Multi-Gravity: Can You Have Too Much of A Good Thing?

J. H. C. Scargill

Oriel College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2016

Abstract

The ongoing quest to map the space of possible, consistent field theories, and to determine their properties, is extremely interesting from a purely theoretical perspective, as well as being a prerequisite for the phenomenological application of any such theories. To that end, this thesis is concerned with various theoretical aspects of multi-gravity theories—that is, theories which contain multiple, interacting spin-2 fields.

The analysis presented here will proceed by using the Stückelberg trick to reintroduce the multiple copies of diffeomorphism invariance which the theory would possess if none of the spin-2 fields interacted, and it is discussed in detail how to apply this to multi-gravity, and in particular how this differs from massive and bi-gravity.

The structure of the interactions which arise from the helicity-0 modes of the massive gravitons is examined in the so-called decoupling limit. It is found that the structure these take in bi-gravity is generalised to that of multi-Galileons, although this is not always manifest; an extension of the so-called Galileon duality is used to probe this.

Multi-gravity theories can be elegantly described using graphs to encode the network of interactions between the different fields, and this forms the basis for the other two questions which will be covered. When the theory contains a cycle of interactions, it is explicitly shown how this leads to the introduction of a ghost-like instability which is not present otherwise, and how this lowers the cutoff of the effective theory. In the absence of cycles, the structural properties of the graph which represents a given theory continue to play a large role in determining the scale at which the effective theory breaks down. This is the final topic, which is studied in detail, and upper and lower bounds on this scale are derived which depend on various properties of the graph.

Acknowledgements

I am, of course, immensely grateful to my two supervisors, Pedro Ferreira and John March-Russell for the invaluable advice and encouragement which they have provided over the course of my DPhil. Meanwhile great thanks also go to my collaborator Johannes Noller, with whom I completed much of the work presented in this thesis, and with whom I have had innumerable useful discussions. Finally I am grateful to all of the above, as well as to James Bonifacio and Macarena Lagos, for reading a draft of this thesis, and for providing useful feedback. Thanks also go to the Science and Technology Facilities Council for funding my DPhil, as well as to Oriel College for additional funding.

Statement of Originality

This thesis is based on original research and contains no material that has already been accepted, or is concurrently being submitted, for any degree or diploma or certificate or other qualification in this university or elsewhere. To the best of my knowledge and belief this thesis contains no material previously published or written by another person, except where due reference is made in the text.

J. H. C. Scargill 2016

Contents

CHAPTER

1

Introduction

The question of which field theories are consistent, and if so what are their properties, is one which is interesting not only for its own sake—to better understand what the rich tapestry of nature and mathematics allows—but also because such theories might prove to have practical applications in models of fundamental physics, which we would not know or fully understand unless these unexplored patches of theory space are examined. To that end, this thesis is devoted to a particular class of field theories which have attracted a great deal of interest over the last five years (though their full history is in fact older): theories in which there are multiple, different spin-2 fields which have non-linear interactions between them (as well as non-linear self-interactions).

Before continuing I wish to make a brief note about nomenclature. These theories are commonly called 'multi-gravity' because the non-linear completion of a self-interacting, massless spin-2 field yields Einstein gravity (see *e.g.* [1, 2], though also [3]), and so in a very real sense the same applied to multiple spin-2 fields (now not all massless, as described below) will yield multiple gravitational theories, interacting amongst themselves. However some people object to this name on the grounds that there is only *one* 'gravity,' which depends on how matter couples to the theory, and considering the rest of the content to just consist of new fields. Whilst I am somewhat sympathetic to that viewpoint (though at the risk of getting bogged down in semantics), I am not aware of any appealing alternatives: 'multi-metric' and 'multi-vielbein' are also used, however these refer to two different theories which are not always equivalent (see section 2.5.3 and chapter 5); thus I shall use the term 'multi-gravity,' and will also refer to the spin-2 perturbations of these fields as 'gravitons.'

There is a powerful no-go result due to Boulanger, Damour, Gualtieri, and Henneaux [4] which states that theories of multiple interacting *massless* gravitons are inconsistent, and so in these theories all but one of the spin-2 fields must be *massive*; thus a theory of multi-gravity necessitates a theory of massive gravity. The linear theory of a massive spin-2 field was developed by Fierz and Pauli in 1939 [5], however it took a while for a satisfactory non-linear completion of this to be found, in part due to an apparent no-go result of Boulware and Deser [6];[1] in 2010 a satisfactory non-linear theory of massive gravity was discovered by de Rham, Gabadadze, and Tolley [9, 10], and subsequently developed by them and others [11–16]; it is this 'dRGT theory,' also called 'ghost-free massive gravity,' which forms the basis for the theories considered here.

In dRGT massive gravity there is just one massive graviton, however it can be easily extended to include also a massless one—a theory sometimes called Hassan-Rosen (HR) bi-gravity [17,18]—and further to true multi-gravity which contains one massless graviton and an arbitrary number of massive ones [19]. For reviews of massive gravity see [20,21], and for bi-gravity see [22].

