol. IV (1982), Section 6. After choosing the radiation gauge $A_0 = 0$ and $\nabla\cdot\mathbf{A} = 0$, a photon is described by a vector $\mathbf{A}(\mathbf{x})$ or, in momentum space, by $\tilde{\mathbf{A}}(\mathbf{k})$, subject to the transversality condition $\mathbf{k}\cdot\tilde{\mathbf{A}}(\mathbf{k}) = 0$. Let us at first neglect the transversality condition. The vector character of $\tilde{\mathbf{A}}(\mathbf{k})$ then corresponds to spin $s = 1$, and the total angular momentum j of a photon is given by the combination of $s = 1$ and of the orbital angular momentum l, with the usual composition rule of angular momenta in quantum mechanics. This means that a state with $j = 0$ can be obtained in only one possible way, i.e. combining the spin $s = 1$ with $l = 1$, while there are three states for each $j \neq 0$, which are obtained with $l = j, j \pm 1$ (for the purpose of this counting we consider a state with momentum j as one single state, regardless of the $2j + 1$ possible components of j_z).

To understand what is the parity of these states it is convenient to write explicitly their wavefunction. Just as the angular dependence of a scalar function can be expanded in terms of the (scalar) spherical harmonics $Y_{lm}(\theta, \phi)$, the angular dependence of a vector function $\mathbf{A}(\mathbf{x})$ can be expressed in vector spherical harmonics. As we will discuss in more detail in Section 3.5.2, see in particular eq. (3.247), these can be written as

$$\mathbf{Y}^l_{jj_z}(\theta,\phi) = \sum_{l_z=-l}^{l} \sum_{s_z=0,\pm1} \langle 1ls_zl_z|jj_z\rangle\, Y_{ll_z}(\theta,\phi)\, \boldsymbol{\xi}^{(s_z)} . \tag{2.198}$$

where $\boldsymbol{\xi}^{(s_z)}$ is the wavefunction of a spin-1 particle with a given value s_z of the projection of the spin on the z axis, and $\langle 1ls_zl_z|jj_z\rangle$ are the Clebsch–Gordan coefficient necessary to combine a spin state $|ss_z\rangle$ with $s = 1$ and an orbital angular momentum state $|ll_z\rangle$, so to obtain a total angular momentum $|jj_z\rangle$. In terms of the unit vectors $\mathbf{e}_x$, $\mathbf{e}_y$ and $\mathbf{e}_z$ of a Cartesian reference frame, the explicit form of the spin wavefunction is

$$\boldsymbol{\xi}^{(\pm1)} = \mp\frac{1}{\sqrt{2}}(\mathbf{e}_x \pm i\mathbf{e}_y)\,, \qquad \boldsymbol{\xi}^{(0)} = \mathbf{e}_z\,. \tag{2.199}$$

Consider a parity transformation, defined so that it changes the sign of the orientation of the axes of the reference frame, $\mathbf{e}_i \to -\mathbf{e}_i$.[34] Under this transformation, the state (2.198) picks a factor $P = (-1)^{l+1}$, where the $(-1)^l$ comes from the transformation of the scalar spherical harmonics and the further minus sign comes from the spin wavefunction. Thus, we can summarize as follows the possible states before imposing the condition $\mathbf{k}\cdot\tilde{\mathbf{A}}(\mathbf{k}) = 0$,

$$\begin{aligned}
j = 0: &\quad \text{one state,} &&(l = 1, P = +)\,,\\
j = 1: &\quad \text{three states,} &&(l = 0, P = -),\ (l = 1, P = +),\ (l = 2, P = -)\,,\\
j = 2: &\quad \text{three states,} &&(l = 1, P = +),\ (l = 2, P = -),\ (l = 3, P = +)\,,
\end{aligned} \tag{2.200}$$

and similarly we have three states for all higher values of j. We now impose the transversality condition $\mathbf{k}\cdot\tilde{\mathbf{A}}(\mathbf{k}) = 0$. This means that we remove a longitudinal state of the form $\tilde{\mathbf{A}}(\mathbf{k}) = \phi(\mathbf{k})\mathbf{k}$. The number of states of this form is therefore the same as the number of states of a scalar particles with wavefunction $\phi(\mathbf{k})$ (or, equivalently, of the scalar degree of freedom described by $\nabla\cdot\mathbf{A}$). When we develop ϕ in spherical harmonics, the total angular momentum j of

