Massive spin-2 theories

Sarah Folkerts[*]*, Cristiano Germani*[†] *and Nico Wintergerst*[‡]

Arnold Sommerfeld Center, Ludwig-Maximilians-Universität, Theresienstr. 37, 80333 München, Germany

Abstract

We give an introduction to massive spin-2 theories (including massive gravity) and the problem of their non-linear completion. We review the Boulware-Deser ghost problem and two ways to circumvent classic no-go theorems. In turn, massive spin-2 theories are not uniquely defined. In the case of truncated theories, we show that the Boulware-Deser ghost may only be avoided if the derivative structure of the theory is not tuned to be Einsteinian.

1 Introduction

If low energy physics is described by the language of effective field theory (EFT), an important question to ask is what (interacting) degrees of freedom can in principle be used on four-dimensional Minkowski space.

In a Poincaré invariant theory, such as an EFT constructed on a four-dimensional flat space, different degrees of freedom as well as their corresponding one-particle states may be labelled by their mass and spin, which are Casimir operators of the Poincaré group.

This classification is of course most useful when the one-particle states under consideration are eigenstates of the full Hamiltonian. However, in almost all cases diagonalization of the interacting Hamiltonian is extremely involved. Therefore, one introduces the concept of *asymptotic states*, eigenstates of the quadratic part of the Hamiltonian. This is due to the fact that a typical measurement involves a scattering experiment where the experimental apparatus may only monitor the *in* and *out* states far away from the scattering interactions. This brings about the concept of the S-matrix

$$P(in, out) = \langle in|S|out \rangle \,, \tag{1}$$

where $P \leq 1$ is the probability that the *in* asymptotic (far away from the scattering process) state scatters into the *out* asymptotic state where, importantly, *in* and *out* states are eigenstates of the *free* Hamiltonian. In almost all cases, these states are also eigenstates of the Casimir operators of the Poincaré group, thus allowing the above classification on the asymptotic states. Here, we will follow this trend.

Fundamental theories must have that for all processes $P \leq 1$. This condition is what is commonly called the unitarity bound. In an EFT instead, the considered action is only an approximation to the full quantum effective action; it corresponds to the first terms of a perturbative expansion in terms of the dimensionless parameters E/Λ_s and ϕ/Λ_s, where E is the energy, ϕ represents any field content and Λ_s is the so-called cut-off or strong coupling scale. Of course, if $\Lambda_s \to \infty$, the theory is valid up to any energy scale and it is called *renormalizable*.

Whenever the perturbative expansion breaks down, for sufficiently high transfer-energies $E > \Lambda_s$, one obtains $P > 1$. In this case, if one insists on this theory still being described in terms of a perturbative expansion, it must be rewritten in terms of new degrees of freedom materializing at an energy scale $E > \Lambda_s$. This is what is commonly called a Wilsonian UV completion.

---

[*]E-mail: Sarah.Folkerts@physik.uni-muenchen.de

[†]E-mail: cristiano.germani@lmu.de

[‡]E-mail: nico.wintergerst@physik.uni-muenchen.de

Theories with massive and massless spin 0 and spin 1/2 fields together with massless spin 1 fields can be constructed without violation of unitary. This is, in fact, the content of the Standard Model of particle physics.

Any other degrees of freedom will inexorably violate unitarity at some energy scale (non-renormalizable theories). If it was not for gravity, a massless spin 2 particle, we could just ignore these other degrees of freedom. However, as gravitational interactions *do* exist, one may ask whether other non-renormalizable degrees of freedom also exist up to energy scales that are not experimentally probed yet.

A self-interacting massive vector field with mass m has a typical strong coupling scale $\Lambda \sim m$. The violation of unitarity is due to the fact that the longitudinal mode of the vector, which is non-physical in the massless case, has a polarization $\vec{\epsilon} \propto \frac{\vec{k}}{m}$. Thus, the larger is the mass, the later we need to postpone the completion necessary to restore unitarity.

This can be understood by (schematically) considering the interaction in Fourier space

$$\int d^4x\, A_\mu J^\mu \sim \int dt\, d^3k\, J_{\mu;(\mathbf{k})} \sum_{\lambda=1}^{3} \epsilon_{\mathbf{k},\lambda}^{\mu} A_{\mathbf{k}} \,, \tag{2}$$

where the sum extends over the three polarizations defined by the conditions

$$\begin{aligned} \mathbf{k}_\mu \epsilon_{\mathbf{k},\lambda}^{\mu} &= 0 \,, \\ \epsilon_{\mu;\mathbf{k},\lambda} \epsilon_{\mathbf{k},\lambda}^{\mu} &= 1 \,. \end{aligned} \tag{3}$$

The spatial part of the longitudinal polarization is defined to be parallel to the three-momentum, i.e. $\epsilon_{\mathbf{k},3}^{i} \propto \mathbf{k}^i$. Explicitly, it takes on the form

$$\epsilon_{\mathbf{k},3}^{\mu} = \left(\frac{|\mathbf{k}|}{m}, \frac{\mathbf{k}}{|\mathbf{k}|} \frac{E_{\mathbf{k}}}{m} \right) \,, \tag{4}$$

thus giving rise to a vertex

$$J_{\mu;(\mathbf{k})} \epsilon_{\mathbf{k},3}^{\mu} A_{\mathbf{k}} \tag{5}$$

which becomes strong at energies of order m.

Note that for large $|\mathbf{k}|$,

$$\epsilon_{\mathbf{k},3}^{\mu} \approx \frac{1}{m} k^\mu \,. \tag{6}$$

This signals a straightforward way to avoid the violation of unitarity. If the source J_μ is conserved, $\partial_\mu J^\mu = 0$, in the large momentum limit $J_{\mu;(\mathbf{k})} \epsilon_{\mathbf{k},3}^{\mu} \to 0$.