For the most part this thesis will be concerned with purely theoretical aspects of these theories, and thus it is appropriate now to discuss some of their phenomenological applications. As mentioned above, theories of multi-gravity are intimately entwined with those of massive gravity, an original motivation for which was in cosmology to explain the observed accelerated expansion of the universe. There are two aspects to this: on the one hand it is hoped that as the finite mass of the graviton renders gravity finite ranged, it can act essentially as a 'high-pass' filter, de-gravitating the infinite wavelength

[1] It has recently been pointed out [7] that a counter-example in fact had already been discovered a year previously [8], though it seems not to have received any attention at the time.

mode that is the cosmological constant [23–25]; on the other hand dRGT also allows cosmological solutions whose expansion accelerates even in the absence of a cosmological constant (so-called self-acceleration), for example see [26, 27].

Unfortunately there are a few issues which prevent this from being a perfect solution. In four dimensional space-time, a massless spin-2 field has two helicity degrees of freedom (*d.o.f.*) whilst a massive spin-2 field has five; these can be decomposed into those corresponding to massless spin-2 (2 *d.o.f.*), massless spin-1 (2 *d.o.f.*), and spin-0 (1 *d.o.f.*), and in the linear theory the helicity zero part leads to an anomalous bending of light, which remains even as the massless limit is taken—this is the famous van Dam-Veltman-Zakharov (vDVZ) discontinuity [28–30]; a resolution is provided by the Vainshtein screening mechanism [31–34] which involves including non-linear terms that become dominant in the vicinity of a massive object and screen the effects of the troublesome helicity zero mode. Both of these effects will be discussed in more detail in section 2.3, but schematically the way the Vainshtein mechanism works is that the kinetic term for the (classical or quantum) fluctuations φ about a background π becomes $Z^{\mu\nu}(\pi)\partial_\mu\varphi\partial_\nu\varphi$, where $Z^{\mu\nu}(\pi)$, which is large, encodes the effect of the dominant non-linear terms; canonically normalising via $\varphi \to \varphi/\sqrt{Z}$ then has the effect in higher order interaction terms of increasing the suppression scale to $\Lambda\sqrt{Z} \gg \Lambda$ [35]. This Vainshtein screening is thus an observational requirement (and is one reason why the Fierz-Pauli mass term requires a non-linear completion), however unfortunately it is somewhat in conflict with the de-gravitation idea mentioned above. This was studied in [36], and it was found that screening a larger cosmological constant reduces the scale at which the Vainshtein effect is active, thus raising tensions with solar system observations, with the result that a vacuum energy only as large as $10^{-3}\,\mathrm{eV}$ could be de-gravitated. Therefore one has to assume some other mechanism for removing a large vacuum energy.

A second issue is that these theories typically become strongly-coupled at an energy scale $\sim (m^2 M_{\mathrm{Pl}})^{1/3}$, where m is the mass of the graviton, however explaining the current acceleration requires $m \sim H_0 \sim 10^{-33}\,\mathrm{eV}$, leading to a very low strong coupling scale of the order of $(10^3\,\mathrm{km})^{-1}$, which some may deem unacceptable. However it is worth noting three things: i) as explained above, this value is environment dependent [21, 37], being redressed, and raised, by effects from a non-trivial background;[2] ii) a low strong-coupling scale is required for the Vainshtein mechanism to function, as it relies on non-linearities becoming dominant,[3] and the Vainshtein effect is a phenomenological necessity; iii) this is just the scale at which perturbative unitarity is lost and the theory becomes strongly coupled, whilst the scale at which new physics enters may be much higher. On this last point see [39] in which two effective theories were studied—chiral perturbation theory (for variable number of colours and flavours) and gravity coupled to many massless scalar fields—which both have the property that violation of tree-level unitarity does not lead to new physics, but that the effective theory can 'self-heal' and restore unitarity via iteration of scattering diagrams. Also, see [40, 41] for recent interesting work on the possibility of increasing the strong coupling scale to $(mM_{\rm Pl})^{1/2} \sim (0.1\,\mathrm{mm})^{-1}$, which is more palatable.

A further intruiging application of bi-gravity is the possibility for the massive graviton to be the dark matter particle, heretofore observed only gravitationally. This has been examined in recent work, in which it was found that observational constraints could be satisfied with a graviton mass $\sim 0.01\,\mathrm{GeV}$ [42] or $10^3\,\mathrm{GeV} \lesssim m \lesssim 10^{11}\,\mathrm{GeV}$ [43] (where the difference is due to differences between the specifics of the models, such as the production mechanism considered).

Another area in which multi-gravity finds application is in theories to describe spin-2 states in QCD. This was first investigated by Isham, Salam, and Strathdee [44], who proposed a model in which the (ordinary, massless) graviton mixes with a massive, neutral spin-2^+ meson, and this so-called 'strong gravity' received much attention at the time (for example see [45] for a contemporary review). It is interesting to note as well that in a universe in which QCD has a parametrically lightest state with $J^P = 2^+$, then this would be described by a massive-gravity theory, which then receives a UV completion in the form of QCD (though of course, for better or for worse, our universe is not like this).

Finally, multi-gravity theories also arise naturally in the context of dimensional deconstruction. This involves taking a Kaluza-Klein theory, *i.e.* a higher-dimensional theory with one of its dimensions compactified and viewed from the lower-dimensional perspec-

[2] Although it is not clear to what extent this does actually help [38].

