# Gravitation from Field Theory

Gabriel Massoni Salla
gabriel.massoni.salla@usp.br

*Department of Mathematical Physics, University of São Paulo, São Paulo, Brazil*

October 7, 2019

**Abstract**

Using field theoretical argumentation we show that General Relativity is the unique field theory that can explain gravitational phenomena. In addition, we use this result to analyse $f(R)$ and Massive Gravity theories.

## Contents

# Introduction

It is widely believed that General Relativity (GR) is the only theory consistent with the theory of a massless spin 2 particle, the graviton. In the literature one may find several proofs that that this is indeed the case [1, 2, 3, 4, 5]. But after a close inspection, one may also see that all these proofs are somewhat dissatisfying, in the sense that they are incomplete or biased on ideas from GR and none rely entirely on field theoretical methods to achieve the final result, the Einstein-Hilbert action [1, 6]. In this work we present a complete proof, from the very beginning. In particular, the approach we use here is much similar to that in Feynman's book [1], an analogy with gauge theories. There, Feynman used Quantum Electrodynamics (QED) to obtain the linear structure of gravity and from it could start doing calculations. Here we explore this idea further. By comparing $U(1)$ theories with Yang-Mills theories we identify how the non-linearity emerges in gauge theories and we use the same pattern to implement the non-linearity in the linear theory of gravity, while avoiding possible biased ideas or definitions from GR.

Although highly theoretical, this discussion involves many practical applications, for example in the context of modified theories of gravitation that attempt to generalise Einstein's gravity. The motivations for such theories are many [7, 8, 9] and include yet unsolved puzzles like the one of the cosmological constant. Most of those theories are formulated in a field theoretical framework, so in order to discuss them we need a deep understanding of gravity as a field theory. In the second part of this paper we will present very briefly two classes of modified gravity theories as an illustration of the previous discussion.

Our work is structured in the following way. In the first chapter we discuss how should we approach the problem from a field theoretical point of view and present some aspects on gravitation phenomenology which will be essential in our discussion. Then in the second chapter we determine the spin of the graviton. The third chapter is dedicated to construct and study some aspects of the linear theory of the graviton. Finally in the fourth chapter we construct step by step the non-linear theory: from realising that it must be non-linear to the computation of the infinitesimal graviton transformation. The last chapter is devoted to a brief study of two possible alternatives to GR, while pointing out how do they connect to the discussion of the previous chapters.

# 1 How to approach the problem

Our mindset is one of physicists who are already familiar with Quantum Field Theory (QFT) and the Standard Model of Particle Physics (SM), but have just recently discovered gravitational phenomena [1]. In such scenario, physicists will obviously search for a field description of gravity, though we must always bare in mind that there may be better theories which are not field theories. To construct such field theory, the first step is to determine the field we are going to deal with. Next we should verify if such field is already included in the SM, because then there would be nothing new to describe. If the gravitational field is not described by the SM, we must construct it based on experimental evidence. In particular, the experiments will point out what's the spin associated with such field. From this point forward the analysis becomes more standard [1]: we will construct the only quadratic Lagrangian consistent with unitarity, couple the field to a yet unknown physical current $T$, compute the free propagator and so on, just like in any other field theory. Some difficulties and subtleties will show up though, but that's to be expected from such mysterious interaction.

Long story short, we will try to answer the following questions:

1. From the observed phenomena, what are the properties of the gravitational interaction?
2. Is there a particle in the SM that can explain all these phenomena?
3. If not, i.e., if the gravitational interaction is not described by the SM, what's the spin of this new gravitational field?
4. What is the most general quadratic Lagrangian we can build consistent with our previous knowledge from QFT? What are the EoM and the free propagator?
5. How does this field interact with other fields (and perhaps with itself)?
6. Is our formulation consistent?

The last question, though an obvious one, will prove it worth of being stated at the very beginning of this paper.

Let's begin with what we observe. We gather here the 4 principal experimental facts:

- **Gravitational Potential**

  Mainly through the measure of celestial bodies' motions, we can conclude that there is a purely attractive force between them. This force is proportional to the product of the masses of the bodies and the proportionality factor is the gravitational constant $G$, that does not depend on the particular configuration [10]. The explicit expression of the observed potential is given by

$$V(r) = -\frac{GM}{r} + \mathcal{O}\left(\frac{1}{r^2}\right), \tag{1.1}$$

  where $M$ is the mass of the body generating such potential and $r$ the radial distance [6, 10].

  The first term is identical to the electromagnetic case, an inverse law proportional to the "charge". The second piece does not exist in classical Electromagnetism (EM), but only in QED when we include loop corrections to the classical potential [11]. That's not the case of Eq. (1.1), because this term is observed at the classical level, hence cannot be generated by radioactive correction coming from Quantum Mechanics. Such terms, in particular the $1/r^2$ one, are important for example in the perihelion precession of Mercury, which cannot be explained with a $1/r$ law alone [6, 12].

- **Principle of Equivalence**

  In standard geometrical formulation of gravity the equivalence principle plays a key role in the formulation of GR, it allows us to connect the gravitational interaction with General Coordinate Transformations (GCT) [13, 14]. The literature provides us with many distinct formulations of the same principle, for example the weak and strong versions [6]. But such details do not concern us, we focus on its very essence: gravity couples to everything in the same way, i.e., $G$ is an universal constant and every gravitational interaction will depend on it. In terms of field theory this means that all other fields that couples to gravity will do so with the same coupling constant $\lambda = \lambda(G)$.

  There are many experiments that validate the equivalence principle at the classical/macroscopic realm. At the quantum level experiments with bouncing neutrons [15, 16] and neutron interferometry [16, 17] reinforce the validity of this principle even in the quantum realm.

- **Coupling with Electromagnetism**

  The potential in Eq. (1.1) is proportional to the mass of the source, and the force is also proportional to the mass of the interacting particle. Does this means that massless particles like photons do not interact gravitationaly? We suspect that this is not true, since special relativity tells us that mass is energy and photon, despite massless, have energy. Indeed, photons interact with gravity and the first measurements that confirmed GR were of light deflection due to gravitational fields [18]. In particular, the angle by which a photon is deflected under the influence of a mass $M$ is

  $$\Delta\varphi = \frac{4MG}{r_0}, \tag{1.2}$$

  where $r_0$ is the smallest distance between the photon and the massive object [6, 13]. Note that there is no dependency on the photons momentum and/or its polarisation.

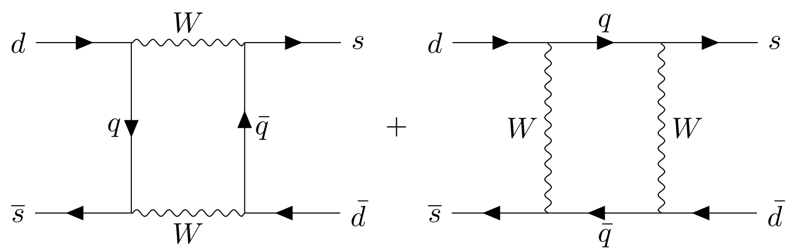
  In conclusion, gravity couples to everything that have energy; this includes both massive and massless particles.

- **Matter and Antimatter**

  We could imagine that gravity couples to matter and antimatter in different ways, just like the gauge bosons of the SM do. To clarify this issue we need a quantum mechanical experiment, in particular one that does not involve electromagnetic (EM) interactions, as they're much stronger than gravitational ones*. A very clear and straightforward example is the neutral kaon oscillation. The neutral kaon, $K^0$, due to its internal composition, is not its own antiparticle [19]. The SM predicts that a $K^0$ at rest has a non-zero probability of turning into its antiparticle, $\bar{K}^0$, and

*The ratio between the gravitational and electromagnetic force between two electrons is of order $10^{-40}$.

vice-versa, through the following amplitude [1, 20, 21]




where $q$ is a up type quark (up, charm or top), $K^0 = \langle \bar{s}d \rangle$ and $\bar{K}^0 = \left\langle \bar{d}s \right\rangle$. With this piece of information, we realise that the Hamiltonian of the system is not diagonal in the $\{K^0, \bar{K}^0\}$ basis, instead the Hamiltonian eigenstates are*

$$K_{1,2} = \frac{K^0 \pm \bar{K}^0}{\sqrt{2}}. \tag{1.3}$$

Since in this reference frame the Hamiltonian accounts only for the energy coming from the masses, the eigenstates have definite and measurable masses, say, $m_1$ and $m_2$. Inverting Eq. (1.3) and evolving $K^0$ in time we obtain

$$K^0 = \frac{K_1 + K_2}{\sqrt{2}} \Rightarrow K^0(t) = \frac{1}{\sqrt{2}}\left(K_1 + e^{-\frac{itc^2}{\hbar}\Delta m} K_2\right), \tag{1.4}$$

with

$$\Delta m \equiv m_2 - m_1. \tag{1.5}$$

It is indeed possible to measure this phase and set upper bounds on $\Delta m$. According to the Particle Data Group (PDG) the bound is [19]

$$\frac{\Delta m}{m_1 + m_2} < 10^{-19}. \tag{1.6}$$

But all experiments are performed under the influence of Earth's gravitational field, therefore, if $K^0$ and $\bar{K}^0$ interact differently with gravity, there would be an extra potential energy in the Hamiltonian, which would raise the bound in Eq. (1.6) [1]. Worst, if we considered the gravitational potential of the whole galaxy or even of the universe, this shift would become arbitrarily large, preventing us from getting anywhere near the bound (1.6). So gravity must couple to matter and antimatter in the same way.

Now we go back and ask ourselves: is gravity described by the SM? The answer is no, it is not. Due to the potential in Eq. (1.1) we conclude that gravity is mediated by a massless particle, called graviton. The massless particles in the SM cannot be the graviton, because they do not respect the principle of equivalence nor couple in the same way with matter and antimatter. Therefore

| No particle in the SM is the graviton |
|---|

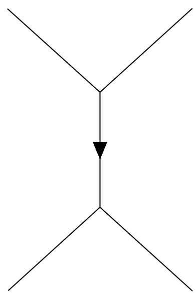
*Without considering $CP$ violation [12]

# 2 Spin of the graviton

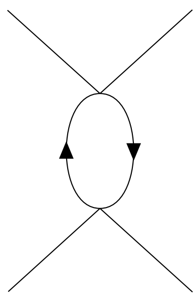
We need to introduce a new field to describe the graviton and the first question is, how does it transform under Lorentz transformations? Or equivalently, what is its spin*? We will consider every possibility in the following discussion.

## 2.1 Half-integer spin

In the 60's the idea that the neutrinos could play the role of the graviton was very popular [1]. The neutrinos are fermions, so the punchline was to have a fermionic mediator, something not present in the SM. Of course a fermionic mediator is possible, but there are a few problems. If we want to calculate the potential due to an exchange of a single (fermionic) graviton we should evaluate the diagram:



Though this isn't a potential, but a scattering! Fermions carry spin angular momentum, so the quantum numbers of the incoming particles will be altered, characterising a scattering. We can solve this issue considering that, not a single, but two gravitons are exchanged:



(2.1)

At first such diagram poses no problem, but it was shown that the potential associated with such process is proportional to $r^{-5}$, i.e. it contradicts Eq. (1.1) [23, 24]. This result can be understood in the low energy limit with dimensional analysis only. Consider an exchange of neutrinos interacting with external leptons via 4-Fermi theory. The first contribution to this potential comes from diagram (2.1), so we may conclude that the potential generated by a neutrino anti-neutrino pair is given by

$$V_{\nu\bar{\nu}} \sim G_F^2, \tag{2.2}$$

where $G_F$ is the Fermi constant. Considering a spherically symmetric potential, the only way to obtain the correct dimension is to include a $r^{-5}$ in Eq. (2.2), hence

$$V(r)_{\nu\bar{\nu}} \sim \frac{G_F^2}{r^5}. \tag{2.3}$$

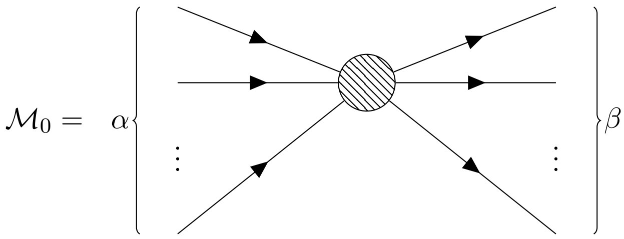
Similar calculations can be done using more general structures than the 4-Fermi theory and also considering vector boson mediation [23, 24]; nevertheless Eq. (2.3) remains valid.

*Of course, in correct terminology massless particles have helicity, not spin [22]

## 2.2 Spin $\geq 3$

As the spin grows, so does the difficulty to construct the appropriate theory. It would be certainly nice we could eliminate higher spins form the discussion. Fortunately we can do it with the soft-emission technique [12, 25, 26].

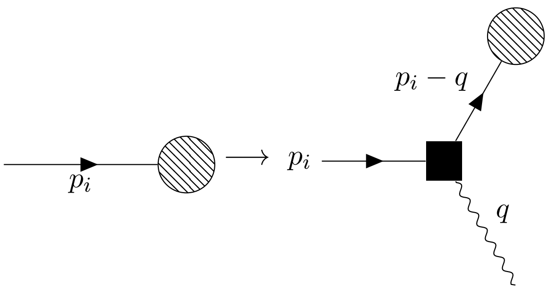
Consider a scattering from the multi-particle state $\alpha$ to an other multi-particle state $\beta$, with amplitude given by:




(2.4)

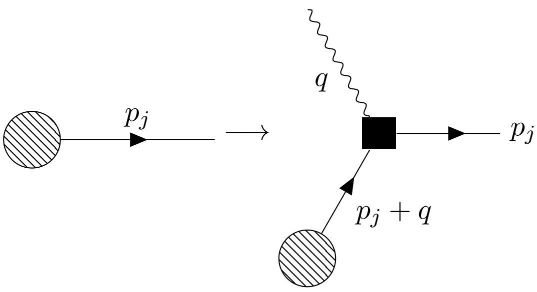
The in-particles with momenta $p_i$ and the out-particles with momenta $p_j$ can be either bosons or fermions.