We can understand this effect in a different way. We can construct one particle states by defining creation and annihilation operators $a_{\mathbf{k}}^{\lambda}$, generating the eigenstates of the non-interacting Hamiltonian. We write

$$A_\mu^\lambda = \int \frac{d^3k}{\sqrt{2E_{\mathbf{k}}}} \epsilon_\mu^\lambda(\mathbf{k}) \left(a_{\mathbf{k}}^{\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} + a_{\mathbf{k}}^{\dagger\lambda} e^{-i\mathbf{k}\cdot\mathbf{x}} \right) \,. \tag{7}$$

At the same time we can define a new scalar field ϕ as

$$\phi = \int \frac{d^3k}{\sqrt{2E_{\mathbf{k}}}} \left[\left(i a_{\mathbf{k}}^{(3)} \right) e^{i\mathbf{k}\cdot\mathbf{x}} + \left(i a_{\mathbf{k}}^{(3)} \right)^{\dagger} e^{-i\mathbf{k}\cdot\mathbf{x}} \right] \,. \tag{8}$$

On an asymptotic state $|\mathbf{k}\rangle$ of four-momentum k^μ one has

$$\partial_\mu \phi |\mathbf{k}\rangle = \frac{1}{\sqrt{2E_{\mathbf{k}}}} k_\mu \left(a_{\mathbf{k}}^{(3)} e^{i\mathbf{k}\cdot\mathbf{x}} + a_{\mathbf{k}}^{\dagger(3)} e^{-i\mathbf{k}\cdot\mathbf{x}} \right) |\mathbf{k}\rangle \,. \tag{9}$$

Henceforth

$$\left(A_\mu^{(3)} - \frac{1}{m}\partial_\mu\phi \right) |\mathbf{k}\rangle \underset{\mathbf{k}\gg m}{\sim} \frac{m}{|\mathbf{k}|}\hat{k}_\mu a_\mathbf{k}^{(3)}|\mathbf{k}\rangle \underset{\mathbf{k}\to\infty}{\to} 0 \;. \tag{10}$$

Here, $\hat{k}_\mu$ is the unit vector pointing in the direction of $\mathbf{k}$. We thus see that the longitudinal polarization in the high energy limit is well described by a scalar field up to corrections $\mathcal{O}\left(\frac{m}{\mathbf{k}}\right)$. This is the essence of the Stückelberg decomposition of the massive vector. There, the field is decomposed into a massless vector and a scalar where the scalar is nothing else than the re-incarnation of the gauge direction of the massless case.

To be precise, in the massless case the action is invariant under the gauge transformation

$$A_\mu \to A_\mu + \frac{\partial_\mu\phi}{m} \;, \tag{11}$$

for any ϕ and some mass scale m. In the massive case however ϕ represents the extra polarization at high energies as in (10). In this case, we can decompose

$$A_\mu = \tilde{A}_\mu + \frac{\partial_\mu\phi}{m} \;. \tag{12}$$

As a consequence, A_μ, and hence any action constructed from it, is invariant under transformations of $\tilde{A}_\mu$ of the type (11) (U(1)) if the change is absorbed by a shift in the scalar ϕ. We see that that the interaction of the scalar degrees of freedom to external conserved sources J^μ is absent:

$$\int d^4x A_\mu J^\mu = \int d^4x \tilde{A}_\mu J^\mu + \frac{1}{m}\,\text{boundary}\;. \tag{13}$$

Thus a massive linearly interacting vector can exist without unitarity problems. The would be strong coupling scale appears in fact only in the boundary term.

The question is now whether interacting higher spins $s \geq 3/2$ can be consistently constructed. First of all, one notes that, because of the non-trivial tensorial structure of $s \geq 3/2$ fields, all interactions must be unitarity violating [1]. Furthermore, in the case of massive fields, the extra longitudinal polarization does not decouple in the massless limit even if the field interacts with a conserved source. This is simply due to the fact that these extra polarizations always carry contributions which are not proportional to the four-momentum even in the high-energy limit.

Let us take as an example a massive spin 2 field $h_{\alpha\beta}$. Similar to the massive vector field discussed above, the properties of this theory can also be investigated through a helicity or *linear* Stückelberg decomposition.

The decomposition of the field into helicity eigenmodes is in complete analogy to the massive vector. For high energies the helicity-1 component (or vectorial polarization) can be described by the derivative of a Lorentz vector (A_μ), whereas the helicity-0 component (or longitudinal polarization) can be described by a scalar field χ [2,3]. The helicity-2 component (the transverse polarization) is described by a tensor $\tilde{h}_{\mu\nu}$.

One then decomposes the massive spin-2 field as

$$h_{\mu\nu} = \tilde{h}_{\mu\nu} + \frac{\partial_{(\mu}A_{\nu)}}{m} + \frac{1}{3}\left(\frac{\partial_\mu\partial_\nu\chi}{m^2} + \frac{1}{2}\eta_{\mu\nu}\chi \right) \;. \tag{14}$$

Here we used the symmetrization convention $a_{(\mu}b_{\nu)} = \frac{1}{2}(a_\mu b_\nu + a_\nu b_\mu)$.

Similar to the massive vector, interactions of the longitudinal polarization violate unitarity. The coupling to the energy momentum tensor is of the form (recall that the energy momentum tensor must be of dimension 4)

$$\frac{1}{\Lambda_s}\int d^4x\, h_{\alpha\beta}T^{\alpha\beta} = \frac{1}{6\Lambda_s}\int d^4x\, \chi T + \dots \;, \tag{15}$$

and violates unitarity at the scale Λ_s.