[3] Note that non-renormalisation properties of the operators important for the Vainshtein effect mean that there exists a regime in which they can be taken to be dominant, but quantum corrections can be safely ignored [21].

tive, and approximating it by a series of interacting lower-dimensional fields corresponding to different points in the compactified dimension, effectively discretising it [46]; the Kaluza-Klein theory has an infinite tower of massive states, which is truncated in the dimensionally deconstructed version.

Before discussing the case of gravity, it is perhaps helpful to first consider Yang-Mills. In that case one arrives at a theory of multiple gauge fields which have nearest-neighbor interactions with one-another through a series of non-linear sigma model fields which transform under the fundamental representation of one gauge group and under the anti-fundamental representation of the next one along [46]. Through the Higgs mechanism these interactions then lead to a theory of one massless spin-1 field and a tower of massive spin-1 fields.

In the case of gravity, take GR in higher dimensions, perform the same procedure, and one is left with a theory of one massless, and a finite tower of massive, spin-2 fields interacting amongst themselves [47–51], furthermore it has been shown that by performing the discretisation in the right way, the interactions at which one arrives are precisely those of dRGT [50]. The case of gravitational dimensional deconstruction is additionally theoretically interesting as, whilst it is simple to discretise the extra dimension, it is not currently known how to perform the inverse procedure of taking the continuum limit and recovering the Kaluza-Klein theory. I will touch on dimensional deconstruction again, later in this thesis, in chapters 5 and 6.

Although throughout this thesis I will focus on multi-gravity theories constructed from dRGT-type interaction terms, I will very briefly mention two ways in which alternative multi-gravity theories may arise [52]. The first is non-commutative geometry: here the set-up consists of the product of a continuous space with a discrete one, and different points of the latter then naturally will lead to different metrics on the former (in a superficially very similar way to the dimensional deconstruction discussed above).

Theories of multi-gravity also naturally arise in the context of brane-world models. Obviously in this case, if the extra dimension is compact, one has the usual Kaluza-Klein excitations, which will include massive spin-2 particles, however I would like specifically to mention the scenario first presented in [53] (and see [54] for a brief review of this and similar models). In an AdS bulk there are two positive tension branes separated by a negative tension one; due to the warp factor in the bulk, which peaks on each of the positive tension branes, in the limit that these are infinitely separated there is a massless graviton state localised on each; when the separation is finite, however, these two states will mix, and in the usual way from quantum mechanics one ends up with one massless and one massive mode. In contrast to the normal single brane picture, in which the massive gravitons all arise as Kaluza-Klein excitations, in this case this first mass gap can be much less than the mass of the next lightest mode, leading to a new effective bi-gravity model.

Another brane-world model which is important in the context of massive gravity is the Dvali-Gabadadze-Porrati (DGP) model [55–57] which involves a four dimensional brane in an infinite, five dimensional bulk. The fact that the bulk is infinite means that there are no Kaluza-Klein excitations, however on the brane the (Källen-Lehman) spectral density of the graviton becomes non-zero at non-zero mass (though still peaked at zero mass), and so in this sense there is a massive graviton on the brane. Just as in other massive gravity theories, this can yield self-accelerating cosmological solutions [58, 59], however degravitation turns out not to function correctly [25]. Another interesting thing to note is that there is a ghost instability around the self-accelerating background [35, 60], however this is *not* the Boulware-Deser (BD) ghost mentioned above, but is one of the five *d.o.f.* of the massless five-dimensional/massive four-dimensional graviton.

The DGP model is also interesting because the helicity-0 mode of the massive graviton (which is encoded in the brane-bending mode) has derivative self-interactions which turn out to take a special form: they are given by the cubic Galileon term $(\partial\pi)^2\Box\pi$ [60], and this led to the (re)discovery of general Galileon terms [61]. Since then, Galileons have appeared in other places, for instance in Dirac-Born-Infeld (DBI) theories as the position modulus of a (slowly moving) four dimensional brane in a five dimensional bulk [62]—they turn up in dRGT massive gravity as well, and will prove to play a crucial role in the sequel.

It behoves me also to mention the issue of superluminality and acausality. Several papers have indicated that dRGT massive gravity potentially allows for superluminal propagation [63–67], and hence acausality, and this possibility would seem to provide very strong theoretical evidence against it, and seemingly by extension against multi-gravity. These claims should not be taken at face value however.

One argument against superluminality being a problem is the fact that it can appear in theories which we know are healthy, for example QED with electrons, in the curved background spacetime of a black hole, shows superluminalities when the electrons are integrated out [68]. Goon and Hinterbichler have examined this [69], and found that it is not possible to induce enough superluminality to produce a difference (compared to luminal propagation) which is large enough to be within the range of validity of the effective theory; however this is in stark comparison with the superluminality problems which plague Galileon theories (which are similar, but not identical to the problems in dRGT), in which the difference does seem to be within the range of validity of the theory. Thus this argument on its own cannot solve the problem.

