

RECEIVED: September 17, 2018

REVISED: October 24, 2018

ACCEPTED: November 28, 2018

PUBLISHED: December 17, 2018

What do gravitons say about (unimodular) gravity?

Mario Herrero-Valea

*Institute of Physics, Laboratory of Particle Physics and Cosmology,
Ecole Polytechnique Fédérale de Lausanne, Lausanne, CH-1015 Switzerland*

E-mail: mario.herrerovallea@epfl.ch

ABSTRACT: We revisit the problem of constraining the weak field limit of the gravitational lagrangian from S-matrix properties. From unitarity and Lorentz invariance of the S-matrix of massless gravitons, we derive on-shell gauge invariance to consist on the transverse part of the linearised diffeomorphisms group. Moreover, by looking to the interaction between sources, we conclude that there exist only two possible lagrangians that lead to a well-defined covariant interaction, corresponding to the weak field limits of General Relativity and Unimodular Gravity. Additionally, this result confirms the equivalence of the S-matrix of both theories around flat space-time.

KEYWORDS: Classical Theories of Gravity, Gauge Symmetry, Models of Quantum Gravity

ARXIV EPRINT: [1806.01869](https://arxiv.org/abs/1806.01869)

Contents

1	Introduction	1
2	Integer spin particles	3
3	On-shell gauge invariance	6
4	The lagrangian description	9
5	Off-shell gauge invariance	15
6	The non-linear theories	18
7	Discussion and conclusions	20

1 Introduction

Perturbative Quantum Field Theory (QFT) has proven to be one of the most powerful tools in making predictions about physical quantities. Starting from a lagrangian and by applying a finite set of rules, we are able to compute observable quantities with a great precision (at least in the perturbative regime). In particular, the formalism allows, via the Quantum Effective Action and the LSZ formula, to obtain the expression for the scattering matrix (S-matrix) of any given theory, an object which presumably contains all possible information on physical observables of an interacting theory about a flat space-time. However, there is a big conceptual gap between the Lagrangian and the S-matrix, and one could question if there exist a single choice of Lagrangian for every given S-matrix. In other words, how many different lagrangians could lead to the same physical predictions or, rephrasing it in yet another way, what are the minimal requirements for a given lagrangian to reproduce what we know from experiments and basic assumptions. This is of particular relevance in fields where our available data is limited or even meagre, as it is the case for Beyond Standard Model (BSM) physics or for the quantum regime of gravitational interactions. The lack of data in these fields have led to a fruitful development of multiple alternative theories and effective field theories, which aim to explain a very reduced number of observations, whenever these exist, and to produce a new set of predictions that could be contrasted in future experiments, altogether with possible solutions to open theoretical problems.

One of these everlasting problems is *the cosmological constant problem* [1, 2]. In particular, the fact that radiative corrections to vacuum energy, even those coming from Standard Model or BSM particles in absence of Quantum Gravity, are brobdingnagian when compared to the experimental value derived from observations. Although in the presence of a

bare cosmological constant in the gravitational lagrangian this is just a fine tuning problem, the fact that we need to adjust the value of the cosmological constant Λ with a precision of 120 orders of magnitude, makes this an unpleasant hierarchy problem much worse than the one of the Higgs boson mass. Likewise the latter, the existence of supersymmetry (SUSY) could relax the numerical clash, but the absence of any clue of the existence of SUSY in LHC data puts this solution in tension with naturalness.

Another possible solution that has attracted attention in the recent years although it is almost as old as GR itself¹ is provided by modifying the infra-red limit of gravitational interactions in Unimodular Gravity (UG) [4–9]. The core idea of UG relies on the fact that the Einstein equations of GR are not independent, but only their traceless part is dynamically non-trivial. Bianchi identities must be always satisfied though, and they provide an extra constraint in order to recover the trace part and the whole set of equations. The cosmological constant, being contained only in the trace part, decouples and it is substituted in the solutions to the equations of motion by an integration constant unrelated to the vacuum energy at all. Thus, fixing its value is a matter of fixing initial and boundary conditions.

From a lagrangian point of view, UG does not differ much from GR [10]. It is obtained from the Einstein-Hilbert action, which must be however varied under the condition of unit determinant for the metric tensor $|g| = 1$ and thus, as a requirement for this constraint to hold, the theory is only invariant under volume preserving diffeomorphisms. Additionally, since $\sqrt{|g|} = 1$, any coupling to the cosmological constant, or to any vacuum energy and its radiative corrections, is dropped out of the dynamics of the theory at any loop order. The effective cosmological constant of the theory is, again, an integration constant obtained when solving the equations of motion for the mean field. In the weak field limit, the dynamics of the theory is described by an action with a gauge symmetry being the product of transverse infinitesimal diffeomorphisms and local rescalings,² a group that has been dubbed *WTDiff* in previous works.

In the recent years there have been various interesting advances in understanding UG and the solidness of its proposal as alternative to GR. These extend from exhaustive studies of how to embed cosmological models into it [11–14] to a thorough research of the ultra-violet regime of the theory [15–19]; passing by several consistency studies of its interaction with matter fields [20–27]. Extensions of the idea of UG to higher spin fields have been also explored [28, 29]. Of particular interest is the work of [30], where it is proven that a *WTDiff* invariant theory is the only other possible Lorentz invariant theory with a massless spin two field, besides the weak limit of GR, that can be completed to a non-linear theory by following Deser’s recipe [31–34]. Of course, the completion they find is Unimodular Gravity.

Motivated by the fact that Unimodular Gravity seems to be as good as General Relativity for classical matters, while it seems to provide a solution to the cosmological constant

¹The equations of motion of Unimodular Gravity were actually proposed by Einstein himself in 1919 [3], although his goal was not to solve the cosmological constant problem but to try to explain the structure of the electron.

²These act by shifting the value of the graviton field $h_{\mu\nu}$ by a quantity proportional to its trace $h_{\mu\nu} \rightarrow h_{\mu\nu} + \phi\eta_{\mu\nu}$ with ϕ an arbitrary function.

problem, we wish to investigate this dichotomy from the viewpoint of massless gravitons and their S-matrix. Explicitly, we wish to revisit the standard lore that relativistic massless gravitons lead uniquely to the weak field limit of GR [35, 36], a fact that cannot be true in full generality precisely because of the existence of UG. Thus, we will study what are the minimal requirements to build a theory of massless gravitons and where the bifurcation appears. Additionally, we will answer two long standing questions about UG that are not well-understood in the literature yet — how does the equivalence principle arise in the theory [37] and what is the relation of its S-matrix with the one of GR.

The content of this paper is organized as follows. In section 2 we review several basic notions about physical states of gravitons and we derive the transformation rules for the polarizations. Later on, in section 3 we derive the form of the on-shell gauge invariance preserved by the physical states by requiring unitarity and Lorentz invariance of the S-matrix. In section 4 we use this to derive a weak field lagrangian which leads to a covariant interaction and we show how this lagrangian corresponds to the gauge fixed limit of known theories in section 5. Finally, we devote section 6 to discuss about the non-linear completion of the theory for gravitons and we draw conclusions in section 7.

2 Integer spin particles

We will closely follow [35, 36] in what follows. Although we are interested in gravity and thus in gravitons, it will be useful to consider also photons, since their construction is similar and they serve as a good simpler example of what we are studying.

We introduce photons and gravitons as massless particles with integer spin that move along a light-like trajectory in flat space-time with momentum q^μ satisfying³

$$q^2 = q^\mu q_\mu = 0. \tag{2.1}$$

The physical polarizations of a given particle with spin s are objects $\epsilon^{\mu_1 \mu_2 \dots \mu_s}$ that transform under the irreducible representations of the Lorentz group that preserve the on-shell condition (2.1). This is the little group \mathbb{L} , whose elements $L^\mu{}_\nu$ thus satisfy

$$L^\mu{}_\nu q^\nu = q^\mu \tag{2.2}$$

where q^μ is any light-like vector. In particular, to give a clear characterization of \mathbb{L} , we can take a particular momentum K^μ in the (t, x^3) plane

$$K^\mu = (\kappa, 0, 0, \kappa) \tag{2.3}$$

with κ arbitrary.

With this given vector, the little group corresponds to all those transformations that take the euclidean plane (x^1, x^2) onto itself. That is, we have $\mathbb{L} = E^2$. Any element of this

³Throughout this work will use the mostly-plus convention for the signature of the flat metric $\eta_{\mu\nu} = \text{Diag}(-1, 1, 1, 1)$. Greek indices run in the range $\{0, 1, 2, 3\}$, while latin indices run only over space directions.

group will consist of a translation plus a rotation, with the latter being represented by a unitary operator

$$U[L] = e^{iJ_3\theta[L]} \quad (2.4)$$

where J_3 is the helicity of the particle along the x^3 direction, which for massless particles is aligned with their spin, and $\theta[L]$ is the rotation angle, different for every element of the little group \mathbb{L} .