Given this amplitude we now let an external leg emit a soft-gauge boson of spin $n$ with momentum $q$; with soft meaning that the energy of $q$ being much smaller than any other energy in the system. For this emission we have two possibilities: either a particle in the state $\alpha$ emits the gauge boson



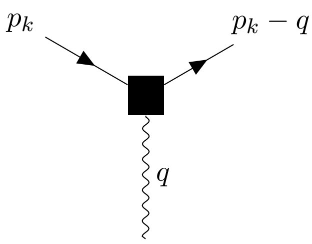

(2.5)

or a particle in the $\beta$ state emits it




(2.6)

We do not know exactly how's the interaction between the particles, so we represent it as a black blob:




(2.7)

The new amplitude for the emission of a single gauge boson through the $k$-th external leg is

$$\mathcal{M}_0 \to -i\epsilon_{\mu_1\cdots\mu_n}(q)\Gamma_k^{\mu_1\cdots\mu_n} \cdot (\text{propagator}) \cdot \mathcal{M}(p \pm q), \tag{2.8}$$

where this tensor quantity $\Gamma_k^{\mu_1\cdots\mu_n}$ that depends on the type of the particle $k$, has $n$ indices because we embedded the spin $n$ gauge boson in a $n-$rank tensor. Besides, Lorentz invariance puts several constrains on $\Gamma$. First, it must be a function of only $\frac{p_k\cdot q}{m_k^2}$, where $p_k$ is the momentum of the particle and $m_k$ its mass. Second, the indices may originate only from $p_k^\mu$, $q^\mu$ or Dirac structures in the case of fermions.

As said before we are interested in calculating the amplitude for the system to emit a gauge boson from one of its external legs and, if possible, to write it in terms of the original amplitude $\mathcal{M}_0$. For this we must sum over all possible emissions in the external legs. Considering first an emission through a bosonic in-external leg, the new amplitude is

$$\mathcal{M}_0 \to -i\Gamma_i^{\mu_1\cdots\mu_n}\epsilon_{\mu_1\cdots\mu_n}(q)\frac{i}{(p_i-q)^2-m_i^2}\mathcal{M}_0(p_i-q). \tag{2.9}$$

All these momenta are on-shell, so $p_i^2=m_i^2$, $q^2=0$ and $q^\mu\epsilon_{\mu\cdots\nu}(q)=0$. Now comes the soft-limit, with which we obtain

$$p_i\cdot q\approx 0, \qquad p_i-q\approx p_i, \tag{2.10}$$

therefore the tensor $\Gamma_i^{\mu_1\cdots\mu_n}$ becomes relevant only at 0. Moreover, $\Gamma_i^{\mu_1\cdots\mu_n}$ is contracted with a polarisation tensor, so the only relevant component of it is

$$\Gamma_i^{\mu_1\cdots\mu_n}(0)\to p_i^{\mu_1}\cdots p_i^{\mu_n}\tilde{\Gamma}_i(0). \tag{2.11}$$

The final amplitude is

$$-p_i^{\mu_1}\cdots p_i^{\mu_n}\epsilon_{\mu_1\cdots\mu_n}\tilde{\Gamma}_i(0)\frac{1}{p_i\cdot q}\mathcal{M}_0. \tag{2.12}$$

For a out-going particle the soft limit give us

$$p_j^{\mu_1}\cdots p_j^{\mu_n}\epsilon_{\mu_1\cdots\mu_n}\tilde{\Gamma}_j(0)\frac{1}{p_j\cdot q}\mathcal{M}_0. \tag{2.13}$$

Equations (2.12) and (2.13) are the same even if the particles were fermions. We can show this with a concrete example, for $n=1$, in which the amplitude (2.8) is given by [11, 12, 25]

$$-ie_k\epsilon_\mu(q)\bar{u}(p)\gamma^\mu\frac{i}{\not{p}\pm\not{q}-m_k}\mathcal{M}(p\pm q), \tag{2.14}$$

where we have suppressed the spin indices and $e_k$ is the (electric) charge of the particle. When we apply the soft limit, this expression becomes proportional to

$$\pm e_k\frac{1}{p\cdot q}\epsilon_\mu(q)p^\mu\bar{u}(p)\mathcal{M}(p)=\pm e_k\frac{1}{p\cdot q}\epsilon_\mu(q)p^\mu\mathcal{M}_0. \tag{2.15}$$

So we still obtain the same answer.

The final amplitude $\mathcal{M}$ of the system emitting a single soft gauge boson is*:

$$\mathcal{M}\approx\mathcal{M}_0\epsilon_{\mu_1\cdots\mu_n}(q)\Big\{\sum_j\tilde{\Gamma}_j(0)\frac{p_j^{\mu_1}\cdots p_j^{\mu_n}}{p_j\cdot q}-\sum_i\tilde{\Gamma}_i(0)\frac{p_i^{\mu_1}\cdots p_i^{\mu_n}}{p_i\cdot q}\Big\}. \tag{2.16}$$

One of the most important facts about gauge theories is the Ward identity; it ensure us that our amplitudes are gauge invariant, and so our cross-sections. We can check the Ward identity substituting the polarisation tensor by

$$\epsilon^{\mu_1\cdots\mu_n}(q)\to q^{\mu_1}\xi^{\mu_2\cdots\mu_n}+\cdots+q^{\mu_n}\xi^{\mu_1\cdots\mu_{n-1}}+\mathcal{O}(q^2), \tag{2.17}$$

*Only external leg emissions are relevant because of the singular factor of $p\cdot q$ in the denominator, which appears only for external legs.

where $\xi$'s are just functions characterising the gauge transformation. As this $\xi$'s are arbitrary, the equation

$$\sum_j p_j^{\mu_2} \cdots p_j^{\mu_n} \tilde{\Gamma}_j(0) = \sum_i p_i^{\mu_2} \cdots p_i^{\mu_n} \tilde{\Gamma}_i(0) \tag{2.18}$$

must hold if $\mathcal{M}$ is Lorentz invariant. Eq. (2.18) is just a conservation law for powers of momenta. If we have another condition on the components of the momenta in addition to the usual 4-momentum conservation, the only way to satisfy Eq. (2.18) is with $p^\mu \equiv 0$, for all momenta. This is clearly unacceptable, so we would rather have

$$\tilde{\Gamma}(0) = 0, \tag{2.19}$$

for all particles, meaning that there is no interaction between them and the gauge bosons. There are three cases where we can save the situation and still have interactions: spin zero, 1 and 2. Spin zero does not transform under Lorentz, so Eq. (2.18) makes no sense in this kind of theory. In a theory of spin 1, Eq. (2.18) give us the conservation of the quantity $\tilde{\Gamma}_k(0)$. In spin 2 we obtain

$$\sum_j p_j^{\mu} \tilde{\Gamma}_j(0) = \sum_i p_i^{\mu} \tilde{\Gamma}_i(0), \tag{2.20}$$

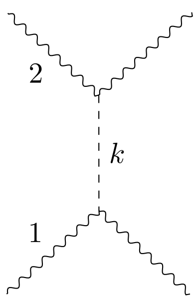
which is still problematic. But if we consider that $\tilde{\Gamma}_i(0) = \tilde{\Gamma}_j(0) = \tilde{\Gamma}(0) = \text{constant}$, we obtain just the 4-momentum conservation, so our theory is interacting and consistent. For other spins, though, nothing can be done.

One more remark on the fact that $\tilde{\Gamma}(0) = \text{constant}$ for spin 2 particles is in order. This means that every field coupled to gravity will have the same coupling strength, $\tilde{\Gamma}(0)$. In other words, the equivalence principle is automatically satisfied when considering a interacting spin-2 theory.

## 2.3 Spin 0

The simplest field is the scalar field that transforms under the $(0, 0)$ representation of the Lorentz group and therefore has only spin zero degree of freedom (d.o.f.). In non-relativistic situations we can model gravity as a scalar field with a massless Klein-Gordon Lagrangian, which gives us the same expression for the $1/r$ potential in Eq. (1.1), so it is not an absurd idea to consider first a scalar theory.

An argument to exclude the spin 0 possibility often seen in the literature is the following [1]. Gravity, of course, couples to mass as in Newton's theory, but it also couples to energy since the photon interacts with it. The quantity that represents all the energy and momentum of a theory is the energy-momentum tensor $T^{\mu\nu}$ [13], though we still do not know exactly how to define it. As explained in detail in Section 3.2, all matter will couple to gravity through its $T^{\mu\nu}$. If gravity were a scalar theory the propagator would be just $\frac{i}{k^2}$, with no Lorentz index. Imagining a diagram




(2.21)

with amplitude given by

$$(T_1)^{\mu}{}_{\mu}\frac{i}{k^2}(T_2)^{\nu}{}_{\nu}, \tag{2.22}$$

The propagator does not connect any Lorentz indices from 1 and 2, so the only way to make the amplitude Lorentz invariant is to take the traces. Of course we know that the energy-momentum tensor of the photon [10, 27],

$$T^{\mu\nu}_{\text{EM}} = F^{\mu\alpha}F_{\alpha\nu} - \frac{1}{4}\eta^{\mu\nu}F^{\alpha\beta}F_{\alpha\beta}, \tag{2.23}$$

is traceless, thus the amplitude is zero, contradicting our experimental observation. But the assumption that the amplitude goes with the energy-momentum tensor of the photon alone is an a posteriori conclusion. The only way we have to compute this tensor is with [12]

$$T^{\mu\nu} = \sum_n \frac{\partial \mathscr{L}}{\partial\, \partial_\mu\, \phi_n}\, \partial^\nu\, \phi_n - \eta^{\mu\nu}\, \mathscr{L}, \tag{2.24}$$

where the sum is over the fields present in the Lagrangian. In the case of a graviton-photon interaction the Lagrangian would be something like

$$\mathscr{L} = \mathscr{L}_{\text{graviton}} + \mathscr{L}_{\text{photon}} + \mathscr{L}_{\text{int}}\,.$$

To say that $T_1$ and $T_2$ are computed only from $\mathscr{L}_{\text{photon}}$ and not from the full Lagrangian is a consequence of a spin 2 theory alone as we will see in the next Chapters.

The correct way to exclude the spin zero theory is given in [4], where the authors calculate generically the scattering in Diagram (2.21) and use it to compute the deflection angle. This angle is given by

$$\Delta\varphi_{\text{scalar}} \propto \frac{4GM}{b^3}\left(\frac{2\left|\vec{\epsilon}\cdot\hat{b}\right|^2 - 1}{|\vec{p}|}\right), \tag{2.25}$$

with $\vec{b}$ the impact parameter, $M$ the mass of the massive body, $\vec{p}$ the photon's 3-momentum and $\epsilon_\mu$ the polarisation vector. Eq. (2.25) contradicts the observed deflection angle $\Delta\varphi$ in Eq. (2.21). Therefore gravity cannot be mediated by a scalar particle.

## 2.4 Spin 1

If the graviton were a vector field, gravity would be just like QED or QCD, which at first is no problem. The problem is that we already know the behaviour of both theories, in particular, their effect on matter and antimatter.

We conclude that the only possible interacting theory capable of describing gravity is one of spin 2. We now proceed in formulating this theory.

# 3 Linear Theory

## 3.1 Lorentz representation and quadratic Lagrangian

Lorentz representations are always obtained from Wigner's classification of the Lorentz group [22]. The case here considered is one of a particle with spin 2, so it contains 5 d.o.f. if massive and 2 if massless. The minimal Lorentz structure that can fit so many d.o.f. is a rank two tensor. A rank 2 tensor $T^{\mu\nu}$ has 16 d.o.f. and transforms as usual:

$$T^{\mu\nu} \to \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}T^{\alpha\beta}. \tag{3.1}$$

But we know that this representation is reducible, in the sense that we can split $T^{\mu\nu}$ in 3 different pieces which are also tensors. Every tensor $T^{\mu\nu}$ can be written as

$$T^{\mu\nu} = S^{\mu\nu} + F^{\mu\nu} + \frac{1}{4}\eta^{\mu\nu}T, \tag{3.2}$$

where $S$ and $F$ are the traceless symmetric and anti-symmetric pieces, respectively, and $T \equiv T^{\mu}_{\mu}$ is the trace of the tensor. Each one transforms according to the transformation rule in Eq. (3.1), hence are also tensors.

In terms of su(2)⊕su(2) classification we have [12, 22, 25]:

$$S^{\mu\nu} \sim (1,1) \tag{3.3}$$

$$F^{\mu\nu} \sim (1,0) \oplus (0,1) \tag{3.4}$$

$$T \sim (0,0) \tag{3.5}$$

The anti-symmetric plays no role here as it describe spin 1 theories. In abelian theories there is an unique quantity, $F_{\mu\nu}$, which is constructed out of first derivatives of the vector field and is also gauge invariant. In non-abelian case, the tensor $F^{a}_{\mu\nu}$ is not per se gauge invariant, but is still the only physical tensor constructed out of first derivatives [20]. They both form the kinetic term for their respective theories, $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ and $-\frac{1}{4}F^{a}_{\mu\nu}F^{\mu\nu a}$. So an antisymmetric tensor cannot describe a spin 2 theory and what's left is the symmetric (traceless or not) piece, that indeed contains spin 2 d.o.f.:

$$(1,1) = 2 \oplus 1 \oplus 0, \tag{3.6}$$

therefore our gravitational field is represented by a symmetric tensor.

We will call the graviton field $h_{\mu\nu}(x)$, a symmetric rank 2 tensor with 10 d.o.f.. The graviton is massless, so it does not have 5 d.o.f., but only 2. We need to somehow get rid of these extra d.o.f.. Before worrying with this problem, we must first build the Lagrangian that describe this field.