If we were instead considering a non-conserved source $J^{\mu\nu}$ of mass dimension $d \geq 2$ [1] we would have in addition the interaction

$$\int d^4x \, h_{\alpha\beta} J^{\alpha\beta} = \frac{\Lambda^{3-d}}{3m^2} \int d^4x \, \partial_{\mu\nu} \chi \tilde{J}^{\mu\nu} + \dots \,, \tag{16}$$

where Λ is the mass scale of the source J and $\tilde{J} = J/\Lambda^{3-d}$ is the dimensionless source. This interaction creates a strong coupling scale $\Lambda_s \equiv \left(\frac{3m^2}{\Lambda^{3-d}}\right)^{\frac{1}{d-1}}$ proportional to the mass m.

What we thus see is that theories describing interactions of fields of spin $s \geq 3/2$, as EFTs on a Minkowski background, cannot be fundamental (unless free) and must be UV completed in the Wilsonian sense if they are to be described in perturbation theory. For a different route to UV completion see, e.g, [4–6].

A Poincaré invariant consistent theory of massless spin-2, at least at the lowest momentum expansion, must be Einstein's theory [7, 8]. Therefore, at least in the low momentum limit, a massless spin-2 theory is unique up to a strong coupling scale Λ_s. The question we would like to address is whether a similar "uniqueness" theorem holds for a massive spin 2 at least up to a generic strong coupling scale Λ_s. As we shall demonstrate, this is not the case.

2 Massive spin-2 without self-interactions

The action of a free massive spin-2 particle is given by what is commonly called the Fierz-Pauli action and is unique [9].

Its construction can be understood most easily by considering the aforementioned helicity or Stückelberg decomposition. Demanding the absence of higher derivatives, which signal the appearance of new degrees of freedom, removes all arbitrariness in the action; only the Fierz-Pauli form allows for this property. It is given by

$$\begin{aligned} S = \int d^4x \mathcal{L} &= \int d^4x \left(\partial_\mu h^{\mu\nu} \partial_\nu h - \partial_\mu h^{\rho\sigma} \partial_\rho h^\mu_\sigma + \frac{1}{2} \partial_\mu h^{\rho\sigma} \partial^\mu h_{\rho\sigma} - \frac{1}{2} \partial_\mu h \partial^\mu h \right. \\ &\quad \left. - \frac{1}{2} m^2 (h^{\mu\nu} h_{\mu\nu} - h^2) \right), \end{aligned} \tag{17}$$

where $h \equiv h^\mu_\mu$.

Inserting (14) into the quadratic action (17) leads to

$$\begin{aligned} \mathcal{L}_{\text{PF}} &= \tilde{h}^{\mu\nu} \mathcal{E}^{\rho\sigma}_{\mu\nu} \tilde{h}_{\rho\sigma} - \frac{1}{8} F_{\mu\nu} F^{\mu\nu} + \frac{1}{12} \chi \Box \chi - \frac{1}{2} m^2 \left(\tilde{h}^{\mu\nu} \tilde{h}_{\mu\nu} - \tilde{h}^2 \right) + \frac{1}{6} m^2 \chi^2 \\ &\quad + \frac{1}{2} m^2 \chi \tilde{h} + m \left(\tilde{h} \partial_\mu A^\mu - \tilde{h}^{\mu\nu} \partial_\mu A_\nu \right) + \frac{m}{2} \chi \partial_\mu A^\mu \,, \end{aligned} \tag{18}$$

where $\tilde{h}^{\mu\nu} \mathcal{E}^{\rho\sigma}_{\mu\nu} \tilde{h}_{\rho\sigma} = \partial_\mu \tilde{h}^{\mu\nu} \partial_\nu \tilde{h} - \partial_\mu \tilde{h}^{\rho\sigma} \partial_\rho \tilde{h}^\mu_\sigma + \frac{1}{2} \partial_\mu \tilde{h}^{\rho\sigma} \partial^\mu \tilde{h}_{\rho\sigma} - \frac{1}{2} \partial_\mu \tilde{h} \partial^\mu \tilde{h}$ describes the linear part of the Einstein action. For $k^2 \gg m^2$, the action becomes diagonal in field space. The individual kinetic terms for $\tilde{h}_{\mu\nu}$ and A_μ correspond to massless linearized Einstein and Maxwell theory, respectively. Thus, in the limit where the mixing of the individual fields can be neglected, $\tilde{h}_{\mu\nu}$ carries precisely the two helicity-2, A_μ the two helicity-1 and χ the single helicity-0 degrees of freedom.

Note that requiring the diagonalization of the kinetic term fixes the relative factor of $1/2$ between the χ-terms in (14). Similarly, the factors of m in (14) normalize the kinetic terms. The coefficient of the kinetic term for χ is determined by the coupling of $h_{\mu\nu}$ to sources: $\int d^4x T^{\mu\nu} h_{\mu\nu}$. The propagator of a massive spin-2 field $h_{\mu\nu}$ between two conserved sources $T_{\mu\nu}$ and $\tau_{\mu\nu}$ is given by

---

[1] For dimension $d < 2$ we just have kinetic or mass mixing.

$$
\begin{aligned}
T^{\mu\nu} D_{\mu\nu,\rho\sigma} \tau^{\rho\sigma} &= T^{\mu\nu} \frac{\left(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho} - \frac{2}{3}\eta_{\mu\nu}\eta_{\rho\sigma}\right)}{p^2 - m^2} \tau^{\rho\sigma} \\
&= T^{\mu\nu} \frac{\left(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho} - \frac{1}{2}\eta_{\mu\nu}\eta_{\rho\sigma}\right)}{p^2 - m^2} \tau^{\rho\sigma} + T^{\mu\nu} \frac{1}{6} \frac{\eta_{\mu\nu}\eta_{\rho\sigma}}{p^2 - m^2} \tau^{\rho\sigma} \ . \quad (19)
\end{aligned}
$$

The first term in the last line corresponds to the helicity-2 state $\tilde{h}_{\mu\nu}$. The second term is an additional interaction from the extra scalar degree of freedom χ and fixes the overall normalization of it in our helicity decomposition. By considering non-conserved sources one can accordingly fix the normalization of A_μ in (14).