[34] In general, we can define parity either changing the sign of vectors with respect to a fixed reference frame, or reversing the orientation of the axes of the reference frame while keeping the vectors fixed. Here we adopt the latter point of view.

such a state is equal to the order l of the spherical harmonic, and its parity is $P = (-1)^j$. Thus, at each level j we must remove from eq. (2.200) one spurious state, with $P = (-1)^j$. Therefore at level $j = 0$ we end up with no physical states, while at all higher levels we end up with two physical states of opposite parity. This shows that for the photon there can be no monopole radiation, because there is no physical photon state with $j = 0$, while for all other values of j we have two physical states. For instance, the states with $j = 1$ correspond to an electric dipole photon ($P = -$) and a magnetic dipole photon ($P = +$).

Having understood the argument for the photon, we can adapt it to the graviton. The graviton is described by a 2×2 traceless symmetric tensor $\tilde{h}_{ij}(\mathbf{k})$ subject to the transversality condition $k^i\tilde{h}_{ij}(\mathbf{k}) = 0$. Again, we neglect at first the transversality condition. A symmetric traceless tensor corresponds to spin $s = 2$, while the parity on a true tensor is $P = (-1)^l$. Then, combining the orbital angular momentum with the spin $s = 2$, we have the following states

$$\begin{aligned}
j = 0: &\quad \text{one state,} &&(l = 2, P = +)\\
j = 1: &\quad \text{three states,} &&(l = 1, P = -),\ (l = 2, P = +),\ (l = 3, P = -)\\
j = 2: &\quad \text{five states,} &&(l = 0, P = +),\ (l = 1, P = -),\ (l = 2, P = +),\\
& && (l = 3, P = -),\ (l = 4, P = +)\,.
\end{aligned} \tag{2.201}$$

Similarly for all higher j there are five states, two with $P = +$, two with $P = -$ and one more with $P = (-1)^j$. We now impose the transversality condition. The most general traceless symmetric tensor which does *not* satisfy the transversality condition has the form

$$\tilde{h}_{ij}(\mathbf{k}) = a_i(\mathbf{k})k_j + a_j(\mathbf{k})k_i + b(\mathbf{k})(k_ik_j - \frac{1}{3}\delta_{ij}|\mathbf{k}|^2)\,, \tag{2.202}$$

with $k^ia_i(\mathbf{k}) = 0$, in order to respect the condition of zero trace. The most general spurious state is therefore parametrized by a scalar b and by a transverse vector a_i. Exactly as with the scalar ϕ found above, expanding b in spherical harmonics we have one state for each j, with parity $P = (-1)^j$. This eliminates the state with $j = 0$ in eq. (2.201), while at level $j = 1$ it leaves us with one state with $P = +1$ and one with $P = -1$, and at all higher j levels we are left with two states with $P = +1$ and two with $P = -1$. Finally, we must remove the spurious states described by $a_i(\mathbf{k})$. However this is a vector, transverse to $\mathbf{k}$, and therefore its states are the same as the photon states discussed above. There is no state at $j = 0$, and two states, with opposite parity, at all other j level. This remove the two states which were left at $j = 1$, and leaves us with two states, with opposite parity, at all higher levels. In conclusion, for the graviton,

$$\begin{aligned}
j = 0: &\quad \text{no state}\,,\\
j = 1: &\quad \text{no state}\,,\\
j = 2, 3, \ldots: &\quad \text{two states, one with } P = +,\ \text{one with } P = -\,.
\end{aligned} \tag{2.203}$$

Therefore, for gravitational wave there can be no monopole nor dipole radiation, since these would correspond to gravitons with $j = 0$ and $j = 1$, respectively. We will come back to the multipole expansion of gravitational waves in Section 3.5.2, where we will show how to express the two states allowed for $j \geqslant 2$ in terms of tensor spherical harmonics, and we will verify again that states with $j = 0$ or $j = 1$ are not allowed.