There are however many other objections which one can raise, and de Rham covers them very well in her review [21], and so I will here just briefly summarise her comments. Essentially all of the examples demonstrating superluminality either rely on unphysical backgrounds (as is the case for [63]), depend on local solutions for which no known sensible global extension exists [64], rely on an incorrect use of characteristic analysis [65], or lie beyond the regime of validity of the theory [66, 67], and no examples of closed timelike curves which lie within the regime of validity have been discovered. Furthermore causality is ultimately determined by the *front* velocity [70], *i.e.* the speed of propagation of the first (potentially infinitesimal) disturbance to reach an observer after the field is perturbed elsewhere, and this requires considering modes with arbitrarily high frequency, and hence adequate determination of the (a)causality of a theory cannot rely on purely classical calculations.[4]

Finally, it is also worth pointing out that some Galileon theories which exhibit superluminal behaviour are known to be dual (in the sense of the Galileon duality, described in section 3.3) to a free field [73], which is trivially healthy. This of course does not mean that superluminalities in all Galileon theories (and by extension massive gravity theories) are necessarily spurious, but it does indicate the need for more careful consideration. See [74–76] for further discussion on this point.

[4] This also provides another resolution of the issue of superluminality in QED on a curved background [71, 72].

This thesis is structured in the following way: chapter 2 continues the introduction by describing in more detail the dRGT theory of massive gravity, its extension to multi-gravity, as well as a useful way of describing the structure of a given theory in terms of a graph; chapter 3, which is partly based on [77], discusses how to analyse these theories using the Stückelberg trick—which is the approach that will be used throughout this thesis—and highlights those aspects which are different from the case of massive and bi-gravity; this is closely followed by chapter 4—based on [78]—in which these theories are examined in the decoupling limit, which effectively separates out the different helicity components of the massive gravitons, and contact is made with the so-called Galileon dualities; the next two chapters, which are from [79] and [80] respectively, then examine effects due to the graph structure of the interactions—5 considers cycles of interactions and explicitly demonstrates, in such a circumstance, the appearance of a ghost-like instability in the metric version of the theory which is absent in the vielbein version, whilst chapter 6 considers to what extent the strong-coupling scale of a theory is controlled by properties of its theory graph; finally chapter 7 concludes; four appendices contain additional details on some of the calculations.

CHAPTER

2

Multi-Gravity

In this chapter, after first covering the linear theory of spin-2 fields, I will briefly explain the link between multi- and massive gravity in 2.2, as well as some of the classic problems facing massive gravity in 2.3, before introducing in detail the dRGT theory of massive gravity in 2.4, in both its metric and vielbein formulations. In section 2.5 its extension to bi- and multi-gravity is covered, and I discuss how one can use graphs to describe the structure of a given multi-gravity theory. I will conclude by quoting some results from the literature concerning the coupling of matter to these theories, the possibility of introducing non-minimal kinetic terms, the effect of quantum corrections, and so-called partially massless spin-2 theories, in sections 2.6 to 2.9.

2.1 The linear theory of spin-2 fields

Let me start by covering the linear theory (by which I mean that the action is quadratic in the fields) of spin-2 fields, both massless and massive, as studied by Fierz and Pauli many years ago [5]; I will keep the dimension of spacetime, D, arbitrary.

A massless spin-2 field $h_{\mu\nu}$, which is symmetric in its two indices, propagating on Minkowski space has the following Lagrangian

$$\mathcal{L}_{m=0} = -h^{\alpha}_{\mu}\mathcal{E}^{\mu\ \ \nu}_{\ \alpha\ \beta}h^{\beta}_{\nu}, \tag{2.1}$$

where indices are raised by the Minkowski metric $\eta_{\mu\nu}$, and the kinetic operator is

$$\mathcal{E}^{\mu\ \ \nu}_{\ \alpha\ \beta} = \frac{1}{2}\delta^{\mu\nu\lambda}_{\alpha\beta\rho}\partial^{\rho}_{\lambda} = \delta^{\mu}_{[\alpha}\delta^{\nu}_{\beta]}\Box + \delta^{\nu}_{[\alpha}\partial^{\mu}_{\beta]} + \delta^{\mu}_{[\beta}\partial^{\nu}_{\alpha]}, \tag{2.2}$$

where I have used the shorthand notation $\partial_{\mu\nu\cdots} = \partial_\mu\partial_\nu\cdots$, square brackets around indices denote anti-symmetrisation, and the generalised Kronecker delta is defined by

$$\delta^{\mu_1\cdots\mu_n}_{\nu_1\cdots\nu_n} \equiv \frac{1}{(D-n)!}\epsilon^{\mu_1\cdots\mu_n\lambda_1\cdots\lambda_{D-n}}\epsilon_{\nu_1\cdots\nu_n\lambda_1\cdots\lambda_{D-n}}. \tag{2.3}$$

The kinetic term (which one recognises as the linearisation of the Einstein-Hilbert Lagrangian) takes this form because it allows (2.1) to be invariant under the following gauge transformation

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu\xi_\nu + \partial_\nu\xi_\mu \tag{2.4}$$

(which one recognises as a linearised diffeomorphism). The equations of motion (*e.o.m.*) arising from (2.1) are

$$-2\mathcal{E}^{\mu\ \ \nu}_{\ \alpha\ \beta}h^{\beta}_{\nu} = 0, \tag{2.5}$$

and one notes that, due to the symmetry properties of (2.2), the 00 and $0i$ components of this will contain 0 and 1 time derivatives respectively, and thus are constraint equations. Therefore, of the initially $\frac{1}{2}D(D+1)$ *d.o.f.* in $h_{\mu\nu}$, D are eliminated by constraints, and a further D are eliminated by the gauge invariance, leaving one with $\frac{1}{2}D(D-3)$ *d.o.f.*—precisely the correct number for a massless spin-2 field (see *e.g.* [81]).