The most general massless Lagrangian one can build with two derivatives of $h_{\mu\nu}$ is

$$\mathscr{L} = a\,\partial^{\alpha}\,h_{\mu\nu}\,\partial_{\alpha}\,h^{\mu\nu} + b\,\partial_{\mu}\,h^{\mu\nu}\,\partial^{\alpha}\,h_{\alpha\nu} + c\,\partial_{\mu}\,h\,\partial_{\nu}\,h^{\mu\nu} + d\,\partial_{\alpha}\,h\,\partial^{\alpha}\,h, \tag{3.7}$$

where $a, b, c$ and $d$ are real numbers and $h$ is the trace of $h_{\mu\nu}$. How to determine these coefficients? There are many arguments that lead to the same answer. One particularly interesting is one of unitarity, i.e., that our theory preserves probability. First we decompose the Fock space and write $h_{\mu\nu}$ as the following

$$\begin{array}{ccccccccc} h_{\mu\nu} & = & h_{\mu\nu} & + & \partial_{\mu}\pi_{\nu} & + & \partial_{\nu}\pi_{\mu} & + & \partial_{\mu}\partial_{\nu}\pi \\ & & \downarrow & & \searrow & & \swarrow & & \downarrow \\ (1,1) & = & 2 & \oplus & & 1 & & \oplus & 0 \end{array}$$

That is, we write $h_{\mu\nu}$ in terms of a spin 2, a spin 1 and a spin 0 fields explicitly [12]. This is always possible since we still haven't done anything to remove the extra d.o.f..

Note that we have enough redundancy to choose $\partial_\mu h^{\mu\nu} = 0$ and $\partial_\mu \pi^\mu = 0$. Now we can insert this into Eq. (3.7) and obtain

$$\begin{aligned}\mathscr{L} &= a\,\partial_\alpha(h_{\mu\nu} + \partial_\mu \pi_\nu + \partial_\nu \pi_\mu + \partial_\mu \partial_\nu \pi)\,\partial^\alpha(h_{\mu\nu} + \partial_\mu \pi_\nu + \partial_\nu \pi_\mu + \partial_\mu \partial_\nu \pi) + \\ &\quad + b\,\partial_\mu(\partial_\mu \pi_\nu + \partial_\mu \partial_\nu \pi)\,\partial_\alpha(\partial_\alpha \pi_\nu + \partial_\alpha \partial_\nu \pi) + c\,\partial_\mu(h + \Box\pi)\,\partial_\nu \partial_\mu \partial_\nu \pi + \\ &\quad + d\,\partial_\alpha(h + \Box\pi)\,\partial_\alpha(h + \Box\pi) \\ &= a\,\partial^\alpha h_{\mu\nu}\,\partial_\alpha h^{\mu\nu} + b\,\partial_\mu h^{\mu\nu}\,\partial^\alpha h_{\alpha\nu} + c\,\partial_\mu h\,\partial_\nu h^{\mu\nu} + d\,\partial_\alpha h\,\partial^\alpha h + \\ &\quad + (2a + b)\pi_\mu \Box^2 \pi^\mu + (-a - b - c - d)\pi\Box^3\pi + (-c - 2d)h\Box^2\pi.\end{aligned}$$

The spin 2 part is identical to Eq. (3.7) but we notice the appearance of some dangerous, higher derivatives terms. Unitarity imposes that the propagator in momentum space cannot decrease faster than $\frac{1}{p^2}$, a result that follows from the Källén-Lehmann spectral representation [12, 20, 25]. So higher derivatives in the Lagrangian that would induce higher powers of momentum in the propagator, for example

$$\Box^2 \to \frac{1}{p^4},$$

are forbidden. We can use this to determine the coefficients of the Lagrangian:

$$\begin{aligned} 2a + b &= 0, \\ a + b + c + d &= 0, \\ c + 2d &= 0. \end{aligned}$$

Choosing $a = \frac{1}{2}$ we have the correct Fierz-Pauli Lagrangian [28]

$$\mathscr{L}_2 = \frac{1}{2}\,\partial^\alpha h_{\mu\nu}\,\partial_\alpha h^{\mu\nu} - \partial_\mu h^{\mu\nu}\,\partial^\alpha h_{\alpha\nu} + \partial_\mu h\,\partial_\nu h^{\mu\nu} - \frac{1}{2}\,\partial_\alpha h\,\partial^\alpha h. \tag{3.8}$$

This procedure guarantees that the Lagrangian (3.8) is unique, up to total derivatives.

With these coefficients the Lagrangian becomes invariant under the transformation

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \pi_\nu + \partial_\nu \pi_\mu, \tag{3.9}$$

which characterises a gauge transformation*. It is not a surprise since every massless theory is redundant [25]. From now on, we will refer to Eq. (3.9) as the gauge transformation of the graviton field.

⋆⋆⋆

Before proceeding to the next section, we could ask ourselves if a "strength field tensor" of the graviton field exists, just like $F^{\mu\nu}$ is for the photon field $A^\mu$. Here we understand such tensor as the one constructed from derivatives of the field that is also gauge invariant. These requirements come from the analogy to EM, in which the same conditions are satisfied, and will be modified when the theory becomes non-linear. It's straightforward to see that the only possibility is [9, 26]

$$R_{\alpha\beta\mu\nu} = \frac{1}{2M_P}\Big[\partial_\alpha \partial_\mu h_{\beta\nu} - \partial_\beta \partial_\mu h_{\alpha\nu} - \partial_\alpha \partial_\nu h_{\beta\mu} + \partial_\beta \partial_\nu h_{\alpha\mu}\Big], \tag{3.10}$$

with $M_P$ the Planck mass. This tensor has some symmetries in its indices: anti-symmetric in $\alpha\beta$ and $\mu\nu$, and symmetric in the exchange of the pairs $(\alpha\beta) \leftrightarrow (\mu\nu)$. We must also pay attention to the Lorentz representation of such tensor. In QED we have

$$A^\mu \sim \Big(\frac{1}{2}, \frac{1}{2}\Big) \to F^{\mu\nu} \sim (1, 0) \oplus (0, 1), \tag{3.11}$$

*Now $\partial_\mu \pi^\mu$ and $\partial_\mu h^{\mu\nu}$ are not necessarily zero. Also, $\pi$ here is dimensionless.

whereas for the gravitational field [20]

$$h_{\mu\nu} \sim (1,1) \to R_{\alpha\beta\mu\nu} \sim (2,0) \oplus (0,2). \tag{3.12}$$

One remarkable property of QED is the fact that it is possible to write the kinetic term only in terms of the strength field tensor. Once again using the electromagnetic analogy we expect to be able to write $\mathscr{L}_2$ in terms of $R_{\mu\nu\alpha\beta}$ only up to total derivatives. We arrive at

$$\mathscr{L}_2 = M_P\Big(h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h\Big)R^{\alpha}{}_{\mu\alpha\nu} \equiv M_P\bar{h}^{\mu\nu}R_{\mu\nu}, \tag{3.13}$$

with

$$R_{\mu\nu} = R^{\alpha}{}_{\mu\alpha\nu}, \tag{3.14}$$

$$\bar{h}^{\mu\nu} = h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h. \tag{3.15}$$

We note that there is a left over piece $\bar{h}_{\mu\nu}$ together with $R_{\mu\nu}$, so one may argue that we do not succeeded in rewriting $\mathscr{L}_2$ only in terms of $R_{\alpha\beta\mu\nu}$ and hence $R_{\alpha\beta\mu\nu}$ is not a strength field tensor as we thought. Strictly speaking yes, but we must face that we are dealing with a totally new interaction and we should expect some subtleties. In other words, analogies have its limits, whence let's not be hasty and discard a *very* good answer in E.q (3.13).

### 3.2 Equations of Motion and gauge fixing

Our situation now is very similar to Electrodynamics: we then had 4 components in $A_\mu$, from which only two have physical meaning. To remove this redundancy we used gauge invariance and transversality of on-shell photons to remove one d.o.f. each. We will proceed in the exact same way, but obviously with some technicalities.

Let us begin computing the EoM

$$\delta S_2[h_{\mu\nu}] = \delta \int d^4x \;\; \mathscr{L}_2 = 0. \tag{3.16}$$

We need to be careful in the variation of the action because the $\delta h_{\mu\nu}$ is symmetric, so only the symmetric part of the integrand needs to be zero. In the Lagrangian $\mathscr{L}_2$ we have only terms with derivatives, so

$$\int d^4x \;\; \partial_\mu \frac{\partial \mathscr{L}_2}{\partial\, \partial_\mu\, h_{\alpha\beta}} \delta h_{\alpha\beta} = 0, \tag{3.17}$$

where

$$\frac{\partial \mathscr{L}_2}{\partial\, \partial_\mu\, h_{\alpha\beta}} = \partial^\mu\, h_{\alpha\beta} - \eta^{\mu\beta}\,\partial_\sigma\, h^{\alpha\sigma} - \eta^{\mu\alpha}\,\partial_\sigma\, h^{\beta\sigma} + \eta^{\alpha\beta}\,\partial_\sigma\, h^{\mu\sigma} + \frac{1}{2}\Big[\eta^{\mu\beta}\,\partial^\alpha\, h + \eta^{\mu\alpha}\,\partial^\beta\, h\Big] - \eta^{\alpha\beta}\,\partial^\mu\, h,$$

is already symmetrized.

$$\Rightarrow \partial_\mu \frac{\partial \mathscr{L}_2}{\partial\, \partial_\mu\, h_{\alpha\beta}} = \Box h^{\alpha\beta} - \partial^\beta\,\partial_\sigma\, h^{\alpha\sigma} - \partial^\alpha\,\partial_\sigma\, h^{\beta\sigma} + \eta^{\alpha\beta}\,\partial_\mu\,\partial_\sigma\, h^{\mu\sigma} + \partial^\alpha\,\partial^\beta\, h - \eta^{\alpha\beta}\Box h = 0,$$

which can also be rewritten as

$$\Big[\Box h^{\alpha\beta} - \frac{1}{2}\eta^{\alpha\beta}\Box h\Big] - 2\Big[\partial^{(\beta}\,\partial_\sigma\, h^{\alpha)\sigma} - \frac{1}{2}\eta^{\alpha\beta}\,\partial_\mu\,\partial_\sigma\, h^{\mu\sigma}\Big] + \Big[\partial^\alpha\,\partial^\beta - \frac{1}{2}\eta^{\alpha\beta}\Box\Big]h = 0.$$

The above equation is not in its most practical presentation. In order to reduce it we introduce the following notation to simplify all calculations. Define the operations [1]

$$X^{(\mu\nu)} \equiv \frac{1}{2}\Big(X^{\mu\nu} + X^{\nu\mu}\Big), \tag{3.18}$$

$$\bar{X}^{\mu\nu} \equiv X^{(\mu\nu)} - \frac{1}{2}\eta^{\mu\nu}X. \tag{3.19}$$

The "bar" operation in Eq. (3.19) has some properties:

- Linear $\rightarrow \overline{(X+aY)}_{\mu\nu} = \bar{X}_{\mu\nu} + a\bar{Y}_{\mu\nu}$;
- $\bar{\eta}_{\mu\nu} = -\eta_{\mu\nu}$;
- $\bar{\bar{X}}_{\mu\nu} = X_{(\mu\nu)}$;
- $\bar{X} = -X$,

remembering we are denoting the trace of the tensor $X_{\mu\nu}$ simply by $X$. Using this notation we arrive at

$$\Box\bar{h}_{\alpha\beta} + \overline{\partial_\alpha\,\partial_\beta}h - 2\,\partial^\sigma\,\overline{\partial_{(\alpha}\,h}_{\beta)\sigma} = 0. \tag{3.20}$$

Now we may "bar" this expression, because all tensors we are working with are symmetric and the "bar" operation is involutive with respect to symmetric tensors. Hence

$$\Rightarrow \Box h_{\alpha\beta} - 2\,\partial^\sigma\,\partial_{(\alpha}\,\bar{h}_{\beta)\sigma} = 0. \tag{3.21}$$

Eq. (3.21) is the most compact form of the EoM.

We expect that our field can be decomposed into Fourier modes

$$h_{\mu\nu}(x) \sim \int \frac{d^3p}{\sqrt{2\omega_p}}\;\epsilon_{\mu\nu}(p)e^{ip\cdot x} + \text{h.c.}, \tag{3.22}$$

so that a plane wave solution is given by

$$h_{\mu\nu}(x) = \epsilon_{\mu\nu}(p)e^{\pm ip\cdot x}, \tag{3.23}$$

with the on-shell condition $p^2 = 0$. The tensors $\epsilon_{\mu\nu}(p)$ are the polarizations of the graviton with respect to the momentum $p$. Substituting back into the EoM we obtain

$$p^2\epsilon_{\alpha\beta}(p) - 2p^\sigma p_{(\alpha}\bar{\epsilon}_{\beta)\sigma}(p) = 0 \Rightarrow p^\sigma\bar{\epsilon}_{\sigma\alpha} = 0. \tag{3.24}$$

These are 4 equations constraining the propagation of on-shell gravitons. Next is the gauge-fixing, which could be something like the gauge Lorentz in analogy with EM. But, if we take a closer look to the EoM, we will see that we obtain a wave equation if we fix

$$\partial_\mu\,\bar{h}^{\mu\nu}(x) = 0, \tag{3.25}$$

also known as *harmonic gauge* [1, 12]. Eqs. (3.24) and (3.25) are in total eight equations, hence from the 10 original d.o.f. only 2 remain for on-shell gravitons, which correspond to the two physical propagating modes.