For concreteness, any element of the little group can be described by the rotation angle $\theta[L]$ and a vector $X = (0, X^1, X^2, 0)$ through the matrix

$$\mathcal{E} = \begin{pmatrix} \cos\theta[L] & \sin\theta[L] & 0 & 0 \\ -\sin\theta[L] & \cos\theta[L] & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -X^1 & X^1 \\ 0 & 1 & -X^2 & X^2 \\ X^1 & X^2 & 1 - \frac{(X^1)^2+(X^2)^2}{2} & \frac{(X^1)^2+(X^2)^2}{2} \\ X^1 & X^2 & -\frac{(X^1)^2+(X^2)^2}{2} & 1 + \frac{(X^1)^2+(X^2)^2}{2} \end{pmatrix}. \quad (2.5)$$

The corresponding polarizations can now be constructed easily in the frame where $q^\mu = K^\mu$ starting from the corresponding ones for a spin one particle. There, the orthogonal euclidean plane can be parametrized by using the following basis

$$e_\pm^\mu = \frac{1}{\sqrt{2}}(0, 1, \pm i, 0) \quad (2.6)$$

which can be directly taken to be the polarizations of a spin one particle (photon herein after) in this frame. For a different momentum q^μ we just need to act on e_\pm^μ with the rotation $R^\mu{}_\nu(q)$ which aligns the x^3 -axis with q^μ

$$\epsilon_\pm^\mu(q) = R^\mu{}_\nu(q)e_\pm^\nu. \quad (2.7)$$

From this construction, it is easy to see that the photon polarization satisfies the following properties

$$\begin{aligned} (\epsilon_\pm^\mu(q))^* \epsilon_{\pm\mu}(q) &= 1 \\ \epsilon_\pm^\mu(q) \epsilon_{\mu\pm}(q) &= 0 \\ (\epsilon_\pm^\mu(q))^* &= \epsilon_\mp^\mu(q) \\ \epsilon_\pm^0(q) &= q_\mu \epsilon_\pm^\mu = 0. \end{aligned} \quad (2.8)$$

Polarizations for higher spin particles belong to irreducible representations of the direct product of $\epsilon_\pm^\mu(q)$. For the particular case of a spin two particle (graviton herein after), the two polarizations can be taken to be

$$\epsilon_\pm^{\mu\nu}(q) = \epsilon_\pm^\mu(q)\epsilon_\pm^\nu(q) \quad (2.9)$$

and they satisfy a set of properties inherited from (2.8). The ones that will be relevant for the rest of this work are

$$\begin{aligned} \eta_{\mu\nu}\epsilon_\pm^{\mu\nu}(q) &= 0 \\ q_\mu\epsilon_\pm^{\mu\nu}(q) &= 0 \\ \epsilon_\pm^{0\nu}(q) &= 0 \\ \epsilon_\pm^{\mu\nu}(q) &= \epsilon_\pm^{\nu\mu}(q). \end{aligned} \quad (2.10)$$

Note that altogether these properties imply that the polarizations can actually be written by using purely space indices as $\epsilon_{\pm}^{ij}(q)$ and that they are both traceless and transverse. In this section we will keep the full space-time greek indices thought, but we will resort to space indices later on this work.

Although it carries a Lorentz index, the photon polarization $\epsilon_{\pm}^{\mu}(q)$ is *not* a vector, since it does not transform homogeneously when acting with a Lorentz transformation upon it. In other words

$$\alpha^{\mu}_{\nu} \epsilon_{\pm}^{\nu}(q) \neq \epsilon_{\pm}^{\mu}(\alpha q) \quad (2.11)$$

where α^{μ}_{ν} is an arbitrary Lorentz transformation.

In order to find what is the real transformation rule, let us go to the frame where we take the momentum to be K^{μ} by means of a Lorentz transformation defined as

$$q^{\mu} = \Lambda^{\mu}_{\nu}(q) K^{\nu}. \quad (2.12)$$

Here $\Lambda^{\mu}_{\nu}(q)$ is the result of combining the rotation $R^{\mu}_{\nu}(q)$ in (2.7) with a boost along the x^3 -axis which takes the vector to the appropriate length $\Lambda^{\mu}_{\nu}(q) = R^{\mu}_{\alpha}(q) B^{\alpha}_{\nu}(q)$. By using it, we rewrite the former inequality as

$$\alpha^{\mu}_{\nu} \Lambda^{\nu}_{\rho}(q) e_{\pm}^{\rho}(q) \neq \Lambda^{\mu}_{\nu}(\alpha q) e_{\pm}^{\nu} \quad (2.13)$$

where we must note that the boost leaves any polarization unaffected, since they are blind to the length of the momentum.

Multiplying by the inverse of the operator in the l.h.s,

$$e_{\pm}^{\mu} \neq [\Lambda^{-1}(q)]^{\mu}_{\nu} [\alpha^{-1}]^{\nu}_{\rho} [\Lambda(\alpha q)]^{\rho}_{\sigma} e_{\pm}^{\sigma}. \quad (2.14)$$

Therefore, in order to study how $\epsilon_{\pm}^{\mu}(q)$ behaves under a Lorentz transformation, it is enough to find what is the action of the operator in the r.h.s. of the former expression

$$\Xi^{\mu}_{\sigma} = [\Lambda^{-1}(q)]^{\mu}_{\nu} [\alpha^{-1}]^{\nu}_{\rho} [\Lambda(\alpha q)]^{\rho}_{\sigma}. \quad (2.15)$$

By taking into account the definition (2.12), we see that Ξ^{μ}_{ν} is an element of the little group \mathbb{L} , satisfying $\Xi^{\mu}_{\nu} K^{\nu} = K^{\mu}$. Its action over e_{\pm}^{μ} it is obtained from (2.5) and can be written as a rotation, as given in (2.4), plus a translation along K^{μ}

$$\Xi^{\mu}_{\nu} e_{\pm}^{\nu} = e^{\pm i\theta[\Xi]} e_{\pm}^{\mu} + X_{\pm} K^{\mu} \quad (2.16)$$

where X_{\pm} are the components of the translation and $J_3 = 1$ for a spin one massless particle.

Acting with Ξ^{μ}_{ν} on e_{\pm}^{μ} and applying $\Lambda^{\mu}_{\nu}(q)$ on both sides, this can be rewritten as

$$[\alpha^{-1}]^{\mu}_{\nu} \epsilon_{\pm}^{\nu}(\alpha q) = e^{\pm i\theta[\Xi]} \epsilon_{\pm}^{\mu}(q) + X_{\pm} q^{\mu}. \quad (2.17)$$

The form of X_{\pm} can be obtained by setting $\mu = 0$ in this expression, which allows us to rewrite it as

$$\left([\alpha^{-1}]^{\mu}_{\nu} - [\alpha^{-1}]^0_{\nu} \frac{q^{\mu}}{|q|} \right) \epsilon_{\pm}^{\nu}(\alpha q) = e^{\pm i\theta[\Xi]} \epsilon_{\pm}^{\mu}(q) \quad (2.18)$$

where $|q|$ is the length of the space component of the momentum q^i .

Note that if it were not because of the second term on the l.h.s., we would have found an homogeneous transformation, with the Lorentz transformation corresponding to a rotation around the x^3 axis. This second term is precisely what makes the polarization a non-covariant object and corresponds, note the 0 index, to a boost, yielding the transformation group non-semi-simple. However, this is a no-loss situation for us, since requiring the non-homogeneous parts to cancel for any physical quantity will lead to constraints that will reduce the space of possible physical theories compatible with our assumptions.

In the case of higher spin particles, the corresponding transformation rules can be derived by considering the action of the operator Ξ^μ_ν on the different pieces of the direct product.

3 On-shell gauge invariance

Let us now turn our attention to the interacting theory. In such a QFT where a set of massless particles interact in flat space-time, any individual process can be described by the corresponding element in the S-matrix. For a process with n incoming and m outgoing particles, this element will take the following schematic form in momentum space⁴

$$S_{n \rightarrow m} = (\epsilon^{i_1}(q_{i_1}))^* (\epsilon^{i_2}(q_{i_2}))^* \dots (\epsilon^{i_m}(q_{i_m}))^* \epsilon^{j_1}(q_{j_1}) \epsilon^{j_2}(q_{j_2}) \dots \epsilon^{j_n}(q_{j_n}) M_{i_1 i_2 \dots i_m}^{j_1 j_2 \dots j_n} \quad (3.2)$$

where the various $\epsilon^k(q_k)$ represent the polarizations of the different particles, with i_k, j_k interpreted as sets of indices. These are contracted with the scattering amplitude $M_{i_1 i_2 \dots i_m}^{j_1 j_2 \dots j_n}$, which might be computed, in the perturbative regime, with the help of Feynman diagrams if we had a given lagrangian at our disposal. The scattering amplitude is thus a true tensor, transforming homogeneously under the action of the Lorentz group, and can only depend on the momentum and helicities of the particles involved in the scattering process. The fact that the full S-matrix element, and not only the perturbative amplitude, can be expressed in this way is proven in appendix B of [35].

Consider now a simple process of emission of a graviton of momentum q^μ by the interaction of a set of fields. The S-matrix element of such a process will be just

$$S_\pm(q) = (\epsilon_\pm^{\mu\nu}(q))^* M_{\mu\nu}. \quad (3.3)$$

This structure is pretty simple. All details of the interacting theory and of the sources leading to graviton emission will be contained in the scattering amplitude $M_{\mu\nu}$, while contraction with the polarization selects only those terms which produce the physical graviton with light-like momentum.