A massive spin-2 field meanwhile has the Lagrangian

$$\mathcal{L}_{m\neq 0} = -h^{\alpha}_{\mu}\mathcal{E}^{\mu\ \ \nu}_{\ \alpha\ \beta}h^{\beta}_{\nu} - \frac{1}{2}m^2\left(h_{\mu\nu}h^{\mu\nu} - (h^{\mu}_{\mu})^2\right), \tag{2.6}$$

which no longer possesses the gauge symmetry (2.4) as the mass term is not invariant. Its *e.o.m.* are

$$-2\mathcal{E}^{\mu\ \ \nu}_{\ \alpha\ \beta}h^{\beta}_{\nu} - m^2\left(h^{\mu}_{\alpha} - h\delta^{\mu}_{\alpha}\right) = 0, \tag{2.7}$$

whose divergence and trace yield

$$\partial_\mu h^\mu_\alpha - \partial_\alpha h = 0, \qquad \text{and} \qquad \Box h - \partial^\mu_\nu h^\nu_\mu - \frac{D-1}{D-2} m^2 h = 0, \tag{2.8}$$

respectively. These can be combined to yield the following $D+1$ constraints[1]

$$\partial^\mu h_{\mu\nu} = 0 \qquad \text{and} \qquad h = 0, \tag{2.9}$$

which reduce the number of *d.o.f.* in $h_{\mu\nu}$ to $\frac{1}{2}(D+1)(D-2)$—precisely the correct number for a massive spin-2 field (again see *e.g.* [81]).

Note that if the relative coefficient between the two parts of the mass term in (2.6) were anything other than -1, one would not have the constraint $h = 0$, and the theory would propagate one-too many *d.o.f.* (and the extra *d.o.f.* turns out to be ghostly). This is called *Fierz-Pauli tuning*, and it is a requirement of any non-linear theory of massive spin-2 that it possesses this in the linearised limit.

2.2 The necessity of massive gravity

A priori one may wonder about the link between multi-gravity and massive gravity—after all, could one not imagine a theory in which there are different spin-2 fields, between which there exist non-linear interactions, and yet which nonetheless possess a gauge symmetry enforcing them to be massless? If such a theory did exist, then one would have a theory of multi-gravity without any need for a theory of massive gravity, and thus potentially hope to avoid the several problems associated with such—the BD ghost, the vDVZ discontinuity, *etc.*—as described in the next section.

Unfortunately it turns out that such a theory does not, and cannot, exist, under some fairly uncontroversial assumptions, as proven by Boulanger *et al.* in 2000 [4]. Their starting point is a free field theory consisting of multiple, decoupled copies of the linear action for a massless spin-2 field (2.1), which possesses a gauge symmetry consisting of multiple, decoupled linear gauge symmetries (2.4) ensuring the masslessness of each field. They then go on to prove that, under the assumptions of

[1]Note that it is still the case that D of the *e.o.m.* are constraints, but these will be degenerate with D of the constraints presented here.

- locality;
- Poincaré invariance;
- spacetime dimension $D > 2$;
- at most two derivatives in each term in the Lagrangian,

the only non-linear deformation of the initial theory which is consistent, *i.e.* for which the deformed action is invariant under the deformed gauge symmetry, consists of multiple, *decoupled* copies of the Einstein-Hilbert action. Furthermore, they showed that it is not possible for the different gravitational sectors to interact indirectly, through matter (in the form of a scalar field) which couples to two or more gravitons simultaneously. The proof of this theorem, which is based on BRST cohomology, is fairly long and involved, and so, despite its importance, I simply refer the reader to the original paper.

Whilst the first three assumptions seem innocuous,[2] the fourth one, demanding that the Lagrangian contain at most two derivatives per term, might give one pause for thought. Having higher derivatives in the Lagrangian will generically lead to issues with instabilities and negative energy states [83, 84], as I discuss in more detail in section 3.2.1.1, and thus it is probably better not to do away with this assumption if one wishes to retain a healthy theory. That being said, in [4] it is also proven that without this assumption on the number of derivatives it remains the case that the only consistent non-linear deformation of the initial theory has a deformed gauge symmetry algebra which, to first order in the deformation parameter, is simply a direct sum of independent diffeomorphism (diff) algebras (*i.e.* it is not modified in a non-trivial way compared with the case of decoupled fields).

Therefore one sees that, under the four assumptions described above, a consistent theory of multi-gravity, *i.e.* a theory of multiple spin-2 fields with interactions between them, can contain at most one massless spin-2 field (for each decoupled sector), with the rest being massive. Thus, a consistent theory of multi-gravity necessitates a consistent, non-linear theory of a massive spin-2 field, *i.e.* a theory of massive gravity. Fortunately such a theory exists, and it is presented in section 2.4.

[2] Though of course the denizens of 'Lineland' and 'Pointland' [82] might take issue with the third one!