With the gauge fixed EoM we can calculate the propagator of the graviton. Consider a source $S_{\mu\nu}$ coupled to $h_{\mu\nu}$ in the following way:

$$S_{\text{int}} = \lambda\int d^4x\;h^{\mu\nu}(x)S_{\mu\nu}(x), \tag{3.26}$$

with $\lambda$ a coupling constant. This is the only way to couple a (tensorial) source to the graviton and, if $S_{\mu\nu}$ does not depend of the graviton field, it implies that $S_{\mu\nu}$ must be symmetric and conserved in the usual sense,

$$\partial^\mu\,S_{\mu\nu}(x) = 0, \tag{3.27}$$

because of gauge invariance. The EoM become then

$$\Box h_{\mu\nu} = \lambda\bar{S}_{\mu\nu} \tag{3.28}$$

and we want to find ${\Delta_{\mu\nu}}^{\alpha\beta}(p)$ such that

$$h_{\mu\nu}(p) = i\lambda {\Delta_{\mu\nu}}^{\alpha\beta}(p) S_{\alpha\beta}(p) \tag{3.29}$$

in momentum space. The momentum dependency is trivial, identical to the one in scalar theory, namely $\frac{i}{p^2}$. The tensor expression is more complicated, we need to discover what rank 4 tensor is equivalent to the action of the "bar" operation. We can easily calculate it through the derivative of a barred tensor with respect to its unbarred:

$$\delta \bar{X}^{\mu\nu} = \delta X_{\alpha\beta} \frac{\delta \bar{X}^{\mu\nu}}{\delta X_{\alpha\beta}} \tag{3.30}$$

$$\begin{aligned} \rightarrow \frac{\delta \bar{X}^{\mu\nu}}{\delta X_{\alpha\beta}} &= \frac{\delta}{\delta X_{\alpha\beta}} \Big[ X^{(\mu\nu)} - \frac{1}{2}\eta^{\mu\nu} X \Big] \\ &\equiv P^{\mu\nu\alpha\beta}, \end{aligned}$$

where we've defined

$$P^{\mu\nu\alpha\beta} = \frac{1}{2}\Big[\eta^{\mu\alpha}\eta^{\nu\beta} + \eta^{\mu\beta}\eta^{\nu\alpha} - \eta^{\alpha\beta}\eta^{\mu\nu}\Big]. \tag{3.31}$$

It is now straightforward to see that

$$P_{\mu\nu\sigma\lambda}P^{\sigma\lambda\alpha\beta} = \frac{1}{2}\Big[\eta^{\alpha}{}_{\mu}\eta^{\beta}{}_{\nu} + \eta^{\alpha}{}_{\nu}\eta^{\beta}{}_{\mu}\Big], \tag{3.32}$$

i.e., that the bar operation is involutive with respect to symmetric tensors.

The final form of the graviton propagator in momentum space and in harmonic gauge is

$$\Delta^{\mu\nu\alpha\beta}(p) = \frac{i}{p^2 + i\epsilon} P^{\mu\nu\alpha\beta}. \tag{3.33}$$

⋆⋆⋆

What is the physical meaning of $S_{\mu\nu}$? Based on our experimental facts we know that the graviton couples to energy and the rank 2 tensor that coincidentally represents the total energy and momentum of a system is the energy-momentum tensor $T_{\mu\nu}$, which has dimension 4. The conserved charges, $Q_\mu$, associated with such tensor are

$$Q_\mu = \int d^3x \; T_{\mu 0}. \tag{3.34}$$

These charges are none other than the 4-momentum itself, since the definition of $T_{\mu\nu}$ is the flux of $p_\mu$ across surface of constant $x_\nu$ [13]. But for a theory to conserve 4-momentum it needs to be not only Lorentz, but also Poincaré invariant [12, 25]. So, if $\xi$ is a constant four vector,

$$\frac{\delta \mathscr{L}}{\delta \xi^\mu} = \partial_\mu \mathscr{L} = \sum_n \frac{\partial \mathscr{L}}{\partial \phi_n}\frac{\delta \phi_n}{\delta \xi^\mu} = \partial_\nu \Big( \sum_n \frac{\partial \mathscr{L}}{\partial\, \partial_\nu\, \phi_n}\, \partial_\mu\, \phi \Big), \tag{3.35}$$

where we have used the EoM. Hence the conserved tensor due to translational invariance is in fact the same energy-momentum tensor defined in Eq. (2.24).

Another important point is the coupling $\lambda$. Since $S_{\mu\nu}$ is identified with $T_{\mu\nu}$, by dimensional analysis we conclude that $[\lambda] = -1$. Relevant to gravitational phenomena is the Planck mass $M_P$, which is the fundamental scale of gravity, therefore $\lambda$ must be proportional to $M_P^{-1}$.

In short, from now on, we need to pay attention to three points:

- the energy-momentum tensor is intrinsically connected with the Poincaré group;
- from gauge transformations of the graviton, the introduced energy-momentum tensor must be conserved;
- The dimensionfull constant $\lambda$ has dimension -1 and is proportional to $\frac{1}{M_P}$.

# 4 Non-Linear Theory

## 4.1 Why non-linearity?

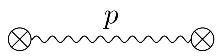
At first glance, the free theory looks just fine and it remains only to couple it with some matter field to make experimental predictions. However, we must not overlook the phenomenology of the gravitational interaction previously discussed, in particular the one that determines the potential between two sources. It was given by Eq. (1.1)

$$V(r) = -\frac{GM}{r} + \mathcal{O}\left(\frac{1}{r^2}\right),$$

where $M$ is the mass of the source (here we are considering the dynamics of stellar bodies, for example).

The machinery of Feynman diagrams help us understand how non-linearity emerges and why it solves the issue with $V(r)$. Take for example Mercury's perihelion precession. The first order, Fourier transformed potential between the Sun and Mercury is given by the following diagram

$$iM_m\tilde{V}_1(p) = \qquad (4.1)$$




where $M_m$ is the mass of Mercury. The crossed dots in the above diagram means *current insertions*, in other words something that generate gravitons [11, 12]. In our example a classical source of graviton are masses, hence the insertion of $M_\odot$, the mass of the Sun, generates a graviton with momentum $p$ and the other mass, $M_m$, absorbs it. In terms of the energy density $\rho$, these stationary sources are given by

$$\rho(\vec{x}) = M\delta^3(\vec{x} - \vec{y}), \tag{4.2}$$

with $\vec{y}$ the location of the mass $M$. Therefore, the only non-vanishing component of the energy-momentum tensor is the 00 component [6], thus, according to Eq. (3.33), the amplitude in Eq. (4.1) is given by

$$iM_m\tilde{V}_1(p) = \frac{1}{M_P^2}\tilde{T}_\odot^{00}\tilde{T}_m^{00}P_{0000}\frac{i}{p^2} = \frac{M_\odot}{M_P}\frac{M_m}{M_P}\frac{i}{p^2}, \tag{4.3}$$

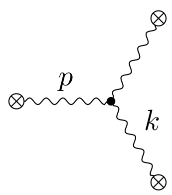
where we introduced the $\frac{1}{M_P^2}$ factor because of dimensional analysis. The above potential in coordinate space is indeed the Newtonian potential

$$V_1(r) = -\frac{GM_\odot}{r}, \tag{4.4}$$

where we have identified $G = \frac{1}{M_P^2}$.

With the theory $\mathscr{L}_2$ alone we can never generate a term that falls off like $1/r^2$. This is obvious from the diagrammatic representation in Eq. (4.1). The only way is to admit the idea of self-interactions between gravitons, which is not an absurd one since gravitons can carry energy and momentum. By carrying energy they induce an energy-momentum tensor and therefore couple via a term $\sim hT$, which is non-quadratic. We are not interested in the explicit form of this $T^{\mu\nu}$; here it will be enough to use just dimensional analysis and physical reasoning. By self interactions we mean that gravitons interact with each other, so if the source $M_\odot$ emit two distinct gravitons, they may first interact with each other and then arrive at $M_m$. Diagrammatically

$$iM_m\tilde{V}_2(p,k) = \qquad . \qquad (4.5)$$

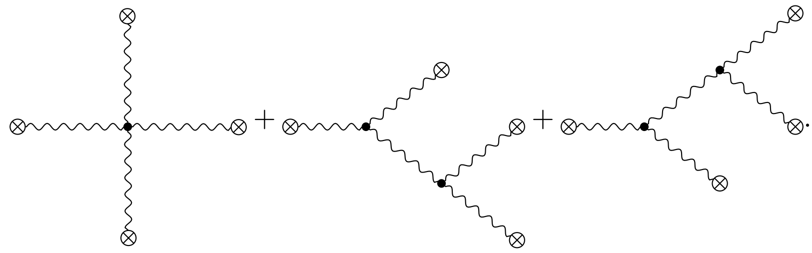


At the right side of the diagram are the two gravitons emitted from the Sun. The vertex gives the same factor of $\lambda$ as in Eq. (3.26), while the rest is straightforwardly given by

$$V_2(r) = +\frac{\lambda}{M_P^3}\frac{M_\odot^2}{r^2} \propto \frac{G^2 M_\odot^2}{r^2}. \tag{4.6}$$

The potential in Eq. (4.6) is exactly the one predicted by GR if the proportionality constant is 1 [6, 13]. Other theories of gravitation, like Brans-Dicke theory, predict the same potential structure but with distinct proportionality factors [6]. As experimental tests of this phenomena favours GR, the coupling constant of gravitation is given by $\lambda = M_P^{-1}$.

If we wanted to calculate even higher order corrections to the Newtonian potential, we would need to calculate the following amplitude

$iM_m\tilde{V}_3 =$  (4.7)

Note that we've included a quartic interaction, which we have no reason to not include. The correct interactions must be calculated from Eq. (3.26), i.e. to calculate the energy-momentum tensor explicitly, something we will always avoid in this paper. Another remark on the above amplitude is that it is of order $G^3$ and as none of our present experiments probed any effects of this order, it's meaningless to calculate $\tilde{V}_3$ [6, 12].

The conclusion here is simple: a theory of gravitation must be non-linear, i.e. must include graviton self-interactions. With the perihelion shift of Mercury we can infer the 3-point graviton vertex, but nothing more. We still do not know if there is higher dimensional operators in the Lagrangian or if it stops at cubic order. Next we shall determine how much non-linear gravity is.

## 4.2 First consistency check

In this section we invoke the sixth question we've made in the first Chapter: is our theory consistent? Based on the previous section we have realised that $\mathscr{L}_2$ alone is not a consistent theory, because we know from observation that the classical action must introduce higher order operators. But the situation is a bit more subtle as noted firstly by Gupta [29] and the later again by Deser [3].

A free graviton described by $\mathscr{L}_2$ carry energy and momentum, so it has a non-vanishing energy-momentum tensor $T_2^{\mu\nu}$. As stated by Eq. (3.26), the existence of such tensor induces a self-coupling

$$S_3 = \lambda \int d^4x \; h_{\mu\nu} T_2^{\mu\nu}, \tag{4.8}$$

which is, based in Eq. (2.24), a cubic Lagrangian in $h$. The new interacting theory of a graviton is now given by

$$S = S_2 + S_3. \tag{4.9}$$

The punchline is: $S_3$ also carries energy and momentum. Since the gravitons now interact, they have energy and momentum which were not considered in $T_2^{\mu\nu}$. So we need to once again compute an energy-momentum tensor, this time from $S_3$. This gives a new tensor $T_3^{\mu\nu}$ that couples as

$$S_4 = \lambda \int d^4x \; h_{\mu\nu} T_3^{\mu\nu}. \tag{4.10}$$

This term is quartic and of order $\lambda^2$, because $S_3$ was of order $\lambda$. And so on, every new piece of the action contains new energy and momentum which were not considered by other tensors, hence induce new higher order couplings. It is then obvious that this procedure never stops, i.e. we will need to compute an infinite series of energy-momentum tensors. Note that only then the theory becomes consistent, because all self-interactions of the graviton would be considered.

More precisely formulated we have the following. Consider two functions: $\eta$, that compute the energy-momentum tensor of a given action; and $\kappa$, that takes an energy-momentum tensor and couples it to the graviton via Eq. (3.26). Note that we do not assume that $\eta$ is given by Eq. (2.24) for reasons that will become clear in the next Chapters, this means we are not yet interested in its explicitly form. The composition of these functions give

$$(\kappa \circ \eta)(S_2) = S_3, \tag{4.11}$$

$$\Rightarrow (\kappa \circ \eta)^n(S_2) = S_{2+n}. \tag{4.12}$$

From the definition, both functions are linear, so we may instead apply them in the following way

$$(\kappa \circ \eta)(S_2 + \cdots + S_n) = S_3 + \cdots + S_{n+1}. \tag{4.13}$$

Since the information of $S_2$ is lost on the right hand side, we define a new function $\xi$ by

$$\xi(\cdot) = S_2 + (\kappa \circ \eta)(\cdot), \tag{4.14}$$

hence

$$\xi\left(\sum_{n=2}^{N} S_n\right) = \sum_{n=2}^{N+1} S_n. \tag{4.15}$$

Composing $\xi$ $N$ times over $S_2$ we obtain

$$(\xi \circ \cdots \circ \xi)(S_2) = \sum_{n=2}^{N+2} S_n \xrightarrow[N\to\infty]{} S = \sum_{n=2}^{\infty} S_n, \tag{4.16}$$

and naturally $S$ is the final action of the gravitational interaction. It has the important fixed-point property

$$\xi(S) = S, \tag{4.17}$$

which guarantees us a kind of uniqueness of $S$, but nothing that really aids us to find the functional form of $S$. At the same time, Eq. (4.17) ensures us that $S$ is the final action sought, because it means that every energy and momentum of the graviton are considered in the theory, i.e., our theory is consistent. Performing such calculations is a hopeless job; every derivation in the literature of GR starting from QFT uses some shortcuts biased on ideas of GR to not perform the sum $S$ [1, 3, 4], or used some other argumentation that did no allow to calculate $S$ explicitly [2, 5].