In the following we will assume that the QFT producing such a S-matrix element is unitary. Moreover, since we are looking to transitions with different number of external

⁴This representation corresponds to a choice of normalization for the states in the Hilbert space of the form

$$\langle q|p\rangle = 2q^0(|q|) \delta^{(3)}(q-p). \quad (3.1)$$

legs, while keeping all the details of the interaction into the undetermined amplitude $M_{\mu\nu}$, these elements have to independently lead to unitary transitions.⁵ Thus, S_{\pm} has to satisfy

$$S_{\pm} S_{\pm}^{\dagger} = \mathbb{I} \tag{3.4}$$

in any physical frame. In particular, this implies that acting with a Lorentz transformation on $S_{\pm}(q)$ can, as most, modify it by adding a phase

$$S_{\pm}(q) = e^{\pm 2i\bar{\theta}} S(\alpha q) \tag{3.5}$$

where $\bar{\theta}$ depends on the particular Lorentz transformation applied.

However, according to the particular representation (3.3), when acting with a Lorentz transformation and using (2.9) and (2.18), we find instead

$$S_{\pm}(q) = e^{\pm 2i\theta[\Xi]} \left([\alpha^{-1}]^{\mu}_{\rho} - [\alpha^{-1}]^0_{\rho} \frac{q^{\mu}}{|q|} \right)^* \left([\alpha^{-1}]^{\nu}_{\sigma} - [\alpha^{-1}]^0_{\sigma} \frac{q^{\nu}}{|q|} \right)^* (\epsilon_{\pm}^{\rho\sigma}(\alpha q))^* M_{\mu\nu}. \tag{3.6}$$

For an infinitesimal Lorentz transformation $\alpha = \delta + \omega$ this reduces to

$$S_{\pm}(q) = e^{\pm 2i\theta[\Xi]} S(\alpha q) - e^{\pm 2i\theta[\Xi]} \frac{\omega^0}{|q|} [(\epsilon_{\pm}^{\mu\alpha}(q))^* q^{\nu} + (\epsilon_{\pm}^{\nu\alpha}(q))^* q^{\mu}] M_{\mu\nu} + O(\omega^2). \tag{3.7}$$

As we see, the first term precisely corresponds to the transformation (3.5) imposed by unitarity. The second piece, coming from the boost part in the transformation (2.18) of the polarization, must thus drop out. Since the Lorentz transformation that we used is completely arbitrary, the only possible way to solve this discrepancy is to ask the scattering amplitude to satisfy

$$(\epsilon_{\pm}^{\mu\alpha}(q))^* q^{\nu} M_{\mu\nu} = 0 \tag{3.8}$$

which enforces the condition

$$q^{\mu} M_{\mu\nu} = \sigma q_{\nu} \tag{3.9}$$

where σ is an undetermined proportionality factor.

What we have derived here is no other thing than the on-shell Ward identities for the scattering amplitude. These are required, when going from the lagrangian description to the S-matrix, to drop the non-physical polarizations out of any physical process. The word on-shell here, stating that this is satisfied only by physical states where $q^2 = 0$, is important, since off-shell and on-shell gauge invariance do not need to agree, as we will see later.

In order to understand this point, let us repeat the same construction for the S-matrix of emission of photons, whose polarization is $\epsilon_{\pm}^{\mu}(q)$. The computation is identical to the one just showed for gravitons, one just needs to substitute the polarizations at all steps

⁵Details such as intermediate processes in perturbation theory are hidden inside $M_{\mu\nu}$, which represents ideally the full non-perturbative amplitude. Moreover, we are not considering internal propagation of gravitons in $M_{\mu\nu}$, which justifies that the S-matrix is diagonal in the external base of states. Alternatively, in [35] it is shown that (3.5) also follows from the transformation law for one-particle states.

and take into account that the scattering amplitude now carries a single index M_μ . The on-shell Ward identity equivalent to (3.9) is

$$q^\mu M_\mu = 0 \tag{3.10}$$

which agrees with the usual requirement of transversality for the scattering amplitude. This implies that any shift of the polarization of the form

$$\epsilon_\pm^\mu(q) \rightarrow f(q)q^\mu \tag{3.11}$$

with $f(q)$ an arbitrary function, will vanish and won't contribute to the S-matrix and, therefore, to any physical process. The polarization is not universally determined but instead there is a redundancy in choosing it corresponding to what we usually call a gauge symmetry. Actually, the transversality condition $q_\mu \epsilon_\pm^\mu = 0$ can be regarded as a gauge choice for this symmetry, which is of course preserved by the physical states due to the fact that $q^2 = 0$, which makes the shift (3.11) irrelevant.

The situation is slightly more subtle for gravitons, though. In this case, the on-shell Ward identity (3.9) allows for a similar gauge redundancy as with the photon

$$\epsilon_\pm^{\mu\nu}(q) \rightarrow q^\mu f^\nu + q^\nu f^\mu \tag{3.12}$$

where, in order to preserve the gauge choice $q_\mu \epsilon_\pm^{\mu\nu} = 0$, the vector f^μ must have only transverse components, satisfying $q_\mu f^\mu = 0$.

If we were deriving this S-matrix from General Relativity in its weak field version, the Fierz-Pauli lagrangian of metric perturbations $h^{\mu\nu}$, we would have assumed that the gauge symmetry had to be the linearised diffeomorphisms group, corresponding to

$$h^{\mu\nu} \rightarrow q^\mu l^\nu + q^\nu l^\mu \tag{3.13}$$

which looks similar to what we have found except for a little difference. The vector l^μ is not constrained to be transverse but it also contains a longitudinal part. However, this longitudinal part can be always written as

$$l_L^\mu = q^\mu L(q) \tag{3.14}$$

with $L(q)$ corresponding to a function of the momentum. This piece satisfies the constraint $q_\mu l_L^\mu = 0$ as an identity due to the on-shell condition $q^2 = 0$ and thus it does not need to be imposed as a constraint on the scattering amplitude for what it concerns to the S-matrix. In other words, on-shell states are blind to the existence of this longitudinal piece and therefore its presence is not required for consistency of the on-shell symmetry. Indeed, it was shown in [38, 39] that the transverse transformations are enough to reduce the number of polarizations of the massless graviton down to two.

The transverse transformations form a group that has been named *TDiff* before [39] and they consist on the linearised version of volume preserving diffeomorphisms. Adding the longitudinal piece represents, when off-shell, the embedding of this in a bigger group, the linearised version of the whole diffeomorphism group. This means that, although the

extra piece is invisible for on-shell dynamics, it will make a difference when defining off-shell fields, since their action needs to be invariant under the whole symmetry group that we choose to carry on.

This is one of the main points that we want to address here. Lorentz invariance of the S-matrix does not enforce the longitudinal part of the linearised diffeomorphism group to be a symmetry of physical states. Of course, it can always be taken to be so, as it is done in [35, 36], but then this would lead eventually to the weak field limit of GR as the only possible lagrangian theory to describe the system. Here instead we will not impose the longitudinal part as a symmetry but we will solely stick to the transverse transformation imposed by (3.9).

4 The lagrangian description

Let us now to try to derive the most general off-shell lagrangian which leads to the dynamics of gravitons described by the polarizations $\epsilon_{\pm}^{\mu\nu}(q)$ and enjoying the on-shell gauge invariance that we have just derived. In order to do this, let us introduce the concept of a vacuum state $|0\rangle$ and ladder operators $a^{\dagger}(q), a(q)$ which create and annihilate quanta of particles with momentum q^{μ} . The fields that give rise to gravitons are then given by

$$\bar{A}^{\mu\nu}[x] = \sum_{\pm} \int \frac{d^3p}{(2\pi)^3} \epsilon_{\pm}^{\mu\nu}(p) \left(a(p)e^{ip \cdot x} + a^{\dagger}(p)e^{-ip \cdot x} \right) \quad (4.1)$$

where we are summing over polarizations in order to create a parity even field.⁶

At this point, we can give a new interpretation to the on-shell gauge invariance (3.12). Since the field is built in terms of the polarization, it means that it will not transform homogeneously under Lorentz transformations, but instead it will enjoy a transformation inherited from (2.18). At an infinitesimal level $\alpha = \delta + \omega$, this transformation reads

$$\epsilon_{\pm}^{\mu\nu}(q) = \epsilon_{\pm}^{\mu\nu}(\alpha q) - e^{\pm 2i\theta[\Xi]} \frac{\omega_{\alpha}^0}{|q|} \left[(\epsilon_{\pm}^{\mu\alpha}(q))^* q^{\nu} + (\epsilon_{\pm}^{\nu\alpha}(q))^* q^{\mu} \right] + O(\omega^2) \quad (4.2)$$

which can be rewritten as

$$\epsilon_{\pm}^{\mu\nu}(q) = \epsilon_{\pm}^{\mu\nu}(\alpha q) + q^{\mu} f^{\nu} + q^{\nu} f^{\mu} \quad (4.3)$$

with f^{μ} satisfying $q_{\mu} f^{\mu} = 0$, and given by

$$f^{\mu} = -e^{\pm 2i\theta[\Xi]} \frac{\omega_{\alpha}^0}{|q|} (\epsilon_{\pm}^{\nu\alpha}(q))^* . \quad (4.4)$$

We again find, although from a different point of view, the already known result that on-shell gauge invariance is precisely the redundancy required in order to cancel the non-physical degrees of freedom which obstruct Lorentz invariance of the theory.