2.3 Obstructions to massive gravity

Before discussing the non-linear theory of massive gravity which forms the bedrock for this thesis, I will discuss briefly some of the classic problems facing massive gravity which were mentioned in the introduction: the vDVZ discontinuity, and its resolution via the Vainshtein mechanism, and the Boulware-Deser ghost. See *e.g.* [20] for further detail.

2.3.1 vDVZ discontinuity

The essence of the vDVZ discontinuity [28–30] is that, when coupled to matter, the massless limit of the linear theory of massive spin-2 described by (2.6) is not smooth. To arrive at this conclusion one must first couple the theory to the energy-momentum tensor for matter fields, $T^{\mu\nu}$, which is done, in four dimensions, by adding to the Lagrangian a term $\frac{1}{M_{\rm Pl}}h_{\mu\nu}T^{\mu\nu}$. Now consider the field external to a stationary point mass, *i.e.* set $T^{\mu\nu} = M\delta^\mu_0\delta^\nu_0\delta^{(3)}(\mathbf{x})$, and solve for $h_{\mu\nu}(\mathbf{x})$—qualitatively one finds $h \sim \frac{1}{4\pi r}\mathrm{e}^{-mr}$, which exhibits the expected Yukawa suppression. Photon trajectories in the metric $g_{\mu\nu} = \eta_{\mu\nu} + \frac{1}{M_{\rm Pl}}h_{\mu\nu}$ will be warped, and the value one finds for the deflection angle, in the limit of small mass, is

$$\alpha = \frac{3GM}{b}, \tag{2.10}$$

where $G = (8\pi M_{\rm Pl}^2)^{-1}$, and b is the impact parameter. This should be compared to the corresponding expression in GR: $4GM/b$, and one sees that an arbitrarily small, yet non-zero, mass for the graviton yields deflection angles which vary by 25% from the case of an exactly zero graviton mass. One may wonder whether this discrepancy could be eliminated by simply redefining $G \to \frac{4}{3}G$, however, as shown below, it is not possible to do this without introducing further discrepancies. This circumstance is obviously unappealing theoretically, but is catastrophic phenomenologically, since the deflection angle (2.10) flies in the face of observations.

In order to understand how to remedy the situation, it is useful to examine it in a different way. Anticipating the discussion in section 3.1, the longitudinal polarisations of the massive graviton can be extracted by introducing new fields via

$$h_{\mu\nu} \to h_{\mu\nu} + \frac{1}{m}\left(\partial_\mu A_\nu + \partial_\nu A_\mu\right) + \frac{2}{m^2}\partial_{\mu\nu}\pi. \tag{2.11}$$

At energies far above the graviton mass A_μ will encode the dynamics of the vector modes of the massive graviton, and π the dynamics of the scalar mode. Assuming that the energy-momentum tensor is conserved, the Lagrangian (2.6) after introducing these new fields, and performing suitable partial integrations, is

$$\mathcal{L}_{m\neq 0} - \frac{1}{2}F_{\mu\nu}F^{\mu\nu} - \left(h_{\mu\nu}\left[m\partial^\mu A^\nu + \partial^{\mu\nu}\pi\right] - h^\mu_\mu\left[m\partial^\nu A_\nu + \partial^\nu_\nu\pi\right]\right) + \frac{1}{M_{\rm Pl}}h_{\mu\nu}T^{\mu\nu}. \quad (2.12)$$

The field redefinition $h_{\mu\nu} \to h_{\mu\nu} - \frac{1}{2}\pi\eta_{\mu\nu}$ then introduces a suitable kinetic term for π, and upon taking the limit $m \to 0$, the Lagrangian becomes

$$\mathcal{L}_{m=0} - \frac{1}{2}F_{\mu\nu}F^{\mu\nu} - \frac{3}{4}(\partial\pi)^2 + \frac{1}{M_{\rm Pl}}h_{\mu\nu}T^{\mu\nu} - \frac{1}{2M_{\rm Pl}}\pi T^\mu_\mu. \quad (2.13)$$

One sees that, whilst the vector modes have completely decoupled, the scalar mode has gained a coupling to matter which does not vanish in the massless limit. It is precisely this extra coupling of matter to gravity which means that the theory will not reproduce the results of GR, even in the massless limit. In particular, since this extra coupling is just to the trace of the energy-momentum tensor, whereas $h_{\mu\nu}$ couples to the full energy-momentum tensor, it is not possible to eliminate it by redefining G. For example, doing this to get the correct bending of light around a point mass, will mean that the response of the metric to traceless energy-momentum tensors does not agree with GR.

2.3.2 Vainshtein mechanism

A resolution to this problem is suggested however by (2.13): if the dynamics of π can be impeded in the vicinity of a source, then the extra coupling to matter effectively disappears. There is a large literature on possible ways of achieving this in general theories which contain extra, light degrees of freedom which modify gravity on cosmological scales—for example see [85] for a short review—and of these so-called screening mechanisms I will only focus on the one initially suggested by Vainshtein [31], not long after the vDVZ discontinuity was discovered, as this is the one which is active in massive gravity.

The basic idea is that at distances shorter than some scale, non-linearities in the behaviour of π become important, and so (2.13) on its own should not be trusted; furthermore, the effect of the non-linear terms is to screen the fifth force arising from the extra *d.o.f.* and prevent its influence from being felt on short distances.