We won't make this sum either, but we can use our knowledge of QFT to infer the form of the final action. Back in Section 3.2 we found a compact form for the quadratic action:

$$S_2 = \frac{1}{\lambda^2} \int d^4x \left(\eta^{\mu\nu} + \lambda \bar{h}^{\mu\nu}\right) R^{\alpha}{}_{\mu\alpha\nu}, \tag{4.18}$$

where we have added a total derivative $\eta R$ term*. The statement is that the final form of $S$ is actually the same as of $S_2$, in other words, the non-linear $R^{\alpha}{}_{\mu\beta\nu}$ and $\eta^{\mu\nu} + \lambda \bar{h}^{\mu\nu}$ are relevant physical quantities and not just random definitions. The basis of our argumentation is an analogy with $SU(N)$ gauge-theories. Take for instance the only linear rank 2 tensor constructed out of spin 1 fields and its derivatives invariant under $U(1)$ gauge

*The justification of such artificial step will become clear in the next section.

transformation, $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. This tensor is analogous to $R_{\mu\nu\alpha\beta}$, in the sense that it is linear, gauge invariant and unique. If we upgrade $U(1)$ to a $SU(N)$ group, just $F^{\mu\nu}$ isn't adequate to describe the theory anymore; it must become non-linear

$$F_{\mu\nu} \to F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu. \tag{4.19}$$

Despite losing gauge invariance and linearity, $F^a_{\mu\nu}$ still is the fundamental quantity in Yang-Mills theory, because it not only describes free-propagation of gluons but also their self-interactions, which make the theory consistent [11, 20]. As consequence, the Yang-Mills Lagrangian may be written only in terms of it. In the case of gravity the insertion of self-interactions will break the linearity and gauge invariance of $R_{\mu\nu\alpha\beta}$, but since it is the relevant physical quantity of the theory, the Lagrangian will still look like (3.7), just as in Yang-Mills theory.

Not only $R_{\mu\nu\alpha\beta}$ should be modified by non-linearity, but also the other term in the Lagrangian, $\eta_{\mu\nu} + \lambda\bar{h}_{\mu\nu}$, should also become non-linear, so that the Lagrangian $S$ will be written in the following way

$$S = \frac{1}{\lambda^2}\int d^4x \text{ (Non-linear } \eta^{\mu\nu} + \lambda\bar{h}^{\mu\nu})(\text{Non-linear } R^\alpha{}_{\mu\alpha\nu}). \tag{4.20}$$

We must not forget that Yang-Mills theory, in contrast to $U(1)$ gauge theories, has non-linear gauge transformations. Therefore Eq. (3.9) will also be modified by the non-linearity of the theory. The determination of such transformation is the key point: if we compute the infinitesimal (first order) transformation of the graviton in the full non-linear regime, we will be able to determine what is the new symmetry group and then construct explicitly the non-linear version of Eq. (4.16). This procedure is the same for Yang-Mills theories, but naturally in such theories only terms up to the quartic order appear, nothing like the infinite non-linearity of gravity, hence the nature of the new group of transformations will be much more wild and complicated than a gentle $SU(N)$.

## 4.3 Second consistency check and Covariant Derivatives

In Section 3.1, the construction of the interacting Lagrangian in Eq. (3.26) together with gauge transformation of Eq. (3.9) implied the usual conservation of the tensorial source $S_{\mu\nu}$

$$\partial_\mu S^{\mu\nu} = 0. \tag{4.21}$$

As the same thing happens for the electromagnetic current $j^\mu$, one could think that there is no issue with Eq. (4.21) and that indeed we should have conserved energy-momentum tensors. That is not the case, as noted by Weinberg and Witten in [30], where they prove that a theory with a massless spin 2 particle cannot have a conserved energy-momentum tensor. This theorem, known as Weinberg-Witten theorem for spin 2 [12], justifies why we did not wanted to use Eq. (2.24) for the energy-momentum tensor at all, else our calculations would be automatically inconsistent. One should not be surprised by this outcome. Following the analogy between abelian and non-abelian gauge theories, from the moment that the theory acquires a global non-abelian symmetry, under which massless spin 1 particles are charged (i.e. they self interact), no gauge invariant conserved current can possibly exist*. This is clear in Yang-Mills theories, where the EoM imply not conservation, but covariant conservation†

$$D_\mu F^a_{\mu\nu} = j^a_\nu \Rightarrow D_\nu j^a_\nu = 0. \tag{4.22}$$

*Also known as Weinberg-Witten theorem for spin 1.

†Note we are sloppy with up-down index notation, leaving implicit that the contraction of repeated indices is made in the usual Lorentz invariant way no matter their upper/lower position.

So taking the Weinberg-Witten theorem for spin 2 into consideration, we must ask ourselves once again: is our theory consistent? Our theory is consistent as long as we determine *how* the energy-momentum tensor is conserved, if it is in some sense. Obviously, $T^{\mu\nu}$ should be covariantly conserved, because the existence of a strength field tensor implies the existence of covariant derivatives, although we still do not know what exactly covariant here means.

It is well know that the most fundamental definition of a strength field tensor is [11, 12, 13, 14]

$$[D_\mu, D_\nu] = (\text{Field Strength Tensor}), \tag{4.23}$$

where $D_\mu$ is the covariant derivative. We have already determined $R_{\mu\nu\alpha\beta}$ from alternative argumentation, therefore we are in position to actually determine the covariant derivative $D_\mu$. Take a scalar field $\phi$ and apply the commutator of covariant derivatives

$$[D_\mu, D_\nu]\phi = \,? \tag{4.24}$$

On the right-hand side we must obtain $\sim \phi R_{\mu\nu\alpha\beta}$ with only two free indices*, hence the only possibility is

$$[D_\mu, D_\nu]\phi \sim \phi R^{\alpha}{}_{\alpha\mu\nu} = 0, \tag{4.25}$$

because of the antisymmetric property of $R_{\mu\nu\alpha\beta}$. The term $\phi R_{\mu\nu}$ is not considered because $R_{\mu\nu}$ (the Ricci tensor) is symmetric in $\mu\nu$, while the covariant derivative is antisymmetric. We cannot then determine anything from a scalar field; we choose instead a vector field $V_\alpha$. The commutator of covariant derivatives, as the definition of a strength field tensor, should give *all* the information contained in $R_{\mu\nu\alpha\beta}$, therefore on the right-hand side of Eq. (4.23) $R_{\mu\nu\alpha\beta}$ must have no contracted indices†. The only possibility, while taking into account the symmetry properties of the strength field tensor, is

$$[D_\mu, D_\nu]V_\alpha \sim V_\sigma R_{\sigma\alpha\mu\nu}. \tag{4.26}$$

Before proceeding, we note something quite odd. In gauge theories, be it a $U(1)$ or a Yang-Mills one, the covariant derivative is a very well determined object, in the sense that we may write it as [11, 12]

$$D_\mu^{\text{YM}} = \partial_\mu - igT^a A_\mu^a. \tag{4.27}$$

The covariant derivative $D_\mu^{\text{YM}}$ is given by Eq. (4.27) regardless if applied on a Lorentz scalar, vector or tensor. In the case here considered we see that this is not true; if we state that the covariant derivative we want to compute is given by

$$D'_\mu = \partial_\mu + \xi_\mu, \tag{4.28}$$

with a fixed field $\xi_\mu$, we will obtain on the one hand that

$$[\partial_\mu \xi_\nu - \partial_\nu \xi_\mu]\phi = 0 \tag{4.29}$$

by Eq. (4.25) and on the other hand that

$$[\partial_\mu \xi_\nu - \partial_\nu \xi_\mu]V_\alpha \sim V_\sigma R_{\sigma\alpha\mu\nu}, \tag{4.30}$$

by Eq. (4.26). The above equations will clearly result in distinct structures for $\xi_\mu$, contradicting our hypothesis in Eq. (4.28). In other words, the covariant derivative we ought to determine depends on the Lorentz structure of the object it is applied to, else there is no

*We do not introduce terms with derivatives of the fields, since there are no such terms in gauge theories.

†The scalar field is obviously an exception to this rule, but it will become clear that the covariant derivative of a scalar field is just an ordinary derivative, so there is indeed no information on the strength field tensor.

way to have Eq.s (4.25) and (4.26) simultaneously. As noted in [20], the analogy between gravity and Yang-Mills theories breaks down completely at this point and the only way to determine the covariant derivative in the general case is by means of GR techniques.

Notwithstanding we will proceed with a bit more care, in the sense that we should also note that Eq. (4.26) implies that the connection "$\xi_\mu$" of the covariant derivative mixes the Lorentz indices, because $V_\sigma$ appears in the right-hand side of Eq. (4.26) and not $V_\alpha$ as in the left-hand side. Hence, at least for a vector, we can write the covariant derivative generically as

$$\nabla_\mu V_\nu = \partial_\mu V_\nu - \Gamma^\sigma_{\mu\nu} V_\sigma, \tag{4.31}$$

where $\Gamma^\sigma_{\mu\nu}$ is the connection*. With the same line of thought it becomes clear now that since a scalar field has no Lorentz indices, then the covariant derivative is just the usual derivative:

$$\nabla_\mu \phi = \partial_\mu \phi. \tag{4.32}$$

From Eq. (4.32) we find the first information on the connection $\Gamma$. Expanding Eq. (4.25):

$$[\nabla_\mu, \nabla_\nu]\phi = \nabla_\mu(\partial_\nu \phi) - \nabla_\nu(\partial_\mu \phi) = \left(\Gamma^\alpha_{\mu\nu} - \Gamma^\alpha_{\nu\mu}\right)\partial_\alpha \phi = 0 \tag{4.33}$$

$$\Rightarrow \Gamma^\alpha_{\mu\nu} = \Gamma^\alpha_{\nu\mu}. \tag{4.34}$$

For a rank two tensor the situation is a bit more complicated, but also straightforward. The commutator of covariant derivatives applied to a rank two tensor is given in general by three terms

$$[\nabla_\mu, \nabla_\nu]T_{\alpha\beta} \sim aTR_{\alpha\beta\mu\nu} + bT_{\alpha\sigma}R_{\sigma\beta\mu\nu} + cT_{\sigma\beta}R_{\sigma\alpha\mu\nu}, \tag{4.35}$$

with $a$, $b$ and $c$ constants, and $T \equiv T^\sigma_{\ \sigma}$. The term with the trace of $T_{\alpha\beta}$ must vanish. To see this consider two cases, one with a $T_{\alpha\beta}$ which is symmetric and the other with a antisymmetric tensor. For the first case, since on the right-hand side of Eq. (4.35) $R_{\alpha\beta\mu\nu} = -R_{\beta\alpha\mu\nu}$ and on the left-hand side $T_{\alpha\beta} = T_{\beta\alpha}$, we must have $a = 0$. For an antisymmetric tensor, this term is obviously zero, since $T$ in this case is zero. The coefficients $b$ and $c$ must be equal by similar argumentation. Eq. (4.35) becomes then

$$[\nabla_\mu, \nabla_\nu]T_{\alpha\beta} \sim T_{\alpha\sigma}R_{\sigma\beta\mu\nu} + T_{\sigma\beta}R_{\sigma\alpha\mu\nu}. \tag{4.36}$$

With the same reasoning as for a vector field, the connection of a covariant derivative applied to a tensor must have the following structure

$$\nabla_\mu T_{\alpha\beta} = \partial_\mu T_{\alpha\beta} - \gamma^\sigma_{\mu\alpha} T_{\sigma\beta} - \gamma^\sigma_{\mu\beta} T_{\alpha\sigma}, \tag{4.37}$$

where $\gamma^\sigma_{\nu\mu} = \gamma^\sigma_{\mu\nu}$ because of symmetry. The coefficients of both terms must also be equal by symmetry. Note that we did not write $\gamma = \Gamma$; they are obviously related in some way, but we cannot guarantee that they are the one and the same.

Let us compute the commutator in Eq. (4.26), which is now written as

$$\begin{aligned}
[\nabla_\mu, \nabla_\nu]V_\alpha &= \partial_\mu(\nabla_\nu V_\alpha) - \gamma^\sigma_{\mu\alpha}\nabla_\nu V_\sigma - \gamma^\sigma_{\mu\nu}\nabla_\sigma V_\alpha - (\mu \leftrightarrow \nu) \\
&= \partial_\mu(\partial_\nu V_\alpha - \Gamma^\sigma_{\nu\alpha}V_\sigma) - \gamma^\sigma_{\mu\alpha}\left(\partial_\nu V_\sigma - \Gamma^\lambda_{\nu\sigma}V_\lambda\right) - \gamma^\sigma_{\mu\nu}\left(\partial_\sigma V_\alpha - \Gamma^\lambda_{\sigma\alpha}V_\lambda\right) - \\
&\quad -(\mu \leftrightarrow \nu) \\
&= \left(\partial_\nu \Gamma^\sigma_{\mu\alpha} - \partial_\mu \Gamma^\sigma_{\nu\alpha}\right)V_\sigma + \left(\Gamma^\sigma_{\mu\alpha} - \gamma^\sigma_{\mu\alpha}\right)\partial_\nu V_\sigma - \left(\Gamma^\sigma_{\nu\alpha} - \gamma^\sigma_{\nu\alpha}\right)\partial_\mu V_\sigma + \mathcal{O}(\Gamma\gamma).
\end{aligned}$$

*The minus sign was chosen just to match the definition from GR, but is totally arbitrary since $\Gamma$ is still unknown.