⁶Alternatively we could construct a parity odd field by using $\epsilon_{+}^{\mu\nu} - \epsilon_{-}^{\mu\nu}$. However, parity odd fields couple to anti-symmetric sources, which do not exist in flat-space. Thus, we will refrain to study those in what follows and we will stick to the even combination.

From the transformation rule of the polarization, we derive the corresponding one for the field

$$\bar{A}^{\mu\nu}[x] \rightarrow \bar{A}^{\mu\nu}[x] + \partial^\mu \Phi^\nu + \partial^\nu \Phi^\mu \tag{4.5}$$

where Φ^μ is the Fourier transform of f^μ as given in (4.4), satisfying $\partial_\mu \Phi^\mu = 0$.

At this point, let us note that the field here constructed satisfies $\bar{A}^{0\mu} = \bar{A}^{\mu 0} = 0$, besides being traceless, symmetric, and transverse on its both indices. In what follows we will forget about full space-time indices and we will separate time, denoted by a zero, from space coordinates, which we will denote with latin letters. Therefore, we introduce the field⁷ \bar{A}^{ij} to describe the physical polarizations of the graviton, satisfying the following equations

$$\partial_i \bar{A}^{ij} = 0 \tag{4.6}$$

$$\delta_{ij} \bar{A}^{ij} = 0 \tag{4.7}$$

$$\square \bar{A}^{ij} = (\partial_i \partial^i - \partial_t^2) \bar{A}^{ij} = 0 \tag{4.8}$$

where the last line is just the statement that the field is on-shell, satisfying the Klein-Gordon equation and propagating massless degrees of freedom with a dispersion relation $(q^0)^2 = |q|^2$, which was our starting assumption.

We wish to find which possible lagrangians allow for external gravitons of the kind we have described. For that, we will consider weak field lagrangians of at most second order in the fields and of the form

$$\mathcal{L}_{\text{phys}} = -\frac{1}{2} \bar{A}^{ij} \mathcal{D}_{ijkl} \bar{A}^{kl} + \bar{A}^{ij} J_{ij} \tag{4.9}$$

with \mathcal{D}_{ijkl} a differential operator and J_{ij} being the source to which the graviton field couples.

If we were given this Lagrangian and insert it in a path integral, we could integrate out the fields by means of its propagator, which basically amounts to invert the operator \mathcal{D}_{ijkl} , and obtain a effective interaction between the sources of the form

$$V[J] = \frac{1}{2} J_{ij} (\mathcal{D}_{ijkl})^{-1} J_{kl} . \tag{4.10}$$

Here we will do the opposite. Starting from what we know of gravitons we will compute the interaction and, by requiring it to be Lorentz covariant, we will find which constrains we must impose on the lagrangian description. Note that since we want to derive a fully covariant lagrangian, J_{ij} must be the spatial part of some symmetric covariant source $J_{\mu\nu}$, which couples to the field through

$$\int d^4x \bar{A}^{\mu\nu} J_{\mu\nu} . \tag{4.11}$$

⁷We will omit the x dependence of the fields in position space for the rest of this work. It will be showed only when required for clarity of the discussion.

Moreover, since we do not want the non-physical gauge degrees of freedom to carry any interaction, we must require the coupling between the field and the source to be invariant under the on-shell transformation (3.12)

$$\int d^4x \bar{A}^{\mu\nu} J_{\mu\nu} \rightarrow 2 \int d^4x \partial^\mu \Phi^\nu J_{\mu\nu} = -2 \int d^4x \Phi^\nu \partial^\mu J_{\mu\nu} \quad (4.12)$$

where we have used the symmetric character of the source and integrated by parts in the last step. Therefore, in order to cancel this coupling for arbitrary Φ , we must require the source to satisfy

$$\partial^\mu J_{\mu\nu} = \partial_\nu \mathcal{X}[x] \quad (4.13)$$

so that the integral vanishes after integration by parts, since $\partial_\mu \Phi^\mu = 0$. Here $\mathcal{X}[x]$ is an in principle arbitrary function of the space-time coordinates.

It is important to note that gauge invariance here does not completely fix the form of the source, but instead allows for the existence of an arbitrary function. If we were repeating the same computation with a photon, or by assuming that the on-shell gauge invariance corresponds to the whole linearised diffeomorphisms group, as discussed in the last section; we would have found a different result, implying the whole conservation of the source and thus setting $\mathcal{X}[x] = 0$. However, *TDiff* invariance allows for this extra function and we will keep it arbitrary in what follows.

Now we focus on computing the propagator for the field \bar{A}^{ij} . This is easy, since we know that \bar{A}^{ij} is transverse and traceless, and it propagates massless degrees of freedom, thus its propagator must have a pole in $q^2 = 0$. The only option in momentum space is⁸

$$G^{ijkl}(q) = -\frac{\Pi^{ijkl}}{q^2 - i\epsilon} \quad (4.15)$$

where we are using the usual $i\epsilon$ prescription for the Feynman propagator.

The projector Π^{ijkl} must project over those fields that are both symmetric in the pairs $\{i, j\}$ and $\{k, l\}$, as well as under $\{i, j\} \leftrightarrow \{k, l\}$; while at the same time being traceless and transverse in both pairs of indices. It can be constructed in the following way. We introduce the transverse projector Π^{ij}

$$\Pi^{ij} = \delta^{ij} - \frac{q^i q^j}{|q|^2} \quad (4.16)$$

which ensures that we will select the transverse part of any element which we act upon with it. Thus, the following construction satisfies all the required conditions

$$\Pi^{ijkl} = \frac{1}{2} \left(\Pi^{ik} \Pi^{jl} + \Pi^{il} \Pi^{jk} - \Pi^{ij} \Pi^{kl} \right). \quad (4.17)$$

⁸We will define the propagator as

$$\mathcal{D}_{ijkl} G^{klab} = -(\delta^{ia} \delta^{jb} + \delta^{ib} \delta^{ja}). \quad (4.14)$$

Now, this is not covariant, in particular because of the presence of only space components, so it won't lead to a covariant interaction when contracted with the sources. In order to guess what is happening here, let us use the trick of [36] and introduce a time-like vector $n^\mu = (1, 0, 0, 0)$ so that we can rewrite the spatial momentum q^i in terms of the full four-momentum q^μ

$$(0, q^i) \equiv q^\mu - n^\mu q^0. \quad (4.18)$$

The transverse projector can equally be rewritten as

$$\Pi^{\mu\nu} = \begin{pmatrix} 0 & 0 \\ 0 & \Pi^{ij} \end{pmatrix} \equiv g^{\mu\nu} + \frac{n^\mu q^\nu + n^\nu q^\mu}{|q|^2} q^0 + \frac{n^\mu n^\nu}{|q|^2} q^2 - \frac{q^\mu q^\nu}{|q|^2} \quad (4.19)$$

and therefore the propagator can be cast into the form

$$G^{ijkl}(q) \equiv G^{\mu\nu\alpha\beta}(q) = -\frac{1}{2} \left(\Pi^{\mu\alpha} \Pi^{\nu\beta} + \Pi^{\mu\beta} \Pi^{\nu\alpha} - \Pi^{\mu\nu} \Pi^{\alpha\beta} \right) \frac{1}{q^2 - i\epsilon} \quad (4.20)$$

which, let us stress, is not a covariant object due to the presence of the time-like vector n^μ . The advantage of this form is that we can now compute the effective action for $J_{\mu\nu}$ easily as

$$V[J] = \frac{1}{2} J_{\mu\nu} \frac{\Pi^{\mu\nu\alpha\beta}}{q^2 - i\epsilon} J_{\alpha\beta} \quad (4.21)$$

which splits in three parts

$$\begin{aligned} V[J] &= \frac{1}{2} \frac{1}{q^2 - i\epsilon} \left(J_{\mu\nu} J^{\mu\nu} - \frac{1}{2} J^2 + 2\mathcal{X}J - 4\mathcal{X}^2 \right) \\ &\quad + \frac{1}{2|q|^2} \left(2J_\mu^0 J_0^\mu + 2\mathcal{X}^2 - J\mathcal{X} - JJ_{00} \right) \\ &\quad - \frac{q^2}{2|q|^4} \left(\frac{1}{2} (J_{00})^2 + \mathcal{X}J_{00} + \frac{1}{2} \mathcal{X}^2 \right) \end{aligned} \quad (4.22)$$

where $J = \eta^{\mu\nu} J_{\mu\nu}$ and we have used the source conservation constraint $\partial_\mu J^{\mu\nu} = \partial^\nu \mathcal{X}$.

Before going further, let us note that the interaction (4.22) can be rewritten in a simpler form

$$V[J] = \frac{1}{2} \frac{1}{q^2 - i\epsilon} \left(\tilde{T}_{\mu\nu} \tilde{T}^{\mu\nu} - \frac{1}{2} T^2 \right) + \frac{1}{2|q|^2} \left(2\tilde{T}_{i0} \tilde{T}_0^i - \tilde{T}_{00} \tilde{T}_{00} - \tilde{T}_{ii} \tilde{T}_{00} \right) - \frac{q^2}{4|q|^4} \tilde{T}_{00} \tilde{T}_{00} \quad (4.23)$$

where $\tilde{T}_{\mu\nu} = J_{\mu\nu} - \mathcal{X} \eta_{\mu\nu}$.