Let us see how this works in an explicit example: consider adding to (2.13) the term $\frac{1}{\Lambda^3}(\partial\pi)^2\Box\pi$ (this is the cubic Galileon term mentioned in the introduction, and which appears in the massive gravity theories considered below). The equations of motion for π are modified to become (absorbing $\mathcal{O}(1)$ numbers into Λ and M)

$$\Box\pi + \frac{1}{\Lambda^3}\left((\Box\pi)^2 - (\partial_{\mu\nu}\pi)^2\right) = \frac{M}{M_{\rm Pl}}\delta^{(3)}(\mathbf{x}), \tag{2.14}$$

which can be integrated to yield

$$\frac{\pi'(r)}{r} + \frac{1}{\Lambda^3}\left(\frac{\pi'(r)}{r}\right)^2 = \frac{M}{M_{\rm Pl}}\frac{1}{4\pi r^3}. \tag{2.15}$$

At large distances the non-linear term is not important and one has the usual behaviour $\pi \sim r^{-1}$, whereas at short distances the non-linear term dominates, and one has $\pi \sim r^{1/2}$, and hence the fifth force is suppressed. The crossover scale between these two regimes is called the *Vainshtein radius* and is given by

$$r_V = \left(\frac{M}{M_{\rm Pl}\Lambda^3}\right)^{1/3} = \left(\frac{M}{m^2 M_{\rm Pl}^2}\right)^{1/3}, \tag{2.16}$$

where the second equality comes from setting $\Lambda = (m^2 M_{\rm Pl})^{1/3}$, its value in massive gravity. For a Hubble scale graviton mass and $M = M_\odot$, this becomes $\sim 10^7$ AU, and so within the solar system one is well within the screened regime, the vDVZ discontinuity is removed, and the theory respects solar system tests of GR.

2.3.3 Boulware-Deser ghost

The previous sub-section showed that a non-linear extension of the linear theory of massive spin-2 is required if the theory is to be phenomenologically viable, and naturally the standard arguments that a theory of massless spin-2 must be non-linearly completed if it is to consistently couple to energy-momentum in accordance with the equivalence principle are also in force here, and thus it is necessary to find a non-linear completion of (2.6).

This is easier said than done. A result of Boulware and Deser [6] showed that any such completion which could be expanded as a power series in the linear mass term would inevitably re-introduce the extra *d.o.f.* which is removed by the Fierz-Pauli tuning in (2.6); furthermore, they showed that, due to this extra *d.o.f.*, if the field equations have Minkowski space as a vacuum solution, then the Hamiltonian of the theory is not bounded from below—this instability is the (in)famous BD ghost.

Fortunately, a theory was discovered which evades this no-go theorem, and eliminates the BD ghost even at the non-linear level, thereby paving the way to a consistent theory of multi-gravity; this is described in detail in the next section.

2.4 dRGT massive gravity

The theory of massive gravity discovered by de Rham, Gabadadze, and Tolley [9, 10] consists of two elements: a metric tensor $g_{\mu\nu}$ and a non-dynamical, fixed, fiducial (or reference) metric $f_{\mu\nu}$ (in initial work this was taken to be the Minkowski metric $\eta_{\mu\nu}$, but it has since been generalised to arbitrary fiducial metric [15])—it is important to note that, in contrast to, for example, the bi-metric formulation of Eddington-inspired Born-Infeld gravity [86], the non-dynamical metric is not an auxiliary field which is varied in the action, but is completely fixed. The dynamical metric has a kinetic term given by its Ricci scalar (as in GR—the case of non-minimal kinetic terms will be considered in section 2.7), and a potential given by various contractions with the fiducial metric.[3] In total, in D dimensions the action looks like

$$\mathcal{S}_{\text{dRGT}} = \frac{M_{\text{Pl}}^{D-2}}{2} \int \mathrm{d}^D x \sqrt{-|g|} \left[R(g) + \frac{1}{2} m^2 \sum_{n=0}^{D} \beta_n e_n \left(\sqrt{g^{-1} f} \right) \right], \tag{2.17}$$

where β_n are constants and $e_n(M)$ is the n-th elementary symmetric polynomial of the eigenvalues of the matrix M. For example the first few are given by

$$e_0(M) = 1, \tag{2.18}$$

$$e_1(M) = \sum_i \lambda_i = [M], \tag{2.19}$$

$$e_2(M) = \sum_{j>i} \lambda_i \lambda_j = \frac{1}{2} \left([M]^2 - [M^2] \right), \tag{2.20}$$

[3]Any contractions of the metric with itself would be trivial, and so this is why the fiducial metric is required. There exist alternative approaches to massive gravity which do away with the need for a second metric, at the expense of being non-local, *e.g.* [87, 88], however as I am ultimately interested in theories with multiple, dynamical fields I will not discuss these here.

where $[M] \equiv \mathrm{tr}(X)$, and e_n vanishes for $n > D$; in general they are given by

$$e_n(M) = \frac{1}{n!}\delta^{\mu_1\cdots\mu_n}_{\nu_1\cdots\nu_n} M^{\nu_1}_{\mu_1}\cdots M^{\nu_n}_{\mu_n}, \tag{2.21}$$

and the generalised Kronecker delta is defined in equation (2.3). Note that $e_0 = 1$ just gives a cosmological constant, and $e_D\left(\sqrt{g^{-1}f}\right) = \left|\sqrt{g^{-1}f}\right|$ just adds a constant term to the action and so does not affect the dynamics, however it is kept here as upon moving to multi-gravity it cannot be ignored in this way.