For $m = 0$, the action (17) describes linearized Einstein gravity and is invariant under linearized diffeomorphisms,

$$
h_{\mu\nu} \to h_{\mu\nu} + \frac{1}{2}(\partial_\mu \xi_\nu + \partial_\nu \xi_\mu) \ , \quad (20)
$$

where $\xi_\mu(x)$ defines the linear coordinate transformation. The gauge redundancy fixes the relative coefficients of the two-derivative terms. Since both vector and scalar appear with derivatives in the Stückelberg decomposition, the only way for their equations of motion to be second order is for these derivative terms to drop out from the two-derivative kinetic term for $h_{\mu\nu}$. In other words, we impose on the kinetic part of the Lagrangian the condition

$$
\mathcal{L}(h_{\mu\nu}) = \mathcal{L}(h_{\mu\nu} + \partial_{(\mu}\tilde{A}_{\nu)} + \partial_{\mu\nu}\tilde{\chi}) + \text{boundaries} \ , \quad (21)
$$

where $\tilde{A}_\mu$ and $\tilde{\chi}$ are respectively a vector and a scalar. This is equivalent to the gauge invariance (20) for a specific ξ_μ.

The uniqueness of said structure can also be understood from a Hamiltonian analysis. Let us first examine the kinetic term. After having integrated by parts such that h_{00} and h_{0i} do not appear with time derivatives, the canonical momenta of the Lagrangian (17) are

$$
\pi_{ij} = \frac{\partial \mathcal{L}}{\partial \dot{h}_{ij}} = \dot{h}_{ij} - \dot{h}_{ii}\delta_{ij} - 2\partial_{(i}h_{j)0}. \quad (22)
$$

The other canonical momenta (π_{00} and π_{0i}) are zero due to the integration by parts. Inverting (22), one obtains

$$
\dot{h}_{ij} = \pi_{ij} - \pi_{kk}\delta_{ij} + 2\partial_{(i}h_{j)0}. \quad (23)
$$

Performing the Legendre transformation and rewriting the Lagrangian in terms of the canonical momenta yields

$$
\begin{aligned}
\mathcal{L} &= \pi_{ij}\dot{h}_{ij} - \mathcal{H} + 2h_{0i}\partial_j\pi_{ij} + h_{00}(\nabla^2 h_{ii} - \partial_i\partial_j h_{ij}) \ , \\
\text{where} \quad \mathcal{H} &= \frac{1}{2}\pi_{ij}^2 - \frac{1}{4}\pi_{ii}^2 + \frac{1}{2}\partial_k h_{ij}\partial_k h_{ij} - \partial_i h_{jk}\partial_j h_{ik} + \partial_i h_{ij}\partial_j h_{kk} - \frac{1}{2}\partial_i h_{jj}\partial_i h_{kk}.
\end{aligned}
$$

$$(24)$$

The canonical momenta for h_{00} and h_{0i} are zero and the variables themselves appear only linearly in terms without time-derivatives. They are Lagrange multipliers which give the constraint equations $\nabla^2 h_{ii} - \partial_i\partial_j h_{ij} = 0$ and $\partial_j\pi_{ij} = 0$. All these constraints commute, in the sense of Poisson brackets, with each other. Hence, the constraints are first class (for an introduction to constrained systems see for example [10, 11]). This is characteristic for theories with a gauge symmetry. The constraints together with the gauge transformations reduce the physical phase space to a four dimensional hypersurface,

which is described by the canonical coordinates of the two physical polarizations of the massless spin-2 graviton and their conjugate momenta.

Adding a mass term to the analysis changes the Hamiltonian and the Lagrangian of (25) in the following way

$$
\begin{aligned}
\mathcal{L} &= \pi_{ij}\dot{h}_{ij} - \mathcal{H} + m^2 h_{0i}^2 + 2h_{0i}\partial_j\pi_{ij} + h_{00}(\nabla^2 h_{ii} - \partial_i\partial_j h_{ij} - m^2 h_{ii}) \,, \\
\text{where} \quad \mathcal{H} &= \frac{1}{2}\pi_{ij}^2 - \frac{1}{4}\pi_{ii}^2 + \frac{1}{2}\partial_k h_{ij}\partial_k h_{ij} - \partial_i h_{jk}\partial_j h_{ik} + \partial_i h_{ij}\partial_j h_{kk} \\
&\quad - \frac{1}{2}\partial_i h_{jj}\partial_i h_{kk} + \frac{1}{2}(h_{ij}h_{ij} - h_{ii}^2).
\end{aligned}
\tag{25}
$$