This new source $\tilde{T}_{\mu\nu}$ is not only symmetric, but also conserved, due to the fact that $\partial^\mu J_{\mu\nu} = \partial_\nu \mathcal{X}$. Since the only conserved symmetric source available in flat space-time is the energy-momentum tensor $T_{\mu\nu}$, this must differ from it at most in a constant piece, which can be always absorbed in the arbitrary function \mathcal{X} . Thus, we can identify $\tilde{T}_{\mu\nu}$ and $T_{\mu\nu}$, having

$$J_{\mu\nu} = T_{\mu\nu} + \mathcal{X} \eta_{\mu\nu}. \quad (4.24)$$

This intermediate result, that \mathcal{X} drops out from the interaction when written in this way, was to be expected, since it is no more than the statement of the *weak equivalence principle* in the weak field approximation — gravitons interact only with the conserved energy-momentum tensor, regardless of what other details or technical tools we are introducing in the theory. What we have done here is deriving the weak version of the equivalence principle from the propagation of massless gravitons. We leave out the question of the strong version of the principle, though. For that, we would need to introduce a self-consistent interaction between gravitons which preserves all the hypothesis previously discussed. Although it seem plausible, we will not get into that matter here.

Now, let us come back to the interaction (4.22). The first element in there represents indeed a covariant interaction propagated by a massless particle, thus the pole in $q^2 = 0$. The other two, however, are more problematic. First of all, they are not covariant, and even worse, they contain poles in $|q|^2$, so they will lead to instantaneous interactions and non-local contributions when taken to position space by a Fourier transform. Clearly, if we were to formulate a theory containing only the variables that we have carried all along until here, we would be in danger, unable to find a covariant interaction.

The problem can be solved by appending to \bar{A}^{ij} a new set of auxiliary variables in such a way that their contribution to the interaction will cancel the dangerous terms in (4.22). It is important to stress, though, that we are not adding new degrees of freedom by doing that, because their propagators, producing the last two lines of (4.22) with opposite sign, will not have poles in q^2 and therefore they will not propagate physical excitations.⁹

We thus introduce three new auxiliary variables, two scalars A^{00} and A^{ii} , and a vector A^{i0} , which must couple to T_{00} , T_{ii} and T_{i0} respectively through a Lagrangian

$$\mathcal{L}_{\text{aux}} = -\frac{1}{2}(A^{00}, A^{ii}, A^{i0})\mathfrak{M} \begin{pmatrix} A^{00} \\ A^{jj} \\ A^{i0} \end{pmatrix} + (A^{00}T_{00} + A^{i0}T_{i0} + A^{ii}T_{jj}) \quad (4.26)$$

where \mathfrak{M} is a matrix valued differential operator. In principle we could choose a different combination of couplings in the scalar sector, for instance

$$(A^{00}, A^{ii}) \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} T_{00} \\ T_{ii} \end{pmatrix} \quad (4.27)$$

with a, b, c and d constants. However, we can always rotate the fields A^{00} and A^{ii} to diagonalize this interaction, ending up with (4.26) after a rescaling. Note that for all these rotations to be equivalent, we must demand the matrix in the middle to be diagonalizable

⁹It is instructive to take a look to the situation for photons. In that case, the interaction will be between a transverse field A^i and a source j^i , and takes the form

$$V[j] = \frac{j_\mu j^\mu}{q^2 - i\epsilon} + \frac{(j_0)^2}{|q|^2}. \quad (4.25)$$

The requirement to cancel the non-covariant piece leads to the introduction of a fourth component A^0 whose equation of motion is the Coulomb's law $\partial_i \partial^i A^0 = -j_0$.

and thus invertible. Singular transformations will play an important role later, but for the moment we will consider only invertible ones.

The value of \mathfrak{M} must be chosen carefully in order to cancel the non-covariant contributions in (4.22). Thus, we have in momentum space

$$\mathfrak{M}^{-1} = \frac{1}{|q|^2} \begin{pmatrix} -\frac{q^2}{|q|^2} - 2 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix}. \quad (4.28)$$

Inverting this matrix we have

$$\mathfrak{M} = \begin{pmatrix} 0 & -|q|^2 & 0 \\ -|q|^2 & -q^2 - 2|q|^2 & 0 \\ 0 & 0 & \frac{|q|^2}{4} \end{pmatrix} \quad (4.29)$$

so the lagrangian in position space is

$$\mathcal{L}_{\text{aux}} = \frac{1}{4} A^{i0} \Delta A^{i0} - 2A^{ii} \Delta A^{jj} + A^{ii} \square A^{jj} - 2A^{ii} \Delta A^{00} + (A^{00} T_{00} + A^{i0} T_{i0} + A^{ii} T_{ii}). \quad (4.30)$$

To this, we need to append the Lagrangian for the physical modes \bar{A}^{ij} . This is simply

$$\mathcal{L}_{\text{phys}} = -\frac{1}{2} \bar{A}^{ij} \square \bar{A}_{ij} + \bar{A}^{ij} T_{ij} + \lambda \left(\partial_i \bar{A}^{ij} \partial_k \bar{A}_j^k \right) \quad (4.31)$$

where λ is a Lagrange multiplier that cancels the longitudinal components out in the limit $\lambda \rightarrow \infty$, ensuring transversality of the field.

The whole lagrangian $\mathcal{L}_{\text{phys}} + \mathcal{L}_{\text{aux}}$ now satisfies all the requirements that we want. It propagates physical massless gravitons and it leads to a covariant interaction in the effective action for the source. However, the interaction piece does not look covariant at first sight. In order to solve this, let us rescale the auxiliary fields in the following way

$$A^{i0} \rightarrow -2A^{i0} \quad A^{ii} \rightarrow \frac{1}{3} A^{ii}. \quad (4.32)$$

Since this is just a global rescaling, it will not lead to any new dynamics, but it allows us to rewrite the lagrangian as

$$\mathcal{L} = A^{i0} \Delta A^{i0} + \frac{2}{9} A^{ii} \Delta A^{jj} + \frac{1}{9} A^{ii} \square A^{jj} - \frac{2}{3} A^{ii} \Delta A^{00} - \frac{1}{2} \bar{A}^{ij} \square \bar{A}_{ij} + \lambda (\partial_i \bar{A}^{ij} \partial_k \bar{A}_j^k) + A^{\mu\nu} T_{\mu\nu} \quad (4.33)$$

with a now covariant interaction term, after defining the field

$$A^{\mu\nu} = \begin{pmatrix} A^{00} & A^{i0} \\ A^{i0} & A^{ij} \end{pmatrix} \quad (4.34)$$

where $A^{ij} = \bar{A}^{ij} + \frac{1}{3} A^{kk} \delta^{ij}$.

However, the real source of the theory is not $T_{\mu\nu}$ but $J_{\mu\nu} = T_{\mu\nu} + \eta_{\mu\nu}\mathcal{X}$. For the physical field there is no difference, since we have seen that all dependence in \mathcal{X} cancels out. However, it makes a difference for the auxiliary fields in the scalar sector, modifying (4.26) to be

$$\mathcal{L}_{\text{aux}} = -\frac{1}{2}(A^{00}, A^{ii}, A^{i0})\mathfrak{M} \begin{pmatrix} A^{00} \\ A^{jj} \\ A^{i0} \end{pmatrix} + \left(A^{00}T_{00} - 2A^{i0}T_{i0} + \frac{1}{3}A^{ii}T_{jj} + (A_{ii} - A_{00})\mathcal{X} \right) \tag{4.35}$$

where we have already rescaled the fields. This will induce an extra contribution in the effective action of the form

$$V_{\text{extra}} = \frac{\mathcal{X} (q^2(2T_{00} - \mathcal{X}) + 2|q|^2(-T_{00} + T_{ii} + 2\mathcal{X}))}{4|q|^4} \tag{4.36}$$

where the dependence on \mathcal{X} has not cancelled.

The most general lagrangian which propagates spin two degrees of freedom in a Lorentz invariant way is thus

$$\mathcal{L} = A^{i0}\Delta A^{i0} + \frac{2}{9}A^{ii}\Delta A^{jj} + \frac{1}{9}A^{ii}\square A^{jj} - \frac{2}{3}A^{ii}\Delta A^{00} - \frac{1}{2}\bar{A}^{ij}\square\bar{A}_{ij} + \lambda(\partial_i\bar{A}^{ij}\partial_k\bar{A}_j^k) + A^{\mu\nu}J_{\mu\nu} \tag{4.37}$$

and leads to the interaction we are looking for but it will also induce the extra piece (4.36) which cannot be cancelled if we keep the arbitrary character of the source.

5 Off-shell gauge invariance

The lagrangian (4.37) leads to the theory of gravitons that we were looking for and couples to the right source, defined by on-shell gauge invariance in (3.12). However, the construction is still not complete, since there is an explicit presence of an undetermined function \mathcal{X} which represents a new source and which induces a new contribution to the effective potential depending on the value of \mathcal{X} .