Note that the higher order terms $\sim \Gamma\gamma$ were rejected, because we have only the first order $R_{\sigma\alpha\mu\nu}$ to compare with. Either way, these corrections would enter in the term proportional to $V_\sigma$ in the following way

$$\partial_\nu \Gamma^\sigma_{\mu\alpha} - \partial_\mu \Gamma^\sigma_{\nu\alpha} + \mathcal{O}(\Gamma\gamma) = \partial_\nu \Gamma^\sigma_{\mu\alpha} - \partial_\mu \Gamma^\sigma_{\nu\alpha} + \gamma^\lambda_{\nu\alpha}\Gamma^\sigma_{\lambda\mu} - \gamma^\lambda_{\nu\alpha}\Gamma^\sigma_{\lambda\mu}, \tag{4.38}$$

hence they do not interfere with the terms proportional to derivatives of $V_\sigma$. Looking at Eq. (4.26) we find no such terms with derivatives of the field, therefore they must vanish. The only way to do so is to impose

$$\gamma = \Gamma. \tag{4.39}$$

Finally, we may write from Eqs. (4.26), (4.38) and (4.39)

$$\partial_\nu \Gamma^\sigma_{\mu\alpha} - \partial_\mu \Gamma^\sigma_{\nu\alpha} + \Gamma^\lambda_{\nu\alpha}\Gamma^\sigma_{\lambda\mu} - \Gamma^\lambda_{\nu\alpha}\Gamma^\sigma_{\lambda\mu} = AR_{\sigma\alpha\mu\nu}, \tag{4.40}$$

with $A$ a constant. Note that

$$R_{\sigma\alpha\mu\nu} \sim \mathcal{O}(\partial^2 h) \Rightarrow \partial\,\Gamma \sim \mathcal{O}(\partial^2 h), \tag{4.41}$$

whence

$$\Gamma\Gamma \sim \mathcal{O}\Big((\partial h)^2\Big), \tag{4.42}$$

so we may indeed neglect the $\Gamma\Gamma$ terms. The equation we must solve is

$$\partial_\nu \Gamma^\sigma_{\mu\alpha} - \partial_\mu \Gamma^\sigma_{\nu\alpha} = AR_{\sigma\alpha\mu\nu}. \tag{4.43}$$

First we redefine $\Gamma \to -A\Gamma$ to get rid of the constant. Second we note that the expression is antisymmetric in $\mu\nu$, hence

$$\epsilon_{\mu\nu\lambda\kappa}[2\,\partial_\mu \Gamma^\sigma_{\nu\alpha} + R_{\sigma\alpha\mu\nu} + \text{symmetric terms}] = 0, \tag{4.44}$$

where $\epsilon$ is the Levi-Civita tensor. Next substitute Eq. (3.10)

$$\epsilon_{\mu\nu\lambda\kappa}\left[2\,\partial_\mu \Gamma^\sigma_{\nu\alpha} + \frac{\lambda}{2}(\partial_\sigma \partial_\mu h_{\alpha\nu} - \partial_\alpha \partial_\mu h_{\sigma\nu} - \partial_\sigma \partial_\nu h_{\alpha\mu} + \partial_\alpha \partial_\nu h_{\sigma\mu}) + \text{symmetric terms}\right] = 0, \tag{4.45}$$

which due to the antisymmetrization is also equal to

$$\epsilon_{\mu\nu\lambda\kappa}[2\,\partial_\mu \Gamma^\sigma_{\nu\alpha} + \lambda(\partial_\sigma \partial_\mu h_{\alpha\nu} - \partial_\alpha \partial_\mu h_{\sigma\nu}) + \text{symmetric terms}] = 0. \tag{4.46}$$

Isolating the derivative in $x^\mu$:

$$\epsilon_{\mu\nu\lambda\kappa}\,\partial_\mu\,[2\Gamma^\sigma_{\nu\alpha} + \lambda(\partial_\sigma h_{\alpha\nu} - \partial_\alpha h_{\sigma\nu}) + \text{symmetric terms} + \text{constant term}] = 0, \tag{4.47}$$

but since the fields fall to zero at infinity, such constant term is zero. The symmetric term is one that contracted by the Levi-Civita together with $\partial_\mu$ vanishes, i.e. must be $\partial_\nu h_{\alpha\sigma}$ (remember Eq. (4.41)). Considering also the symmetry of Eq. (4.34), the final form of the connection is

$$\Gamma^\sigma_{\alpha\nu} = \frac{\lambda}{2}(\partial_\alpha h_{\sigma\nu} + \partial_\nu h_{\sigma\alpha} - \partial_\sigma h_{\alpha\nu}). \tag{4.48}$$

⋆⋆⋆

One remark before proceeding regarding the up and down index notation. As a field theorist, one is used to write every index up or down, because it does not make any difference as long as attention is paid. Here we can already notice that the index position is not *so* arbitrary. Take for example the following object:

$$\nabla_\mu(V_\alpha V^\alpha). \tag{4.49}$$

As usual $V^2$ is a scalar, hence the covariant derivative is just the usual derivative, according to Eq. (4.32). At the same time

$$\nabla_\mu (V_\alpha V^\alpha) = V_\alpha \nabla_\mu V^\alpha + V^\alpha \nabla_\mu V_\alpha, \tag{4.50}$$

which follows directly from Eq.s (4.31), (4.36) and (4.37). Explicitly we can write

$$V_\alpha \nabla_\mu V^\alpha + V^\alpha \Big( \partial_\mu V_\alpha - \Gamma^\sigma_{\mu\alpha} V_\sigma \Big) = \partial_\mu \left( V^2 \right). \tag{4.51}$$

The equation above is only satisfied if the covariant derivative of a contravariant vector has covariant derivative given by

$$\nabla_\mu V^\alpha = \partial_\mu V^\alpha + \Gamma^\alpha_{\sigma\mu} V^\sigma, \tag{4.52}$$

which is almost identical to Eq. (4.31), but with a plus sign. In conclusion, when applying covariant derivatives, we should be careful with up and down indices, whereas for other cases we may use the standard sloppy notation.

## 4.4 Diffeomorfism Group

In this last section we will use the result of Eq. (4.48) to compute explicitly what is the new transformation of the graviton and then determine the new symmetry group.

Previously, we stated that transformation in Eq. (3.9) will be modified to the more general expression

$$h_{\mu\nu} \to h_{\mu\nu} + \delta h_{\mu\nu} \tag{4.53}$$

due to the non-linear nature of gravity. To determine such transformation we will use once again the Weinberg-Witten theorem. If not conserved, the energy-momentum tensor should then be covariantly conserved. As we already determined the covariant derivative of a rank 2 tensor, we are in position to do such calculation.

Consider a gauge invariant (i.e. that does not depend on the graviton itself) energy-momentum tensor $T^{\mu\nu}$ coupled to the graviton

$$S' = \lambda \int d^4x \ h_{\mu\nu} T^{\mu\nu}. \tag{4.54}$$

After a transformation given by Eq. (4.53), the action $S'$ remains the same, therefore

$$\delta S' = \lambda \int d^4x \ \delta h_{\mu\nu} T^{\mu\nu} = 0. \tag{4.55}$$

The transformation $\delta h_{\mu\nu}$ is the one defined to satisfy Eq. (4.55), using only the fact that

$$\nabla_\mu T^{\mu\nu} = 0, \tag{4.56}$$

which is our consistency input.

Once more, let's use the analogy with gauge theories. The gauge transformation of a $U(1)$ theory is given by

$$A_\mu \to A_\mu + \frac{1}{e} \partial_\mu \alpha. \tag{4.57}$$

The non-linear version of transformation (4.57) is given by the Yang-Mills gauge transformation

$$A^a_\mu \to A^a_\mu + \frac{1}{g} D_\mu \alpha^a. \tag{4.58}$$

Transformations (4.57) and (4.58) are related, given some details, by the exchange

$$\partial_\mu \leftrightarrow D_\mu. \tag{4.59}$$

Using this piece of information and Eq. (3.9) we write $\delta h$ as

$$\delta h_{\mu\nu} = \nabla_\mu \pi_\nu + \nabla_\nu \pi_\mu, \tag{4.60}$$

and we must check if this is consistent with Eq. (4.55). Expanding the covariant derivatives and also paying attention to the contravariant nature of the indexes, we find that

$$\delta S' = -2\lambda \int d^4x \; \pi_\nu \Big( \nabla_\mu T^{\mu\nu} - \frac{\lambda}{2}\, \partial_\alpha \, h T^{\nu\alpha} \Big), \tag{4.61}$$

which is not necessarily zero because there is a left-over term with the trace of $h_{\mu\nu}$. If such term wasn't present, everything would proceed nicely, so we need to further investigate it.

Note that based on our previous argumentation we would expect $\delta S'$ to be given only in terms of $\nabla_\mu T^{\mu\nu}$, which is zero for an arbitrary transformation parameter $\pi_\mu$. With this in mind we rearrange Eq. (4.61) as

$$\int d^4x \; (\delta h_{\mu\nu} - \lambda \pi_\mu \, \partial_\nu \, h) T^{\mu\nu} = -2 \int d^4x \; \pi_\nu \nabla_\mu T^{\mu\nu}. \tag{4.62}$$

Integrating the term with $h$ by parts in the left-hand side yields

$$\int d^4x \; (\delta h_{\mu\nu} + \lambda \, \partial_\nu \, \pi_\mu h T^{\mu\nu} + \lambda \pi_\mu h \, \partial_\nu \, T^{\mu\nu}) = -2 \int d^4x \; \pi_\nu \nabla_\mu T^{\mu\nu}. \tag{4.63}$$

The covariant conservation allows us to replace $\partial T$ by $\Gamma T$, but then $h\Gamma T$ would be a second order operator, hence we may neglect it. Also neglecting second order terms we rewrite the above expression as

$$\int d^4x \, \Big( 1 + \frac{\lambda}{2} h \Big) \delta h_{\mu\nu} T^{\mu\nu} = -2 \int d^4x \; \pi_\nu \nabla_\mu T^{\mu\nu}, \tag{4.64}$$

where we have taken into consideration the explicitly form of Eq. (4.60). We may interpret this formula in two distinct ways. The most obvious way is to re-define Eq. (4.53) with

$$\delta \tilde{h}_{\mu\nu} = \Big( 1 + \frac{\lambda}{2} h \Big) \delta h_{\mu\nu}, \tag{4.65}$$

but then we would be forcing the result to appear. The other less intuitive way is the following. Eq. (4.60) gives us explicitly that

$$\delta h_{\mu\nu} = \partial_\mu \, \pi_\nu + \partial_\nu \, \pi_\mu - \lambda \pi_\alpha (\partial_\mu \, h_{\alpha\nu} + \partial_\nu \, h_{\alpha\mu} - \partial_\alpha \, h_{\mu\nu}). \tag{4.66}$$

Note first the last term of Eq. (4.66)

$$h_{\mu\nu} \to h_{\mu\nu} + \lambda \pi_\alpha \, \partial_\alpha \, h_{\mu\nu} + \cdots \,, \tag{4.67}$$

and remember that, as already stressed out at the end of Section 3.2, $T^{\mu\nu}$ is intrinsically connected with the Poincaré group. Hence the obvious interpretation of such term is

$$h_{\mu\nu}(x) \to h_{\mu\nu}(x + \lambda \pi(x)) + \cdots \,. \tag{4.68}$$

In other words, given the symmetry in question (translational symmetry) we identify the parameter that characterises the transformation. This quantity is

$$\zeta_\mu(x) \equiv x_\mu + \lambda \pi_\mu(x). \tag{4.69}$$

The function $\zeta_\mu(x)$ is nothing more than a coordinate transformation and since it is a quite general one, there is no need for the integration measure to remain invariant; it will transform as

$$d^4x = \Big| \frac{\partial x}{\partial \zeta} \Big| d^4\zeta. \tag{4.70}$$

We conclude from Eqs. (4.55), (4.61) and (4.70) that we forgot to consider the transformation of the measure. Since we have argued that Eq. (4.64) should be the correct form of $\delta S'$ (sustained by the analogies with gauge theories), it is clear now that the induced transformation on the measure in terms of $h_{\mu\nu}$ is

$$d^4\zeta = \left(1 - \frac{\lambda}{2}h\right)d^4x. \tag{4.71}$$

At the end, when we perform a transformation (4.53), we obtain

$$S' \to \delta S' = \lambda \int d^4\zeta\ \delta h_{\mu\nu}T^{\mu\nu} = -2\lambda \int d^4x\ \pi_\nu \nabla_\mu T^{\mu\nu}, \tag{4.72}$$

which is zero as needed.

The next step now, given the new transformation parameter $\zeta(x)$, is to rewrite $\delta h_{\mu\nu}$ in terms of $\zeta$ alone, because it is the physical parameter of the general coordinate transformation symmetry group. Note that we will not redefine Eq. (4.53), but only write it in a different form; we are in position to do so due to the integral in Eq. (4.55), therefore we may integrate by parts Eq. (4.53) to obtain:

$$\begin{aligned}\lambda^{-1}\delta h_{\mu\nu}T^{\mu\nu} &= -\pi_\alpha(\partial_\mu h_{\alpha\nu} + \partial_\nu h_{\alpha\mu})T^{\mu\nu} + \cdots \\ &= [(\partial_\mu \pi_\alpha)h_{\alpha\nu} + (\partial_\nu \pi_\alpha)h_{\alpha\mu}]T^{\mu\nu} + \pi_\alpha h_{\alpha\nu}\,\partial_\mu T^{\mu\nu} + \pi_\alpha h_{\alpha\mu}\,\partial_\nu T^{\mu\nu} + \cdots,\end{aligned}$$

where we ignored second order terms comming from $d^4\zeta$. Once more, the (covariant) conservation of $T^{\mu\nu}$ allows us to substitute $\partial T$ by $\Gamma T$, but since there is already an $h_{\alpha\beta}$ in the expression, such term would be of second order, hence can be neglected. At the end, the final form of the graviton transformation is

$$\delta h_{\mu\nu} = \partial_\mu \pi_\nu + \partial_\nu \pi_\mu + \lambda(\partial_\mu \pi_\alpha)h_{\alpha\nu} + \lambda(\partial_\nu \pi_\alpha)h_{\alpha\mu} + \lambda\pi_\alpha\,\partial_\alpha h_{\mu\nu}, \tag{4.73}$$

and now we will see how we can write it in terms of $\zeta_\mu$ alone.