Note that the conjugate momenta are unchanged by the additional mass term. However, the structure of the Lagrangian is different and h_{0i} is no longer a Lagrange multiplier. Nevertheless, it is still non-dynamical and its equation of motion yields the algebraic relation

$$
h_{0i} = -\frac{1}{m^2}\partial_i\pi_{ij}.
\tag{26}
$$

h_{00} still is a Lagrange multiplier and it enforces the constraint

$$
\nabla^2 h_{ii} - \partial_i\partial_j h_{ij} - m^2 h_{ii} = 0
\tag{27}
$$

which is now of second class. Requiring that the constraint is conserved in time, i.e. that it commutes with the Hamiltonian, gives rise to a secondary constraint. Since h_{0i} is determined by (26) and h_{00} gives two second class constraints (one primary and one secondary), the resulting physical phase space is then ten dimensional describing the five physical polarizations of the massive spin-2 particle and their conjugate momenta. Departing from the Fierz-Pauli mass term introduces nonlinearities in h_{00} and the constraint which fixes the trace h_{ii} to zero is lost resulting in either a tachyonic or ghost-like sixth degree of freedom [12, 13].

Let us briefly mention coupling to sources. Adding a source term to the Lagrangian (17) of the form $h_{\mu\nu}T^{\mu\nu}$ does not change the linear constraint analysis. No matter whether the source is conserved, $\partial_\mu T^{\mu\nu} = 0$, or not, the source coupling will only introduce h_{00} and h_{0i} linearly and without time derivatives and therefore it will not affect the number of constraints. Note that this holds true for any linear coupling of $h_{\mu\nu}$ to sources.

3 Self-interacting theories

We now focus on the question of self-interactions in massive spin-2 theories. We address subtleties in the construction and inquire whether uniqueness theorems can exist similar to the case of the Einstein theory for a massless spin-2 field.

3.1 Boulware-Deser ghost

Boulware and Deser (BD) argued in [13] that simply introducing a mass term for the full nonlinear theory of general relativity reintroduces the sixth degree of freedom which could be tuned away in the Fierz-Pauli theory. Although this result turned out to be not generic, it is instructive to see their reasoning.

Let us first consider pure general relativity. Using the ADM formalism [14] in which a general metric can be re-written as

$$
ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta = -N^2 dt^2 + \gamma_{ij}\left(dx^i + N^i dt\right)\left(dx^j + N^j dt\right) \,,
\tag{28}
$$

where $\gamma_{ij} \equiv g_{ij}$, $N \equiv (-g^{00})^{-\frac{1}{2}}$ (lapse), $N_i \equiv g_{0i}$ (shift). The full action reads (for simplicity we set the Planck mass to one)

$$S = \int d^4x \sqrt{-g}\, R = \int d^4x (\pi_{ij}\dot{\gamma}_{ij} - NR^{(0)} - N_i R^i - 2(\pi^{ij}N_j - \frac{1}{2}\pi N^i + N^{|i}\sqrt{\gamma})_{|j}), \qquad (29)$$

All curvatures are functions of γ_{ij} and π_{ij}, but do not depend on N or N_i. R is the four dimensional Ricci scalar and $-R^{(0)} \equiv {}^3R + \gamma^{-\frac{1}{2}}(\frac{1}{2}\pi^2 - \pi_{ij}\pi^{ij})$ and ${}^3R$ is the three dimensional Ricci scalar with respect to the metric γ_{ij}. $R^i = -2\pi^{ij}_{|j}$, where the bar "|" denotes covariant differentiation with respect to the spatial metric γ_{ij}.

In the massless theory, N and N_i are Lagrange multiplier which enforce first class constraints on the system, thereby eliminating four (and correspondingly eight phase space) degrees of freedom yielding 2 propagating helicities of the massless spin-2 particle. We now introduce the Minkowski background by expanding

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} , \qquad (30)$$

where $\eta_{\alpha\beta}$ is the Minkowski metric and $h_{\alpha\beta}$ is a tensor on the flat background. Its indices are consequently raised and lowered by the Minkowski metric. The inverse metric $g^{\alpha\beta}$ is given by an infinite series of $h_{\alpha\beta}$ and can be obtained from $g^{\alpha\mu}g_{\mu\beta} = \delta^\alpha_\mu$. At linear order $N = 1 - \frac{1}{2}h_{00}$ and $N_i = h_{0i}$ and one recovers the result of the previous section. At nonlinear order, however,

$$N^2 = (1 - h_{00}) - h_{0i}h_{0j}g^{ij} , \qquad (31)$$

whereas N_i remains unchanged.

The Fierz-Pauli mass term $f = (h_{\mu\nu}h^{\mu\nu} - h^2)$ can nevertheless easily be expressed in terms of N_i and the nonlinear N [13],

$$f = h_{ij}^2 - h_{ii}^2 - 2N_i^2 + 2h_{ii}(1 - N^2 - N_iN^i). \qquad (32)$$

In contrast to the linear case, here N (which to linear order is equivalent to h_{00}) appears quadratically in the mass term although still appearing linearly in the full non-linear derivative (Einstein) structure of the theory. Therefore now neither N nor N_i are Lagrange multipliers.

Thus, at the full non-linear level, the trace h_{ii} is no longer constrained since the constraint was related to the fact that N was a Lagrange multiplier. Therefore, there are six degrees of freedom propagating: The so-called Bouleware-Deser ghost propagates on top of the five degrees of freedom of the Fierz-Pauli massive spin-2. We will see that this conclusion, although correct generically, can be avoided for specific theories.

The simplest example is the free Fierz-Pauli theory discussed above. There, since the expansion is truncated at the linear level we have that $N = 1 - \frac{1}{2}h_{00}$ and the mass term in the Lagrangian is

$$f = h_{ij}^2 - h_{ii}^2 - 2N_i^2 + 2h_{ii}(1 - 2N). \qquad (33)$$

Thus, as in the derivative part of the action (the linearized Einstein-Hilbert Lagrangian), N only appears linearly. In other words, the lapse is here again a Lagrange multiplier, as in General Relativity.