Additionally, there is the question of the origin of \mathcal{X} . It is a source for the gravitational field, so it must be constructed out of matter fields. However, the fact that it does not gravitate in the standard sense, the effective action induced by physical degrees of freedom containing only the energy-momentum tensor in order to comply with the equivalence principle, tangles up any effort of giving an explicit definition. Since we want our theory to be precisely defined, we need to eliminate the contribution of \mathcal{X} from the effective potential, erasing the ambiguity.

Looking to (4.36), we see that the simplest way is to enforce $\mathcal{X} = 0$, so that we directly identify the source with the energy-momentum tensor

$$J_{\mu\nu} = T_{\mu\nu}. \tag{5.1}$$

In this case, the source is now fully conserved $\partial^\mu J_{\mu\nu} = 0$ and, reverting the argument in (4.12), we see that the field must be now invariant under the whole set of linearised diffeomorphisms in order to avoid the contribution of gauge modes to the effective action

$$A^{\mu\nu} \rightarrow A^{\mu\nu} + \partial^\mu f^\nu + \partial^\nu f^\mu \tag{5.2}$$

where now f^μ is unconstrained, the transformation containing both transverse and longitudinal diffeomorphisms when evaluated off the mass-shell. Importantly, the longitudinal part of the transformation acts here by protecting the value $\mathcal{X} = 0$, ensuring stability of the action under renormalization in the absence of anomalies.

If we do so, the reduced lagrangian obtained from (4.37) can be derived from the *Fierz-Pauli* theory, the weak field limit of GR, for a symmetric tensor $h^{\mu\nu}$

$$\mathcal{L}_{\text{FP}} = -\frac{1}{4}h^{\mu\nu}\square h_{\mu\nu} + \frac{1}{4}h\square h - \frac{1}{2}\partial_\mu h\partial_\nu h^{\mu\nu} + \frac{1}{2}\partial_\mu h^{\mu\nu}\partial_\alpha h^\alpha_\nu. \tag{5.3}$$

If we fix the gauge symmetry of this theory, corresponding to linearised *Diff* and thus requiring of four conditions, by imposing

$$\partial_i \left(h^{ij} - \frac{1}{3}\delta^{ij}h^{kk} \right) = 0, \quad \partial_i h^{i0} + \frac{2}{3}\partial_0 h^{kk} = 0 \tag{5.4}$$

we recover the lagrangian (4.37), up to the use of the equations of motion and total derivatives. Thus, S-matrix unitarity and Lorentz invariance eventually lead to the Lagrangian of a spin two field that we have known for decades.

Since any combination of the auxiliary scalar fields A_{00} and A_{ii} can be taken to the diagonal form that we have used so far, it seems that this is the only option to get a theory of gravitons — to enforce $\mathcal{X} = 0$ and increase the symmetry to full linearised diffeomorphisms. However, this is only true for the case in which we can diagonalize the matrix in (4.27). In principle, we could also consider rotations of the fields which lead to singular matrices

$$\tilde{A}_{00} = \alpha A_{00} + \beta A_{ii}, \quad \tilde{A}_{ii} = \gamma (\alpha A_{00} + \beta A_{ii}). \tag{5.5}$$

These cannot be inverted and thus represent an isolated set of theories unrelated to the invertible rotations. For these transformations, the coupling to \mathcal{X} becomes

$$(\gamma - 1)(A_{00} - A_{ii})\mathcal{X} \tag{5.6}$$

which vanishes only when $\gamma = 1$, implying $\tilde{A}_{00} = \tilde{A}_{ii}$ or, in other words, that the new $\tilde{A}_{\mu\nu}$ must be traceless. That is, there exist a singular field redefinition, corresponding to defining a theory in which the trace is not dynamical, that solves the problem as well.

This seems overkilling and it is indeed subtle, since the field rotation is not invertible and seems to reduce the number of independent variables in the action. However, the outgoing theory is perfectly regular. Since we can always adjust the value of \mathcal{X} to absorb any difference in the trace without modifying the value of the on-shell effective action, we

could always have avoided to introduce the singular rotation and instead have chosen to define \mathcal{X} it as

$$\mathcal{X} = -\frac{1}{4}J + \mathcal{Y} \quad (5.7)$$

where \mathcal{Y} is again arbitrary. By doing this, the interaction part of the lagrangian takes the form

$$\begin{aligned} \mathcal{L}_{\text{int}} &= \left(J_{\mu\nu} - \frac{1}{4}J\eta_{\mu\nu} \right) A^{\mu\nu} + (A^{00} - A^{ii})\mathcal{Y} = \left(A_{\mu\nu} - \frac{1}{4}A\eta_{\mu\nu} \right) J^{\mu\nu} + (A^{00} - A^{ii})\mathcal{Y} \\ &= \left(A_{\mu\nu} - \frac{1}{4}A\eta_{\mu\nu} \right) T^{\mu\nu} + (A^{00} - A^{ii})\mathcal{Y} \end{aligned} \quad (5.8)$$

where we have used the fact that $J^{\mu\nu} = T^{\mu\nu} + \mathcal{X}\eta^{\mu\nu}$.

In this way we isolate all the arbitrariness in the definition of \mathcal{X} into the interaction with the trace of the graviton field, while the physical source $J_{\mu\nu}$ only couples to the traceless part of $A^{\mu\nu}$. This again shows that it is enough to define a theory for this traceless part in order to recover the interaction (4.22) and at the same time get rid of the problems associated to the arbitrariness of \mathcal{X} , forgetting about introducing any interaction with the trace, which is dynamically irrelevant at this level. Being this equivalent to performing the singular transformation (5.5), the latter must lead to a perfectly regular theory.

Let us then introduce the change of variables that makes $A^{\mu\nu}$ explicitly traceless, corresponding to the covariant realization of the singular field rotation (5.5)

$$A^{\mu\nu} \rightarrow \tilde{A}^{\mu\nu} = A^{\mu\nu} - \frac{1}{4}\eta^{\mu\nu}\eta_{\alpha\beta}A^{\alpha\beta}. \quad (5.9)$$

The result, which can be read from (4.37) after elementary algebra, also corresponds to a known lagrangian. Incidentally, it corresponds to the gauge fixed version of the only other known Lorentz invariant lagrangian which propagates a single spin-two physical degree of freedom, the *WTDiff* lagrangian of [39]

$$\mathcal{L}_{WTDiff} = -\frac{1}{4}h^{\mu\nu}\square h_{\mu\nu} + \frac{3}{32}h\square h - \frac{1}{4}\partial_\mu h\partial_\nu h^{\mu\nu} + \frac{1}{2}\partial_\mu h^{\mu\nu}\partial_\alpha h_\nu^\alpha + T_{\mu\nu}h^{\mu\nu}. \quad (5.10)$$

This is a theory that, besides being invariant under *TDiff*, also enjoys a linearised Weyl invariance $h_{\mu\nu} \rightarrow h_{\mu\nu} + \phi\eta_{\mu\nu}$ with ϕ an arbitrary function, which ensures the irrelevance of the trace and thus the W in its name. By taking the *WTDiff* lagrangian and fixing the gauge by using the traceless version of (5.4), one recovers exactly the traceless version of (4.22). Again, the increased symmetry, in this case the Weyl part of the linearised gauge group, protects the dynamics of the theory not by forbidding \mathcal{X} this time but by ensuring its irrelevance at the dynamical level.

The *WTDiff* theory, besides being completely equivalent at the dynamical level to the Fierz-Pauli theory [4, 40], also keeps the arbitrary character of the source in its core as a consequence of Weyl invariance. Any shift of the form $T_{\mu\nu} \rightarrow T_{\mu\nu} + \eta_{\mu\nu}\mathcal{X}$ is non-physical. As such, it does not source the equations of motion and thus vacuum solutions, defined as solutions to the equations of motion when the sources vanish, are still solutions even for

non-vanishing \mathcal{X} . This is contrary to what happens in General Relativity, or the linearised Fierz-Pauli theory, where \mathcal{X} sources the trace of the (linearised) Einstein equations and modifies the solution. Of course, this comes at the cost of having a more complicated gauge structure, leading to an open algebra in the ghost sector and requiring more sophisticated techniques for its quantization [16], but maybe it is not such a big prize to pay.

Finally, it is worth to comment and emphasize that the two options here discussed, the Fierz-Pauli and *WTDiff* lagrangians, are the only possible theories that describe dynamical gravitons and that can be found with this method,¹⁰ corresponding to invertible transformations and to the unique singular rotation that solves the issue. Any other option will lead to the extra interaction (4.36) which cannot be cancelled unless we set $\mathcal{X} = 0$ or we perform the traceless transformation (5.5) with $\gamma = 1$.

6 The non-linear theories

The two theories that we have derived in the previous section correspond both to the weak field limit of known gravitational theories. In the case of the Fierz-Pauli lagrangian, it is known that it reproduces the lowest order in the weak field limit of General Relativity. Indeed, let us take the Einstein-Hilbert action

$$S = M_p^2 \int d^4x \sqrt{|g|} R \tag{6.1}$$

with M_p being the Planck mass. The weak field theory is defined by expanding the metric around flat space-time

$$g_{\mu\nu} = \eta_{\mu\nu} + \frac{1}{M_p} h_{\mu\nu} \tag{6.2}$$

where the factor of M_p is included in order to get a canonical kinetic term for $h_{\mu\nu}$. After plugging this expansion into (6.1), the second order action exactly corresponds to the Fierz-Pauli lagrangian (5.3), while higher orders will contain interactions leading to reproducing the full dynamics of GR.