We begin our demonstration using Eq. (4.68) and neglecting $\mathcal{O}(\pi^2)$ terms:

$$h'_{\mu\nu}(\zeta) = h_{\mu\nu}(\zeta) + \partial_\mu \pi_\nu + \partial_\nu \pi_\mu + \lambda(\partial_\mu \pi_\alpha)h_{\alpha\nu}(\zeta) + \lambda(\partial_\nu \pi_\alpha)h_{\alpha\mu}(\zeta). \tag{4.74}$$

Next we sum a zero*:

$$\eta_{\mu\nu} - \eta_{\mu\nu} = 0 \tag{4.75}$$

$$\lambda^{-1}\eta_{\mu\nu} + h'_{\mu\nu}(\zeta) = \lambda^{-1}\eta_{\mu\nu} + h_{\mu\nu}(\zeta) + \partial_\mu \pi_\nu + \partial_\nu \pi_\mu + \lambda(\partial_\mu \pi_\alpha)h_{\alpha\nu}(\zeta) + \lambda(\partial_\nu \pi_\alpha)h_{\alpha\mu}(\zeta). \tag{4.76}$$

Now we note that

$$\frac{\partial x_\mu}{\partial x^\nu} = \eta_{\mu\nu}, \tag{4.77}$$

hence

$$\lambda^{-1}\eta_{\mu\nu} + h'_{\mu\nu}(\zeta) = h_{\mu\nu}(\zeta) + \partial_\mu \lambda^{-1}x_\nu + \partial_\mu \pi_\nu + \partial_\nu \pi_\mu + \lambda(\partial_\mu \pi_\alpha)h_{\alpha\nu}(\zeta) + \lambda(\partial_\nu \pi_\alpha)h_{\alpha\mu}(\zeta). \tag{4.78}$$

The next step is to make $\zeta$'s appear. Neglecting $\pi^2$ terms we arrive at

$$\eta_{\mu\nu} + \lambda h'_{\mu\nu}(\zeta) = \frac{\partial(x_\mu + \lambda\pi_\mu)}{\partial x^\alpha}\frac{\partial(x_\nu + \lambda\pi_\nu)}{\partial x^\beta}(\eta_{\alpha\beta} + \lambda h_{\alpha\beta}(\zeta)), \tag{4.79}$$

which is equal to

$$\eta_{\mu\nu} + \lambda h'_{\mu\nu}(\zeta) = \frac{\partial\zeta_\mu(x)}{\partial x^\alpha}\frac{\partial\zeta_\nu(x)}{\partial x^\beta}(\eta_{\alpha\beta} + \lambda h_{\alpha\beta}(\zeta(x))). \tag{4.80}$$

*$\eta$ has dimension zero, that's why the $\lambda^{-1}$ factor.

We rewrote the graviton transformation only in terms of $\zeta$ as desired.

The transformation law in Eq. (4.80) characterises what it is called *group of diffeomorfism*, i.e., a group of differentiable functions with differentiable inverse whose product is the composition between functions. This is indeed our case because it is clear in our derivation that $\zeta(x)$ is an arbitrary function that's always connected to the identity,

$$\zeta_\mu(x) = x_\mu + \lambda\pi_\mu(x), \tag{4.81}$$

therefore can always be inverted. An odd thing in Eq. (4.80), however, is the presence of the metric $\eta_{\mu\nu}$ in the transformation law of the graviton. The graviton by itself does not have a nice transformation property, but the quantity

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + \lambda h_{\mu\nu}(x) \tag{4.82}$$

does have one:

$$g'_{\mu\nu}(\zeta) = \frac{\partial\zeta_\mu(x)}{\partial x^\alpha}\frac{\partial\zeta_\nu(x)}{\partial x^\beta}g_{\alpha\beta}(\zeta(x)). \tag{4.83}$$

Throughout this Chapter it was made clear that the symmetry group in question is the Poincaré group. This group, which at the level of $S_2$ was a global group, becomes now in non-linear regime a local group, as stated in Eq. (4.81). Such change in the nature of the group brought huge changes to our theory: infinite non-linearity; covariant derivatives that depend on the Lorentz structure; qualitatively distinction between up and down indices, and so on. But definitely the most striking change is in Eq. (4.82), which means that the graviton field is not a tensor with respect to the new symmetry group induced by it, but $g_{\mu\nu}$ instead. This is the reason why we have introduced a $\eta R$ term in Eq. (4.18), for the strength field tensor will come up naturally as a tensor from relation (4.23). As a consequence all indices contractions must be made by $g_{\mu\nu}$ and not simply by $\eta_{\mu\nu}$. Not only that, but also the integration measure $d^4x$ is not invariant anymore, since Eq. (4.81) is not linear in $x$. In a few words, the situation becomes a mess.

Fortunately, infinite non-linearity comes with a spectacular advantage: geometrical interpretation [6, 13, 14]. The definition of $g_{\mu\nu}$ in Eq. (4.82) means straightforwardly that the metric of the space-time manifold is not flat anymore. In Differential Geometry (DG) program, all inner properties of a manifold can be computed from the metric $g_{\mu\nu}$, for instance the affine connection $\Gamma^\alpha_{\mu\nu}$, the curvature tensor $R_{\mu\nu\alpha\beta}$, the covariant derivatives $\nabla_\mu$ and also the new invariant integration measure $d^4x\sqrt{-\det g_{\mu\nu}}$. From this geometrical approach we may recover our discussion in weak field limit, i.e. $\lambda h \ll 1$. Although one must note that in our prescription $h_{\mu\nu}$ is not necessarily small.

Despite the method of general covariance together with DG being much more straightforward, elegant and simpler in some sense, the field theory method here described has two advantages. First, it guarantees that the theory given by the action $S$ in Eq. (4.16) is unique, something the usual derivations of the Einstein equations cannot. Second, we already have the action for the non-linear theory, which is given by the Einstein-Hilbert action [1, 6]

$$S_{\text{EH}} = M_P^2\int d^4x\sqrt{-g}R, \tag{4.84}$$

where $g \equiv \det g_{\mu\nu}$ and $R = g^{\mu\nu}R^\alpha_{\ \mu\alpha\nu}$. Action (4.84) is precisely the non-linear version of the action in Eq. (4.18). With DG alone we can determine (infer) the Einstein's equations and at any point there is a connection to field theory, therefore DG cannot determine the action $S_{\text{EH}}$ as precisely as with our argumentation.

# 5 Modified Theories of Gravity

Now we start the second part of this work. Here we will analyse some of the so called *Modified Gravity* (MoG) theories, in other words theories that attempt to generalise Einstein's theory. Although GR is a very successful theory, we search for new theories in order to explain yet unsolved problems in it, for example the Cosmological Constant (CC) problem or issues within cosmology itself. In this sense we seek not for small distances, but instead large distances modifications, also named Infrared (IR) theories. Such theories shall be constrained by the fact that GR is the "effective theory" at distances smaller than $H_0^{-1}$, hence every prediction of a MG should only differ significantly from GR at distances larger then $H_0^{-1}$.

In this paper though, we are not interested on only the phenomenological motivations but also on the theoretical ones, which follows from the discussion from the previous chapters. We ask ourselves, how do we change GR in a non-trivial way if GR is the unique field theory of a massless spin 2 particle? The complete answer to this question is still unknown, we will, however, illustrate it with two examples: $f(R)$-theories [31, 32] and Massive Gravity (MaG) theories [5]. They represent two distinct families of MG theories. The first is a generalisation of GR formulated with field theory and it is smoothly connected to GR, in the sense that we can take some smooth limit to re-obtain GR field equations. The latter is also formulated with field theoretical methods, but is not smoothly connected to GR due to the introduction of a mass to the graviton.

## 5.1 $f(R)$ theories

We begin our discussion with the so called $f(R)$-*theories*, which are defined by the action

$$S_f = M_P^2 \int d^4x \sqrt{-g} f(R), \tag{5.1}$$

where $f$ is the most generic function of the Ricci scalar $R$

$$f(R) = \cdots + \frac{\alpha_{-1}}{R} + \alpha_0 \Lambda + R + \alpha_2 R^2 + \cdots , \tag{5.2}$$

where $\alpha_n$ are arbitrary numerical constants and $\Lambda$ is a possible cosmological constant. Note that by taking all $\alpha_n$ to zero we re-obtain the action given in Eq. (4.84), in other words, usual GR. In this manner that we say that GR can be smoothly obtained from Eq. (5.1).

Now we ask ourselves, is this theory a truthful modification of GR? The answer is surprisingly no, although it is not so simple to prove it. If we are interested in any modification at all, we should assume that $\frac{\mathrm{d}^2 f}{\mathrm{d}R^2} \neq 0$, because else we would have

$$\frac{\mathrm{d}^2 f}{\mathrm{d}R^2} = 0 \Rightarrow f(R) = R + \alpha_0 \Lambda, \tag{5.3}$$

which is just usual GR. Hence we can assume that its second derivative is different from zero. This allows us to perform a Legendre transformation on $f$ with conjugate variable $\phi$ [33]. Note that the $\phi$ variable is a scalar due to the fact that $R$ is also a scalar. Such transformation gives us

$$f(R) = \phi R - V(\phi), \tag{5.4}$$

with $V$ the Legendre transform of $f$. Substituting the above expression in the action (5.1) one obtains

$$S = M_P^2 \int d^4x \sqrt{-g} (R\phi - V(\phi)). \tag{5.5}$$

One remark before advancing in our proof is in order. A particular case of Eq. (5.5) is Brans-Dicke theory [6], for which the potential $V$ is given by

$$V_{\mathrm{BD}}(\phi) = \frac{\omega}{\phi}\nabla_\mu\phi\nabla^\mu\phi, \tag{5.6}$$

with $\omega$ a dimensionless parameter. This theory was proposed by Brans and Dicke in 1961 and was one of the first attempts to modify gravity. Here we see that it is a particular case of a more general situation.

Returning to our proof, it is not yet clear why should Eq. (5.5) be GR plus some extra d.o.f.. The difficulty emerges from the $R\phi$ term, which is not present in usual GR. A way to remove such term from the action is to perform a *conformal transformation* [13], i.e.

$$g_{\mu\nu} \to \tilde{g}_{\mu\nu} \equiv \phi g_{\mu\nu}. \tag{5.7}$$

Eq. (5.7) affects every other quantity that depends on the metric, in particular for 3+1 dimensions

$$\sqrt{-\tilde{g}} = \phi^2\sqrt{-g}, \tag{5.8}$$

$$R = \phi\left[\tilde{R} + 3\frac{\tilde{\nabla}_\mu\tilde{\nabla}^\mu\phi}{\phi} - \frac{9}{2}\left(\frac{\tilde{\nabla}_\mu\phi}{\phi}\right)^2\right], \tag{5.9}$$

where $\tilde{\nabla}_\mu$ is the covariant derivative in the transformed manifold. Substituting the transformations above in Eq. (5.5) we obtain

$$S = M_P^2\int d^4x\sqrt{-\tilde{g}}\left[\tilde{R} + 3\frac{\tilde{\nabla}_\mu\tilde{\nabla}^\mu\phi}{\phi} - \frac{9}{2}\left(\frac{\tilde{\nabla}_\mu\phi}{\phi}\right)^2 - \frac{V(\phi)}{\phi^2}\right], \tag{5.10}$$

from which is already clear that the first term is just the Einstein-Hilbert action (4.84). To make explicit that we are dealing with an ordinary scalar field we redefine $\phi$:

$$\Phi \equiv \sqrt{3}\ln\phi. \tag{5.11}$$

Rewriting $S$ in terms of $\Phi$ we arrive at

$$S_f = M_P^2\int d^4x\sqrt{-\tilde{g}}\left[\tilde{R} + \frac{1}{2}(\tilde{\nabla}_\mu\Phi)^2 - \mathcal{V}(\Phi)\right], \tag{5.12}$$

in which

$$\mathcal{V}(\Phi) = e^{-\frac{2\Phi}{\sqrt{3}}}V(e^{\Phi/\sqrt{3}}). \tag{5.13}$$

We have proved in Eq. (5.12) that the theory defined by Eq. (5.1) is GR plus a scalar field coupled to gravity in the standard way. This result is actually very well established in the literature [31, 32], but with its theoretical worth overlooked. On the one hand one can say that Eq. (5.12) really is a novel theory of gravitation since it has some new physics given by the new scalar d.o.f. $\Phi$. On the other hand, though, we argue that Eq. (5.12) is not a truthful modification of GR, because it only accounts for the addition of some new field $\Phi$. Intrinsically the gravitational interactions are still described by the dynamics of a massless spin 2 particle, namely Eq. (4.84). With this reasoning we start to understand how strong the result of the previous chapters really is. Even the most arbitrary function $f(R)$ just adds a trivial scalar field to the theory and leaves the dynamics of the graviton essentially unchanged.

However surprising, we should have foreseen this. If we construct a field theory which is GR plus something (see Eq. (5.2)), then by our previous result we will certainly obtain GR plus something coupled to it. In order to hope for a true modification of gravity, we *must* break at least one of the assumptions we have made at the beginning.

## 5.2 Massive Gravity

The second class of theories we will consider are the *Massive Gravity* ones. As the name suggests, everything we did so far will remain the same, except the mass of the graviton: in MaG the graviton will acquire a mass $M$ different from zero. We immediately see that those theories must be quite complicated, since well established experimental facts such as the one in Eq. (1.1) must hold still. Nevertheless we still try to comprehend in what respect they will differ from GR.