The philosophy of avoiding the BD ghost will be the same for interacting theories: we will search for theories that can be written in terms of a linear lapse function acting as a Lagrange multiplier. In order to do that and to avoid the BD conclusions, we will have to either deform the derivative structure of the massless theory and/or the non-derivative structure.

3.2 Cubic interactions for a massive spin-2 particle

We will start by considering the simplest possible interaction, a cubic interaction as described in [2].

There, the idea was to consider a cubic interaction that keeps the structure of the *linear* Fierz-Pauli action. In other words, by deviating from the Einsteinian derivative structure at the cubic order, $N = 1 - \frac{1}{2}h_{00}$ remains a Lagrange multiplier also in the nonlinear theory. Non-derivative interactions can then be constructed that preserve this property.

This construction can straightforwardly be achieved by considering the most general cubic interaction with at most two derivatives on $h_{\mu\nu}$. Demanding linearity in h_{00} fixes all coefficients besides respective prefactors of the zero- and two-derivative terms.

The unique structure is found to be [2]

$$
\begin{aligned}
\mathcal{L}^{(3)} &= \frac{k_1}{\Lambda^7}\big(h^{\alpha\beta}\partial_\alpha h^{\mu\nu}\partial_\beta h_{\mu\nu} - h^{\alpha\beta}\partial_\alpha h\partial_\beta h + 4h^{\alpha\beta}\partial_\beta h\partial_\mu h_\alpha^\mu - 2h^{\mu\nu}\partial_\alpha h\partial^\alpha h_{\mu\nu} + h\partial_\mu h\partial^\mu h \\
&\quad -3h^{\mu\nu}\partial_\alpha h_\mu^\alpha\partial_\beta h_\nu^\beta - 4h^{\mu\nu}\partial_\nu h_\mu^\alpha\partial_\beta h_\alpha^\beta + 3h\partial_\mu h^{\mu\nu}\partial_\alpha h_\nu^\alpha + 2h^{\mu\nu}\partial^\alpha h_{\mu\nu}\partial_\beta h_\alpha^\beta \\
&\quad -2h\partial_\alpha h\partial_\beta h^{\alpha\beta} + h^{\mu\nu}\partial_\alpha h_{\nu\beta}\partial^\beta h_\mu^\alpha + 2h^{\mu\nu}\partial_\beta h_{\nu\alpha}\partial^\beta h_\mu^\alpha - h\partial_\alpha h_{\mu\nu}\partial^\nu h^{\mu\alpha} - h\partial_\alpha h^{\mu\nu}\partial^\alpha h_{\mu\nu}\big) \\
&\quad +\frac{k_2}{\Lambda^5}\big(2h_\nu^\mu h_\rho^\nu h_\mu^\rho - 3hh_{\mu\nu}h^{\mu\nu} + h^3\big)\,.
\end{aligned}
\tag{34}
$$

In terms of the components of $h_{\mu\nu}$, for example, the non-derivative part is given by

$$
\mathcal{L} = \frac{3k_2}{2}h_{00}(h_{ii}^2 - h_{ij}^2) + \text{terms independent on } h_{00}.
\tag{35}
$$

Hence, h_{00} and h_{0i} appear in the same way as in the free action. We do not display the explicit expression for the derivative part because the expression is rather lengthy. Still, one can easily check that also there h_{0i} remains non-dynamical and can be solved for algebraically, yielding 3 constraints on $h_{\mu\nu}$. Furthermore, h_{00} appears as a Lagrange multiplier in (34) and accordingly eliminates another two degrees of freedom [2].

The fact that the action (34) propagates five degrees of freedom can also be checked in the helicity decomposition (14). Inserting the decomposition into (34) reveals that the corresponding equations of motion are free of higher time derivatives on the helicity components.

Indeed, this is in direct correspondence to the Hamiltonian analysis outlined above. The components h_{00} and h_{0i} are exactly those components of $h_{\mu\nu}$ which can introduce higher time derivatives on the equations of motion as, in terms of helicities, these correspond to $\partial_0^2\chi$, $\partial_0 A_0$, $\partial_0\partial_i\chi$ and $\partial_0 A_i$. Therefore, any action free of higher derivatives on the helicities, requires h_{00} to be a Lagrange multiplier and h_{0i} to be nondynamical.

Up to boundary terms, one can rewrite the above Lagrangian in a compact form [15] as follows

$$
\mathcal{L}^{(3)} = k_1\epsilon^{\alpha_1\ldots\alpha_4}\epsilon^{\beta_1\ldots\beta_4}\partial_{\alpha_1}\partial_{\beta_1}h_{\alpha_2\beta_2}\ldots h_{\alpha_4\beta_4} + k_2\epsilon^{\alpha_1\ldots\alpha_3\sigma_4}\epsilon_{\sigma_4}^{\beta_1\ldots\beta_3}h_{\alpha_1\beta_1}\ldots h_{\alpha_3\beta_3}\,.
\tag{36}
$$

$\epsilon^{\alpha_1\ldots\alpha_4}$ denotes the totally antisymmetric four-tensor in four dimension. From its antisymmetry properties it is then simple to conclude that the constraint structure of the free Lagrangian is preserved. If there is one h_{00} in (36), then there cannot be any other factor of it in that term. Therefore, h_{00} can only appear as a Lagrange multiplier. Terms with h_{0i} can carry at most one time derivative and one power of h_{0i} or only spatial derivatives and at most two powers of h_{0i}; all other terms have spatial indices. Variation with respect to h_{0i}, thus, leads to a constraint equation for itself which defines it algebraically in terms of the components h_{ij}.