This derivation goes in the direction from GR to Fierz-Pauli. There is a way to perform the converse, though, introduced in [32].¹¹ Starting from the Fierz Pauli lagrangian and by introducing consistent self-interactions, the author in [32] is able to re-sum the series, eventually finding the Einstein-Hilbert action.

For years, it was thought that this path was unique. That starting from graviton self-interactions, the only possible non-linear theory to be found was GR. However, this claim was shown wrong in [30], where the authors show that the *WTDiff* theory is equally consistent and can be completed to the action of Unimodular Gravity

$$S_{UG} = M_p^2 \int d^4x R_{|g|=1} \tag{6.3}$$

¹⁰We allow for the possible existence of some exotic theory that might violate our assumptions here, although it seems difficult for such a theory to exist as a local QFT.

¹¹There are some concerns, however, with the structure of certain terms in the expansion of the graviton, since they seem to be non-analytic in the coupling constant. See discussion in [31, 41].

which is written as the Einstein-Hilbert action constrained to metrics of unit determinant. For computational purposes, it can be recast to an unconstrained form where we make the unimodular condition explicit by a change of variables [42]

$$g_{\mu\nu} \rightarrow |g|^{-\frac{1}{4}} g_{\mu\nu} \tag{6.4}$$

so that we have

$$S_{\text{UG}} = M_{\text{p}}^2 \int d^4x |g|^{\frac{1}{4}} \left(R + \frac{3}{32} \frac{\nabla_{\mu}|g|\nabla^{\mu}|g|}{|g|^2} \right) \tag{6.5}$$

where none of the variables here is constrained any more. Remarkably, this change of variable corresponds, for the weak field approximation, to the traceless transformation (5.9).

In this form, the action of Unimodular Gravity is invariant under volume preserving diffeomorphisms¹² as well as under Weyl transformations

$$g_{\mu\nu} \rightarrow \Omega^2(x) g_{\mu\nu}. \tag{6.6}$$

The linear version of these symmetries are precisely the gauge symmetries of the *WTDiff* theory and its full lagrangian can be obtained from (6.5) by performing a weak field expansion like in (6.2).

As we discussed, the advantage of the *WTDiff* theory over Fierz-Pauli is the fact that, thanks to the arbitrariness of the function \mathcal{X} in the definition of the source, a possible cosmological constant cannot have an effect on the dynamics and does not prevent flat space-time to be a solution of the equations of motion around which to define a sensible perturbative QFT. This property emerges in the action (6.5) because Weyl invariance forbids the coupling $\sqrt{|g|}\Lambda$ at the level of the lagrangian. Moreover, because of Weyl invariance and trace irrelevance, UG only couples to the *traceless* part of $T_{\mu\nu}$, so that the presence of vacuum energy sourced by matter fields is also irrelevant.

Although this looks overkilling, since it seems to forbid the dynamics driven by vacuum energy and thus any effect that could lead to the observed accelerated expansion of the Universe, this is not really true. By deriving the equations of motion of the action (6.5) in the gauge $|g| = 1$ we find the traceless part of the Einstein equations

$$R_{\mu\nu} - \frac{1}{4} g_{\mu\nu} R = \frac{1}{M_{\text{p}}^2} \left(T_{\mu\nu} - \frac{1}{4} g_{\mu\nu} T \right). \tag{6.7}$$

To this, we must append the fact that the second Bianchi identities must be satisfied not only as a dynamical consequence of *TDiff* invariance but also as a geometrical constraint for the space-time manifold

$$\nabla_{\mu} R^{\mu\nu} = \frac{1}{2} \nabla^{\nu} R. \tag{6.8}$$

Taking the derivative of (6.7) and using (6.8) and the conservation of the energy-momentum tensor, we find the following constraint

$$\frac{1}{4} \nabla_{\mu} \left(R + \frac{1}{M_{\text{p}}^2} T \right) = 0 \tag{6.9}$$

¹²That is, those diffeomorphisms that preserve the value of $\sqrt{|g|}$.

which implies

$$R + \frac{1}{M_{\text{p}}^2} T = 4\lambda \tag{6.10}$$

where λ is an *integration constant*.

Plugging this back on the traceless equations (6.7) we recover the whole set of Einstein equations

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu} + \lambda g_{\mu\nu} = \frac{1}{M_{\text{p}}^2} T_{\mu\nu} \tag{6.11}$$

with the integration constant λ taking the role of a cosmological constant. However, unlike Λ or vacuum energy, its value is stable under radiative corrections and it therefore does not suffer from this cosmological constant problem. Its value, whatever it is, must be fixed solely by the given initial conditions when solving the equations of motion or, in the QFT language, by an experiment measuring it, after which the result is not renormalized. In particular, this means that there is no hierarchy problem related to UV sensitivity to the introduction of new degrees of freedom whose vacuum energy gravitates. Moreover, any solution of GR can be always found in UG by tuning the value of λ in the proper manner.

7 Discussion and conclusions

Through this work we have studied how much a lagrangian formulation of a theory of gravitons is constrained by basic properties of the S-matrix. By assuming unitarity and Lorentz invariance of physical quantities, we have seen that the lagrangian describing massless gravitons is not unique but instead we found two possibilities, corresponding to the well-known Fierz-Pauli theory and to the so-called *WTDiff* theory. The key point to understand this duality is that the on-shell gauge invariance for the physical states is the same for both theories, the transverse part of the linearised diffeomorphism group, which we have denoted by *TDiff*. We have also seen that, regardless of the lagrangian, we need to append to the physical degrees of freedom a set of auxiliary ones, which do not lead to the propagation of particles but contribute to the effective action of the sources with Coulomb-like interactions needed to cancel non-covariant pieces.

In the case of Fierz-Pauli, the symmetry group is enhanced by the presence of a longitudinal transformation, which is tautological on-shell but off-shell leads to the conservation of the current $J_{\mu\nu}$ of the theory, which becomes then the energy-momentum tensor $T_{\mu\nu}$. The other possible completion is provided in the *WTDiff* theory, where we drop the interaction with the trace of the graviton field while nevertheless finding the same effective action for the source, which is now the traceless part of $T^{\mu\nu}$ and still satisfies the equivalence principle. Both theories enjoy non-linear completions by following the Deser trick. In the Fierz-Pauli case, General Relativity emerges from self-interactions, while Unimodular Gravity is obtained from the *WTDiff* lagrangian.

At this point it is thus natural to ask ourselves if there is a reason to prefer one theory over the other, apart from historical reasons. As far as we know, both GR and UG are

dynamically equivalent both in the weak field limit and classically, and they share the same S-matrix in flat space. Moreover, they are the only known Lorentz invariant theories that propagate only a massless spin two degree of freedom and also the only ones that can be completed from the weak field approximation to the non-linear level.

However, both theories are not identical. In particular, the problem of the radiative corrections to the cosmological constant appears in a different light in Unimodular Gravity. By not coupling to the trace of $T_{\mu\nu}$, vacuum energy does not gravitate at all and the problem is avoided. The cosmological constant is not a coupling in the lagrangian but an integration constant whose value must be set by initial conditions. It is true that there is left the problem of fixing this and to explain why it does take such an unnatural value on our Universe but at least, once the cosmological constant is fixed, radiative corrections do not modify its value producing a hierarchy problem.

The other difference that can be pointed out at this level is the fact that symmetry groups are different for GR and UG. While the former enjoys the full diffeomorphisms as a gauge symmetry, the latter is invariant only under those that preserve the volume, plus Weyl transformations. This leads to a problem in identifying physically equivalent solutions to the equations of motion, since the gauge orbits are different. In particular, this could lead to radical differences in the interpretation of horizons and the interior of black holes, where certain changes of frames are needed in order to avoid apparent singularities in the metric.

Last but not least, it has been pointed out by Padmanabhan [31] that the QFT formulation of Unimodular Gravity could be actually advantageous over that of General Relativity, since it can be written exactly as resembling a Yang-Mills theory, while this is not completely possible in GR. In any case, as far as we know, UG is as good as GR for many matters and leads to the same physics for as far we have tested gravity in our Universe. It would be good to find a way to discriminate between the two theories.

Acknowledgments

I am grateful for the hospitality of the Instituto de Física Teórica UAM-CSIC, where part of this work was done. I also wish to thank Enrique Álvarez and Carmelo P. Martín for fruitful discussions and/or email exchange and to Sergey Sibiryakov for discussions and very useful comments on a previous version of this text. My work has been supported by the Tomalla Foundation.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

- [1] S. Weinberg, *The Cosmological Constant Problem*, *Rev. Mod. Phys.* **61** (1989) 1 [[INSPIRE](#)].
- [2] J. Martin, *Everything You Always Wanted To Know About The Cosmological Constant Problem (But Were Afraid To Ask)*, *Comptes Rendus Physique* **13** (2012) 566 [[arXiv:1205.3365](https://arxiv.org/abs/1205.3365)] [[INSPIRE](#)].