As said previously, $f(R)$-theories are somewhat smoothly connected to GR and hence resulted to be a trivial modification of the latter. Such is not true for MaG. This statement may be understood by the following argumentation. Take for instance the propagator in Eq. (3.33) for massless graviton in harmonic gauge:

$$\Delta^0_{\mu\nu\alpha\beta}(p) = \frac{i}{p^2} P^0_{\mu\nu\alpha\beta} = \frac{i}{2p^2}[\eta_{\mu\alpha}\eta_{\nu\beta} + \eta_{\mu\beta}\eta_{\nu\alpha} - \eta_{\mu\nu}\eta_{\alpha\beta}],$$

obtained by solving the EoM. The addition of a mass $M$ through the Fierz-Pauli term [28]

$$\mathscr{L}_M = -\frac{M^2}{2}\Big(h_{\mu\nu}h^{\mu\nu} - h^2\Big) \tag{5.14}$$

will result in a different EoM, given by

$$\Box h_{\alpha\beta} - 2\,\partial^\sigma\,\partial_{(\alpha}\,\bar{h}_{\beta)\sigma} + M^2(\bar{h}_{\mu\nu} + \eta_{\mu\nu}h) = \lambda\bar{T}_{\mu\nu}, \tag{5.15}$$

in contrast to Eq. (3.21). The inversion of the above equation is a bit more complicated, but nevertheless straightforward. The propagator is given by [5, 9]

$$\Delta^M_{\mu\nu\alpha\beta}(p) = \frac{i}{p^2 - M^2} P^M_{\mu\nu\alpha\beta} = \frac{i}{2(p^2 - M^2)}\left[\tilde{\eta}_{\mu\alpha}\tilde{\eta}_{\nu\beta} + \tilde{\eta}_{\mu\beta}\tilde{\eta}_{\nu\alpha} - \frac{2}{3}\tilde{\eta}_{\mu\nu}\tilde{\eta}_{\alpha\beta}\right], \tag{5.16}$$

where $\tilde{\eta}_{\mu\nu}$ is a new momentum-dependent tensor defined by

$$\tilde{\eta}_{\mu\nu} = \eta_{\mu\nu} - \frac{p_\mu p_\nu}{M^2}. \tag{5.17}$$

The fact that the propagator structure for the massless and the massive one at zero momentum are not equal,

$$P^M_{\mu\nu\alpha\beta}(p=0) \neq P^0_{\mu\nu\alpha\beta}, \tag{5.18}$$

is known as *van Dam-Veltman-Zakharov* (vDVZ) discontinuity and happens due to the factor of 2/3 in Eq. (5.16). In fact, it follows from Eq. (5.18) that, as the propagation at low energy ($p \to 0$) is distinct, MaG is a priori an intrinsically distinct theory of gravitation. But this is actually disastrous because it contradicts the predictions from GR, in particular Eq. (1.1) follows exactly because the graviton is a massless particle. We have reached a dead-lock, where on the one hand we have found a truthful modification of GR, but one the other hand made all phenomenology inconsistent with experiments. How do we conciliate both? The answer to this question is actually very complicated and involves much more theoretical tools than the ones we are dealing with in this paper. Although the situation is highly non-trivial, we can still have a deeper understanding of it by using the same spin 1 analogy we've used previously. Through this analogy we may, in particular, visualise why the vDVZ discontinuity is a huge problem and how it relates to the physical graviton.

Consider the propagator for a spin 1 particle given by the diagram

$$\mu \sim\!\sim\!\sim\!\sim\!\sim\!\sim \nu \; = \frac{iN_{\mu\nu}}{p^2 - M^2 + i\epsilon}, \tag{5.19}$$

where $N_{\mu\nu} = N_{\mu\nu}(p, M)$ is some tensor and $M$ is the mass for such particle. The denominator of Eq. (5.19) is determined based on unitarity grounds as already said previously in Chapter 3. Not exactly by chance, but the numerator of Eq. (5.19) is also determined by unitarity, it must be the sum of all physical polarisation states of the particle, hence [9, 12]

$$N_{\mu\nu}(p, M) = \sum_i \epsilon^{*i}_\mu(p)\epsilon^i_\nu(p) = -\eta_{\mu\nu} + \frac{p_\mu p_\nu}{M^2} = -\tilde{\eta}_{\mu\nu}. \tag{5.20}$$

Note that in the above expression we have the first piece that depends on neither the momentum nor the mass, and the second that depends on both. This sum can be factorised into two pieces, the one that depends only on the helicity modes and the other that represents the sole longitudinal mode, namely

$$N_{\mu\nu}(p, M) = \sum_{h=+-} \epsilon^{*h}_\mu(p)\epsilon^h_\nu(p) + \epsilon^{*L}_\mu(p)\epsilon^L_\nu(p). \tag{5.21}$$

This is an essential step, because such factorisation has the same foundation as the one we have made in Chapter 3 to determine the coefficients from Lagrangian (3.7), in which we separated the transverse and longitudinal mode for the spin 2 case. For spin 1 this decomposition is given by [12]

$$A_\mu = A^T_\mu + \partial_\mu \pi, \tag{5.22}$$

with $\partial_\mu A^T_\mu = 0$. Gauge invariance is the restored if the longitudinal mode also transforms as $\pi \to \pi + \theta$ under gauge transformations. Therefore we can use the usual gauge invariance of $A^T_\mu$ to actually make the sum of transverse polarisation vectors equal to

$$\sum_{h=+-} \epsilon^{*h}_\mu(p)\epsilon^h_\nu(p) \to -\eta_{\mu\nu}, \tag{5.23}$$

in other words, to drop all momentum dependency on the longitudinal component. From this point of view, the term proportional to $\frac{1}{M^2}$ in Eq. (5.20) must be associated with $\epsilon^{*L}_\mu(p)\epsilon^L_\nu(p)$ alone. Hence, if we are interested in the $M \to 0$ limit, we see that the piece that blows up to infinity is exactly the one that must be removed from the polarisation sum in the massless case, the longitudinal component. In this sense the massive propagator is connected "continuously" to the massless one. We just need to realise that the origin of the singularity comes from the DoF that must be discarded.

As we have already seen, such procedure fails for the graviton, as the momentum independent part of of propagator (5.16) does not match the one obtained from the transverse polarisation sum given in Eq. (3.33). This situation cannot be reversed by performing a gauge transformation, because only momentum-dependent pieces are affected by them. The vDVZ discontinuity indicates that the interactions between the longitudinal modes of the massive graviton are highly non-trivial and interfere in the dynamics of the transverse modes.

From this point onward the development of MaG becomes very technical and not particularly illuminating to the discussion here. What is worth remarking is that the vDVZ discontinuity can indeed be removed through the so called Vainshtein mechanism, which is a screening of the longitudinal modes by non-linear interactions that emerge in the $M \to 0$ limit, and we are left essentially with GR plus scalar fields [5].

A final remark on MaG theories is in order. What all MaG theories have in common is obviously the fact that the graviton acquires a mass, hence experiments that measure such mass can bound those theories. There are many kinds of independent experiments [34, 35, 36]. For example, ones that measure the Yukawa potential, others that comes from gravitational waves detection and even some that measure signals from systems of binary pulsars. From weak gravitational lensing measurements one gets [19] the following upper bound:

$$M \lesssim 6 \cdot 10^{-32} \text{ eV}. \tag{5.24}$$

Note that the value of this mass is extremely small, it is way smaller than the bounds for the photon [19]. Despite the smallness of this value, Eq. (5.24) is still consistent with the idea of a IR modification of GR, because, as stressed out in the beginning of this chapter, the relevant distance scale at which new interactions should appear is $H_0^{-1}$. So we should expect $M \sim H_0 \sim 10^{-42}$ eV. Notwithstanding this value puts MaG theories into a tight spot. Can a value of $10^{-32}$ eV produce significant effects? And if it really is massive, how do we explain the smallness of this mass? Those are question to be answered in the future, when the theoretical and experimental nature of the graviton are better understood.

## Conclusion

On the first part of this work we presented a complete argumentation that General Relativity is the only theory consistent with a massless spin 2 field theory. In particular, we avoided bias from our previous knowledge of GR and made a clear proof relying only on particle and field theoretical methods. This line of thought is also consistent with spin 1 gauge theories, in the sense that the gauge structure and non-linear interactions share the same pattern in both theories.

We have explored on the second part two classes of modifies theories of gravitation. Although both kinds of theories are very different from each other and have distinct relations to GR, they end up being equivalent to GR coupled to some other d.o.f.. This shows us that the uniqueness of GR as a field theory can tell us lots about other theories that attempt to modify it. Consequently, it is essential to understand the subtleties of the field theoretical formulation of gravity if we ever hope to consistently modify Einstein's theory of gravitation.

## Acknowledgements

I would like to thank Professor Enrico Bertuzzo for his support and suggestions that were of great help to the development this paper. I also wish to thank Pedro Bittar O. S. and Renan N. Hirayama for the discussions of the consistency of the energy-momentum tensor in the linear theory.

## References


[1] Richard Feynman. *Feynman lectures on gravitation.* CRC Press, 2018.

[2] Steven Weinberg. Photons and gravitons in perturbation theory: Derivation of maxwell's and einstein's equations. *Physical Review*, 138(4B):B988, 1965.

[3] Stanley Deser. Self-interaction and gauge invariance. *General Relativity and gravitation*, 1(1):9–18, 1970.

[4] David G Boulware and Stanley Deser. Classical general relativity derived from quantum gravity. *Annals of Physics*, 89(1):193–240, 1975.

[5] Claudia de Rham. Massive gravity. *Living reviews in relativity*, 17(1):7, 2014.

[6] Steven Weinberg. *Gravitation and cosmology: principles and applications of the general theory of relativity*, volume 1. Wiley New York, 1972.

[7] Timothy Clifton, Pedro G Ferreira, Antonio Padilla, and Constantinos Skordis. Modified gravity and cosmology. *Physics reports*, 513(1-3):1–189, 2012.

[8] Xavier Calmet and Iberê Kuntz. What is modified gravity and how to differentiate it from particle dark matter? *The European Physical Journal C*, 77(2):132, 2017.

[9] Nima Arkani-Hamed. Institute for Advanced Studies, Lectures: Robustness of GR. Attempts to Modify Gravity., 2011. URL: `https://video.ias.edu/pitp-2011-arkani-hamed1`. Last visited on 2019/04/11.

[10] Nivaldo A Lemos. *Mecânica analítica.* Editora Livraria da Física, 2007.

[11] Michael Edward Peskin. *An introduction to quantum field theory.* Westview press, 1995.

[12] Matthew D Schwartz. *Quantum field theory and the standard model.* Cambridge University Press, 2014.

[13] Sean M Carroll. *Spacetime and geometry. An introduction to general relativity.* 2004.

[14] Robert M Wald. General relativity, chicago, usa: Univ. *Pr. 491p*, 1984.

[15] Valery V Nesvizhevsky, HG Börner, AM Gagarski, AK Petoukhov, GA Petrov, H Abele, S Baeßler, G Divkovic, FJ Rueß, Th Stöferle, et al. Measurement of quantum states of neutrons in the earths gravitational field. *Physical Review D*, 67(10):102002, 2003.

[16] Jun John Sakurai, Jim Napolitano, et al. *Modern quantum mechanics.* Pearson Harlow, 2014.

[17] Roberto Colella, Albert W Overhauser, and Samuel A Werner. Observation of gravitationally induced quantum interference. *Physical Review Letters*, 34(23):1472, 1975.

[18] FW Dyson. Ase eddington a. nd cr davidson. *Phil. Trans. Roy. Soc. A*, 220:221, 1920.

[19] M. Tanabashi *et al.*(Particle Data Group). *˙Physical Review D*, 98, 2018.

[20] Steven Weinberg. *The quantum theory of fields*, volume 2. Cambridge university press, 1995.

[21] Mark Thomson. *Modern particle physics*. Cambridge University Press, 2013.

[22] Wu-Ki Tung. *Group theory in physics: an introduction to symmetry principles, group representations, and special functions in classical and quantum physics*. World Scientific, 2005.

[23] G Feinberg and J Sucher. Long-range forces from neutrino-pair exchange. *Physical Review*, 166(5):1638, 1968.

[24] Gerald Feinberg, Joseph Sucher, and C-K Au. The dispersion theory of dispersion forces. *Physics Reports*, 180(2):83–157, 1989.

[25] Steven Weinberg. *The quantum theory of fields*, volume 1. Cambridge university press, 1995.

[26] Alberto Nicolis. Florence HEP, Lectures: Effective Field Theories for Gravity., 2017. URL: `https://www.youtube.com/playlist?list=PLDxsZU4NC6Z4wQ_qaekc8u6LacOZbakDR`. Last visited on 2019/04/11.

[27] John David Jackson. *Classical electrodynamics*. John Wiley & Sons, 2012.

[28] Markus Fierz and Wolfgang Ernst Pauli. On relativistic wave equations for particles of arbitrary spin in an electromagnetic field. *Proc. R. Soc. Lond. A*, 173(953):211–232, 1939.

[29] Suraj N Gupta. Gravitation and electromagnetism. *Physical Review*, 96(6):1683, 1954.

[30] Steven Weinberg and Edward Witten. Limits on massless particles. *Physics Letters B*, 96(1-2):59–62, 1980.

[31] Antonio De Felice and Shinji Tsujikawa. f (r) theories. *Living Reviews in Relativity*, 13(1):3, 2010.

[32] Thomas P Sotiriou and Valerio Faraoni. f (r) theories of gravity. *Reviews of Modern Physics*, 82(1):451, 2010.

[33] Vladimir Igorevich Arnol'd. *Mathematical methods of classical mechanics*, volume 60. Springer Science & Business Media, 2013.

[34] Claudia de Rham, J Tate Deskins, Andrew J Tolley, and Shuang-Yong Zhou. Graviton mass bounds. *Reviews of Modern Physics*, 89(2):025004, 2017.

[35] Xueli Miao, Lijing Shao, and Bo-Qiang Ma. Bounding the mass of graviton in a dynamic regime with binary pulsars. 2019.

[36] Lee Samuel Finn and Patrick J Sutton. Bounding the mass of the graviton using binary pulsar observations. *Physical Review D*, 65(4):044022, 2002.