3.3 Resummed theories

The second possible route to find nonlinear extensions of the Fierz-Pauli theory is to retain the Einsteinian derivative structure and construct a nonlinear extension of the mass term. In this case one searches for a theory preserving a similar constraint structure of the lapse for the full Einstein theory. This approach was taken in [16].

What we learned from the analysis of BD is that the lapse N cannot be a lagrange multiplier if the the two following assumptions co-exist

– The derivative structure is Einsteinian
– N and N^i are independent variables.

As we are interested in the class of theories that fulfill the former assumption, one needs to relax the latter, thereby keeping N as a Lagrangian multiplier such to eliminate the BD ghost. The theory with this property has been found by [16], the so-called dRGT massive gravity.

In other words, the theory of [16] is a deformation of General Relativity with non-derivative term such that [17]

– The derivative structure is Einsteinian.
– N^i can be fully traded by a new variable $n^i(N, N^i, h_{ij})$.
– After the field redefinition, N only appears linearly in the action and without derivatives.

The second condition forces the redefinition to be of type

$$N^i = \left(\delta^i_j + N D^i_j\right) n^j \,, \tag{37}$$

where D^i_j is an appropriate matrix independent upon N.[2] Of course, any truncation in powers of h_{ij} of this construction would bring back the BD ghost.

For our purposes, we adopt the notation of [17]. We write the resummed theory of [16] in terms of the inverse metric g^{-1} and an auxiliary background metric η. The action of dRGT massive gravity can then be written according to [17] (here we re-introduce the Planck mass M_P)

$$S = M_P^2 \int d^4x \sqrt{-g} \left[R(g) + 2m^2 \sum_{n=0}^{2} \beta_n e_n \left(\sqrt{g^{-1}\eta} \right) \right] \,, \tag{38}$$

where m is the graviton mass and the $e_n(\mathbb{X})$ are functions of matrix traces given by

$$e_0(\mathbb{X}) = 1 \,, e_1(\mathbb{X}) = [\mathbb{X}], e_2(\mathbb{X}) = \frac{1}{2}([\mathbb{X}]^2 - [\mathbb{X}^2]) \,. \tag{39}$$

The square brackets denote the trace and $\beta_0 = 6$, $\beta_1 = -3$ and $\beta_2 = 1$ for dRGT massive gravity [16,17]. Note that the coefficients are chosen such that the action describes a flat background without a cosmological constant. The matrix $\sqrt{g^{-1}\eta}$ is defined by $\sqrt{g^{-1}\eta}\sqrt{g^{-1}\eta} = g^{\mu\nu}\eta_{\nu\rho}$. Since $\eta_{\mu\nu}$ transforms as a rank-two tensor, the action (38) is invariant under general coordinate transformations.

Expanding the action (38) to second order in the metric perturbations $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, one recovers the Fierz-Pauli action (17).

As suggested in [16] and later shown in [17–20], the action (38) indeed only propagates five degrees of freedom. In order to see this, one can redefine the shift as (37). This has been also done for the full nonlinear action in [17]. We will briefly discuss their findings.

---

[2]Note that we would get exactly the same results by not doing the field redefinition. This field redefinition only makes manifest that the lapse is a Lagrange multiplier.

A constraint analysis is most conveniently carried out by using the ADM decomposition [14]. Using (28), the Lagrangian (38) is given by

$$M_P^{-2}\mathcal{L} = \pi^{ij}\partial_t\gamma_{ij} + NR^0 + R_iN^i + 2m^2\sqrt{\det\gamma}\,N\sum_{n=0}^{2}\beta_n e_n\left(\sqrt{g^{-1}\eta}\right)\,. \tag{40}$$

The mass term includes N in a non-linear way and is, therefore, responsible for the seeming loss of the constraint. However, one can redefine the shift N_i by (37) and finds the following Lagrangian [17, 19]

$$M_P^{-2}\mathcal{L} = \pi^{ij}\partial_t\gamma_{ij} - \mathcal{H}_0(\pi^{ij},\gamma_{ij},n_j) + N\mathcal{C}(\pi^{ij},\gamma_{ij},n_j)\,, \tag{41}$$

where $\mathcal{H}_0$ is the Hamiltonian and $\mathcal{C}$ is the additional constraint ensuring that only five of the six components of γ_{ij} are propagating. Thus, we have established that there are three independent variables n_i which are not propagating and algebraically determined by their equations of motion and there is one Lagrange multiplier N which yields a constraint equation for π^{ij} and γ_{ij}. Therefore, there are only five propagating independent degrees of freedom which constitute the massive graviton.

It is, however, important to note that the redefinition (37) can only be used when considering the full non-linear action (38). Whenever truncating the theory, this ceases to be valid and thus one is left with six propagating degrees of freedom.

One might be puzzled by analyzing the action (38) in terms of the helicity decomposition (14). There, indeed, are higher derivatives (apparently signaling new degrees of freedom) appearing on the equations of motion of, e.g., the scalar helicity χ for the nonlinear terms [2]. The lowest suppression scale of these terms is $\Lambda_5 = (m^4 M_P)^{\frac{1}{5}}$. In the full theory, however, this scale is redundant and can be removed by a field redefinition [18][3]. With this field redefinition also the higher derivative terms disappear. It can be shown that this happens for all scales below $\Lambda_3 = (m^2 M_P)^{\frac{1}{3}}$ [18], such that the theory in terms of the redefined fields is free of higher derivative interactions. Note, however, that when truncating the theory to any finite order this is no longer the case.