- [3] A. Einstein, *Do gravitational fields play an essential part in the structure of the elementary particles of matter?* (in German), *Sitzungsber. Preuss. Akad. Wiss. Berlin* **1919** (1919) 349 [[INSPIRE](#)].
- [4] E. Álvarez, *Can one tell Einstein's unimodular theory from Einstein's general relativity?*, *JHEP* **03** (2005) 002 [[hep-th/0501146](#)] [[INSPIRE](#)].
- [5] W.G. Unruh, *A Unimodular Theory of Canonical Quantum Gravity*, *Phys. Rev.* **D 40** (1989) 1048 [[INSPIRE](#)].
- [6] M. Henneaux and C. Teitelboim, *The Cosmological Constant and General Covariance*, *Phys. Lett.* **B 222** (1989) 195 [[INSPIRE](#)].
- [7] M. Kreuzer, *Gauge Theory of Volume Preserving Diffeomorphisms*, *Class. Quant. Grav.* **7** (1990) 1303 [[INSPIRE](#)].
- [8] Y.J. Ng and H. van Dam, *A Small but nonzero cosmological constant*, *Int. J. Mod. Phys.* **D 10** (2001) 49 [[hep-th/9911102](#)] [[INSPIRE](#)].
- [9] W. Buchmüller and N. Dragon, *Einstein Gravity From Restricted Coordinate Invariance*, *Phys. Lett.* **B 207** (1988) 292 [[INSPIRE](#)].
- [10] S. Gielen, R. de León Ardón and R. Percacci, *Gravity with more or less gauging*, *Class. Quant. Grav.* **35** (2018) 195009 [[arXiv:1805.11626](#)] [[INSPIRE](#)].
- [11] G.F.R. Ellis, H. van Elst, J. Murugan and J.-P. Uzan, *On the Trace-Free Einstein Equations as a Viable Alternative to General Relativity*, *Class. Quant. Grav.* **28** (2011) 225007 [[arXiv:1008.1196](#)] [[INSPIRE](#)].
- [12] E. Álvarez and A.F. Faedo, *Unimodular cosmology and the weight of energy*, *Phys. Rev.* **D 76** (2007) 064013 [[hep-th/0702184](#)] [[INSPIRE](#)].
- [13] K. Nozari and S. Shafizadeh, *Cosmological inflation in a generalized unimodular gravity*, *Int. J. Mod. Phys.* **D 26** (2017) 1750107 [[arXiv:1712.09522](#)] [[INSPIRE](#)].
- [14] M. Shaposhnikov and D. Zenhausern, *Scale invariance, unimodular gravity and dark energy*, *Phys. Lett.* **B 671** (2009) 187 [[arXiv:0809.3395](#)] [[INSPIRE](#)].
- [15] E. Álvarez, S. González-Martín, M. Herrero-Valea and C.P. Martín, *Unimodular Gravity Redux*, *Phys. Rev.* **D 92** (2015) 061502 [[arXiv:1505.00022](#)] [[INSPIRE](#)].
- [16] E. Álvarez, S. González-Martín, M. Herrero-Valea and C.P. Martín, *Quantum Corrections to Unimodular Gravity*, *JHEP* **08** (2015) 078 [[arXiv:1505.01995](#)] [[INSPIRE](#)].
- [17] A. Eichhorn, *The Renormalization Group flow of unimodular $f(R)$ gravity*, *JHEP* **04** (2015) 096 [[arXiv:1501.05848](#)] [[INSPIRE](#)].
- [18] I.D. Saltas, *UV structure of quantum unimodular gravity*, *Phys. Rev.* **D 90** (2014) 124052 [[arXiv:1410.6163](#)] [[INSPIRE](#)].
- [19] L. Smolin, *The Quantization of unimodular gravity and the cosmological constant problems*, *Phys. Rev.* **D 80** (2009) 084003 [[arXiv:0904.4841](#)] [[INSPIRE](#)].
- [20] E. Álvarez, S. González-Martín and C.P. Martín, *Unimodular Trees versus Einstein Trees*, *Eur. Phys. J.* **C 76** (2016) 554 [[arXiv:1605.02667](#)] [[INSPIRE](#)].
- [21] C.P. Martín, *Unimodular Gravity and the lepton anomalous magnetic moment at one-loop*, *JCAP* **07** (2017) 019 [[arXiv:1704.01818](#)] [[INSPIRE](#)].

- [22] S. González-Martín and C.P. Martin, *Do the gravitational corrections to the β -functions of the quartic and Yukawa couplings have an intrinsic physical meaning?*, *Phys. Lett. B* **773** (2017) 585 [[arXiv:1707.06667](#)] [[INSPIRE](#)].
- [23] R. de León Ardón, N. Ohta and R. Percacci, *Path integral of unimodular gravity*, *Phys. Rev. D* **97** (2018) 026007 [[arXiv:1710.02457](#)] [[INSPIRE](#)].
- [24] S. González-Martín and C.P. Martin, *Unimodular Gravity and General Relativity UV divergent contributions to the scattering of massive scalar particles*, *JCAP* **01** (2018) 028 [[arXiv:1711.08009](#)] [[INSPIRE](#)].
- [25] S. González-Martín and C.P. Martin, *Scattering of fermions in the Yukawa theory coupled to Unimodular Gravity*, *Eur. Phys. J. C* **78** (2018) 236 [[arXiv:1802.03755](#)] [[INSPIRE](#)].
- [26] R. Percacci, *Unimodular quantum gravity and the cosmological constant*, in proceedings of the *Black Holes, Gravitational Waves and Spacetime Singularities*, Rome, Italy, 9–12 May 2017, *Found. Phys.* **48** (2018) 1364 [[arXiv:1712.09903](#)] [[INSPIRE](#)].
- [27] R. Carballo-Rubio, *Longitudinal diffeomorphisms obstruct the protection of vacuum energy*, *Phys. Rev. D* **91** (2015) 124071 [[arXiv:1502.05278](#)] [[INSPIRE](#)].
- [28] E.D. Skvortsov and M.A. Vasiliev, *Transverse Invariant Higher Spin Fields*, *Phys. Lett. B* **664** (2008) 301 [[hep-th/0701278](#)] [[INSPIRE](#)].
- [29] A. Campoleoni and D. Francia, *Maxwell-like Lagrangians for higher spins*, *JHEP* **03** (2013) 168 [[arXiv:1206.5877](#)] [[INSPIRE](#)].
- [30] C. Barceló, R. Carballo-Rubio and L.J. Garay, *Unimodular gravity and general relativity from graviton self-interactions*, *Phys. Rev. D* **89** (2014) 124019 [[arXiv:1401.2941](#)] [[INSPIRE](#)].
- [31] T. Padmanabhan, *From gravitons to gravity: Myths and reality*, *Int. J. Mod. Phys. D* **17** (2008) 367 [[gr-qc/0409089](#)] [[INSPIRE](#)].
- [32] S. Deser, *Gravity from self-interaction redux*, *Gen. Rel. Grav.* **42** (2010) 641 [[arXiv:0910.2975](#)] [[INSPIRE](#)].
- [33] S. Deser, *Selfinteraction and gauge invariance*, *Gen. Rel. Grav.* **1** (1970) 9 [[gr-qc/0411023](#)] [[INSPIRE](#)].
- [34] S. Deser, *Higher curvature gravities, unlike GR, cannot be bootstrapped from their (usual) linearizations*, *Gen. Rel. Grav.* **49** (2017) 149 [[arXiv:1705.08938](#)] [[INSPIRE](#)].
- [35] S. Weinberg, *Photons and Gravitons in s Matrix Theory: Derivation of Charge Conservation and Equality of Gravitational and Inertial Mass*, *Phys. Rev.* **135** (1964) B1049 [[INSPIRE](#)].
- [36] S. Weinberg, *Photons and gravitons in perturbation theory: Derivation of Maxwell's and Einstein's equations*, *Phys. Rev.* **138** (1965) B988 [[INSPIRE](#)].
- [37] T. Ortín, *Higher order gravities and the Strong Equivalence Principle*, *JHEP* **09** (2017) 152 [[arXiv:1705.03495](#)] [[INSPIRE](#)].
- [38] J.J. van der Bij, H. van Dam and Y.J. Ng, *The Exchange of Massless Spin Two Particles*, *Physica A* **116** (1982) 307 [[INSPIRE](#)].
- [39] E. Álvarez, D. Blas, J. Garriga and E. Verdaguier, *Transverse Fierz-Pauli symmetry*, *Nucl. Phys. B* **756** (2006) 148 [[hep-th/0606019](#)] [[INSPIRE](#)].
- [40] E. Álvarez and M. Herrero-Valea, *Unimodular gravity with external sources*, *JCAP* **01** (2013) 014 [[arXiv:1209.6223](#)] [[INSPIRE](#)].

- [41] L.M. Butcher, M. Hobson and A. Lasenby, *Bootstrapping gravity: A Consistent approach to energy-momentum self-coupling*, *Phys. Rev. D* **80** (2009) 084014 [[arXiv:0906.0926](#)] [[INSPIRE](#)].
- [42] E. Álvarez, *The Weight of matter*, *JCAP* **07** (2012) 002 [[arXiv:1204.6162](#)] [[INSPIRE](#)].