Further reading

- For the quantum field-theoretical approach to gravitation see the *Feynman Lectures on Gravitation* by Feynman, Morinigo, and Wagner (1995) (which collects lectures given by Feynman in 1962–63), and also DeWitt (1967) and Veltman (1976). For explicit computations of graviton–graviton scattering see Grisaru, van Nieuwenhuizen and Wu (1975).
- The possibility of deriving Einstein equation from an iteration of linearized theory is discussed, among others, by Gupta (1954), Kraichnan (1955), Feynman, Morinigo, and Wagner (1995), and Ogievetsky and Polubarinov (1965). An explicit and elegant iteration leading from the equations of motion of linearized theory to the full Einstein equations was performed by Deser (1970) using a first order Palatini formalism. The ambiguity concerning boundary terms is discussed by Padmanabhan (2004).
- Phenomenological limits on the graviton mass are discussed by Goldhaber and Nieto (1974). The discontinuity as the graviton mass goes to zero was found by Iwasaki (1970), van Dam and Veltman (1970) and Zakharov (1970). Massive gravitons have been further discussed by Boulware and Deser (1972). The fact that linearized theory becomes singular as $m_g \to 0$ was discovered by Vainshtein (1972). The radiation of massive gravitons in linearized theory is discussed by van Niewenhuizen (1973). Discussions of the fate of the discontinuity are given in Deffayet, Dvali, Gabadadze and Vainshtein (2002) and in Arkani-Hamed, Georgi and Schwartz (2003). The difficulty of performing the matching to an asymptotically flat solution, and the possibility of matching to a De Sitter solution, is discussed in Damour, Kogan and Papazoglou (2003). The fact that beyond linearized theory the trace h becomes a ghost is discussed by Boulware and Deser (1972) and, in full generality, by Creminelli, Nicolis, Papucci and Trincherini (2005).
- Lorentz-violating mass terms for $h_{\mu\nu}$ are discussed in Arkani-Hamed, Cheng, Luty and Mukohyama (2004), Rubakov (2004) and Dubovsky, Tinyakov and Tkachev (2005). In this case the mass of the scalar perturbations can be zero while the mass of the graviton $h_{ij}^{\rm TT}$ can be non-zero, and the bounds on the graviton mass derived from the Yukawa fall-off of the gravitational potential only refer to the scalar sector. Furthermore, these models do not suffer of the vDVZ discontinuity and do not have ghosts.

  A bound on the mass that refers directly to $h_{ij}^{\rm TT}$ can be obtained from pulsar timing, as recognized in Damour and Taylor (1991) and discussed quantitatively in Finn and Sutton (2002). The possibility of bounding the mass of $h_{ij}^{\rm TT}$ from the observation of inspiraling compact binaries is discussed in Will (1998) and Larson and Hiscock (2000).

3 Generation of GWs in linearized theory

We now consider the generation of GWs in the context of linearized theory. This means that we assume that the gravitational field generated by the source is sufficiently weak, so that an expansion around *flat* space-time is justified. For a system held together by gravitational forces, this implies that the typical velocities inside the source are small. For instance, in a gravitationally bound two-body system with reduced mass μ and total mass m, we have $E_{\rm kin} = -(1/2)U$, i.e.

$$\frac{1}{2}\mu v^2 = \frac{1}{2}\frac{G\mu m}{r}, \tag{3.1}$$

and therefore

$$\frac{v^2}{c^2} = \frac{R_S}{2r}, \tag{3.2}$$

where $R_S = 2Gm/c^2$ is the Schwarzschild radius associated to a mass m. A weak gravitational field means $R_S/r \ll 1$ and therefore $v \ll c$. Thus, for a self-gravitating system, weak fields imply small velocities. On the other hand, for a system whose dynamics is determined by non-gravitational forces, the weak-field expansion and the low-velocity expansion are independent, and in this case it makes sense to consider weak-field sources with arbitrary velocities, as we do in this chapter. This will allow us to understand, in the simple setting of a flat background space-time (and therefore Newtonian or at most special-relativistic dynamics for the sources), how GWs are produced. In Section 3.1 we will derive the formulas for GW production valid in flat space-time, but exact in v/c. Then, expanding the exact result in powers of v/c, we will see that for small velocities the GW production can be organized in a multipole expansion (Section 3.2). In Section 3.3 we discuss in detail the lowest order term, which is the quadrupole radiation. In Section 3.4 we discuss the next-to-leading terms, i.e. the mass-octupole and the current quadrupole radiation, and in Section 3.5 we present the systematic multipole expansion to all orders, using first the formalism of symmetric-trace-free (STF) tensors, and then the spherical tensors formalism. Finally, in a Solved Problems section we discuss some applications of this formalism and we collect additional technical material.

The most interesting astrophysical sources of GWs, such as neutron stars, black holes or compact binaries, are self-gravitating systems. In