In works subsequent to [16, 17], it was furthermore shown that the absence of the sixth ghost like degree of freedom can also be confirmed in the Stückelberg language [20].

4 Conclusions

Within this work we have addressed the question whether theories of a single interacting massive-spin 2 field obey similar uniqueness theorems as in the massless case.

For a long time, it was doubtful whether there even exists one consistent theory that describes self-interactions of a massive spin-2 particle. The fact that adding the Fierz-Pauli mass term to the Einstein-Hilbert action introduces nonlinearities in the lapse into the action was taken as the basis of a no-go-theorem for nonlinear extensions of Fierz-Pauli theory. It was argued that any such extension necessarily leads to the appearance of a sixth unphysical and ghost-like polarization in the theory, the Boulware-Deser ghost.

We have reviewed two possible ways to circumvent this apparent theorem. One is to sacrifice the Einsteinian derivative structure, such that the (00)-component of the tensor field $h_{\mu\nu}$ enters the action only linearly even when self-interactions are added. This ensures that only five degrees of freedom are propagating. We have shown that this property can equivalently be checked in a helicity decomposition of the massive tensor. The found action is characterized by the absence of higher derivatives on the helicity components. It is the unique theory with this property.

---

[3]Note that this field redefinition is in fact necessary in order to define a Hamiltonian in terms of the helicities. Without it, the relation between the canonical momenta and time derivatives of fields is not invertible. This reflects the redundancy of the coupling in the full theory.

The second route is to leave the derivative structure untouched, but instead adding a potential for the massive spin-2 field in such a way that guarantees the presence of a Lagrange multiplier in the system. By casting the action into an appropriate form, this Lagrange multiplier is once again given by the lapse.

The latter approach, since it relies on redundancies of the full action, requires a full resummation of the theory. Any truncation to finite order appears to propagate more than five degrees of freedom. However, the scale at which this additional degree of freedom appears coincides with the scale at which nonlinearities become important. Henceforth, conclusions can only be drawn from the resummed theory.

We have further addressed the issue of higher derivatives in the helicity decomposition in the latter class of theories. While these are present, the fact that redundancies are present prevents one from constructing a Hamiltonian. A field redefinition is necessary in order to be able to invert the canonical momenta; after this redefinition, no more higher derivatives are present. The Hamiltonian of the theory does not suffer from an Ostrogradski linear instability.

The experimental viability of either theory is unknown. The former deviates from the well probed Einsteinian cubic vertex and is therefore not viable as a massive graviton. The latter has the correct vertex structure. However, choosing its mass to be of the order of the Hubble scale leads to a strong coupling already at very low energies, $\Lambda_s = (m^2 M_P)^{1/3} \sim (1000 \text{ km})^{-1}$.

Either theory could in principle describe self-interactions of a massive spin-2 meson and could therefore be of different phenomenological interest.

References

[1] S. Weinberg and E. Witten, Phys. Lett. B **96**, 59 (1980).

[2] S. Folkerts, A. Pritzel and N. Wintergerst, arXiv:1107.3157 [hep-th].

[3] K. Hinterbichler, Rev. Mod. Phys. **84**, 671 (2012) [arXiv:1105.3735 [hep-th]].

[4] G. Dvali and C. Gomez, arXiv:1005.3497 [hep-th].

[5] G. Dvali, S. Folkerts and C. Germani, Phys. Rev. D **84**, 024039 (2011) [arXiv:1006.0984 [hep-th]].

[6] G. Dvali, G. F. Giudice, C. Gomez and A. Kehagias, JHEP **1108**, 108 (2011) [arXiv:1010.1415 [hep-ph]].

[7] S. Weinberg, Phys. Rev. **135**, B1049 (1964). S. Weinberg, Phys. Rev. **138**, B988 (1965).

[8] V. I. Ogievetsky, I. V. Polubarinov, Annals of Physics, Volume 35, Issue 2, November 1965, Pages 167-208,

[9] M. Fierz and W. Pauli, Proc. Roy. Soc. Lond. A **173**, 211 (1939).

[10] M. Henneaux and C. Teitelboim, Princeton, USA: Univ. Pr. (1992) 520 p

[11] Dirac, P. A. Maurice, DoverPublications.com, 2001

[12] P. Van Nieuwenhuizen, Nucl. Phys. B **60**, 478 (1973).

[13] D. G. Boulware and S. Deser, Phys. Rev. D **6**, 3368 (1972).

[14] R. L. Arnowitt, S. Deser and C. W. Misner, Phys. Rev. **117**, 1595 (1960). R. L. Arnowitt, S. Deser and C. W. Misner, Gen. Rel. Grav. **40**, 1997 (2008) [gr-qc/0405109].

[15] K. Hinterbichler, arXiv:1305.7227 [hep-th].

[16] C. de Rham, G. Gabadadze and A. J. Tolley, Phys. Rev. Lett. **106**, 231101 (2011) [arXiv:1011.1232 [hep-th]].

[17] S. F. Hassan and R. A. Rosen, Phys. Rev. Lett. **108**, 041101 (2012) [arXiv:1106.3344 [hep-th]].

[18] C. de Rham, G. Gabadadze and A. J. Tolley, JHEP **1111**, 093 (2011) [arXiv:1108.4521 [hep-th]].

[19] S. F. Hassan and R. A. Rosen, JHEP **1204**, 123 (2012) [arXiv:1111.2070 [hep-th]].

[20] C. de Rham, G. Gabadadze and A. J. Tolley, Phys. Lett. B **711**, 190 (2012) [arXiv:1107.3820 [hep-th]].