When expanding around $g = f$, it is useful to re-express the potential as:

$$\sum_{n=0}^{D}\beta_n e_n\left(\sqrt{g^{-1}f}\right) = \sum_{n=0}^{D}\alpha_n e_n\left(\sqrt{g^{-1}f} - I_D\right), \tag{2.22}$$

where I_D is the $D \times D$ identity matrix, and the two sets of coefficients are related by

$$\alpha_n = \sum_{i=n}^{D} {}^{D-n}C_{D-i}\,\beta_i, \qquad \beta_n = \sum_{i=n}^{D} {}^{D-n}C_{D-i}\,(-1)^{i-n}\,\alpha_i. \tag{2.23}$$

Requiring $g = f$ to be a vacuum solution to the equations of motion means $\alpha_0 = \alpha_1 = 0$, which is equivalent to requiring the absence of a cosmological constant and tadpole term for the fluctuation $h = g - f$. These constraints, along with the fact that β_D has no effect in massive gravity, lead to a family of theories in four dimensions parametrised by α_2, α_3. As an example, these constraints on the parameters are satisfied for the so-called 'minimal model': $\mathcal{L}_{\mathrm{int}}(g,f) = \sqrt{-|g|}\left(3 - \mathrm{tr}\sqrt{g^{-1}f} + \left|\sqrt{g^{-1}f}\right|\right)$.

Another useful property of the potential is its behaviour under exchange of g and f:

$$\sqrt{-|g|}\,e_n\left(\sqrt{g^{-1}f}\right) = \sqrt{-|f|}\,e_{D-n}\left(\sqrt{f^{-1}g}\right), \tag{2.24}$$

which is simple to prove by noting that $\sqrt{f^{-1}g} = \sqrt{g^{-1}f}^{-1}$, and so

$$\begin{aligned}\sqrt{-|f|}\,e_{D-n}\left(\sqrt{f^{-1}g}\right) &= \sqrt{-|g|}\,\left|\sqrt{g^{-1}f}\right| \sum_{i_{D-n}>\cdots>i_1}\frac{1}{\lambda_{i_1}\cdots\lambda_{i_{D-n}}}\\ &= \sqrt{-|g|}\sum_{j_n>\cdots>j_1}\lambda_{j_1}\cdots\lambda_{j_n}\\ &= \sqrt{-|g|}\,e_n\left(\sqrt{g^{-1}f}\right). \end{aligned}\tag{2.25}$$

In order for (2.17) to make sense the matrix square root, defined by $\sqrt{g^{-1}f}^{\mu}{}_{\lambda}\sqrt{g^{-1}f}^{\lambda}{}_{\nu} = g^{\mu\lambda}f_{\lambda\nu}$, must exist; another factor to consider is the fact that it is not unique: in general there are 2^N possibilities for the square root of an $N \times N$ symmetric matrix,[4] which would lead to physically inequivalent theories. Most work so far has focused on the principal branch of solutions, however recently Comelli *et al.* [89] investigated other possible branches and found that whilst only one branch is perturbatively well defined around Minkowski space, if global Lorentz symmetry is spontaneously broken, then other branches become viable. Here I will focus only on those that preserve Lorentz symmetry.

2.4.1 Constraint analysis

As described in section 2.1, a spin-2 field which is massless or massive carries $\frac{1}{2}D(D-3)$ or $\frac{1}{2}(D+1)(D-2)$ *d.o.f.* respectively, whilst the metric tensor $g_{\mu\nu}$ is symmetric and so a priori it has $\frac{1}{2}D(D+1)$ *d.o.f.*, and so in order for it to describe a theory of massless spin-2 (GR) or massive spin-2 (massive gravity), $2D$ or $D+1$ *d.o.f.* respectively must be removed by constraints.

It is easiest to examine this in the Hamiltonian formalism via the procedure of Dirac, an exposition of which can be found in [90]. In a system with constraints, the action is

$$\int \mathrm{d}t \left(\dot{q}^i p_i - H(p_i, q^i) - u^m \phi_m(p_i, q^i)\right), \tag{2.26}$$

where ϕ_m are functions of the phase space variables (p_i, q^i), which are enforced to vanish by the Lagrange multipliers u^m. These are called *primary* constraints, and their vanishing describes a surface within phase space, on which the dynamics according to the equations of motion take place. For consistency, the primary constraints are required to be preserved in time

$$\dot{\phi}_m = \{\phi_m, H\} + u^n\{\phi_m, \phi_n\} \approx 0, \tag{2.27}$$

where $\{A, B\}$ is the usual Poisson bracket, and $\approx$ denotes equality which is only required to hold on the constraint surface. If this equation turns out to be independent of the Lagrange multipliers u, then it gives a further set of *secondary* constraints which are required to hold (on the constraint surface); otherwise this condition merely fixes the Lagrange multipliers. The preservation in time of the secondary constraints may then generate tertiary constraints, and so on. If the theory possesses any gauge invariances,

[4] This is most easily seen by diagonlising the matrix.

then, as well as showing up in the form of constraints, there will also be a further set of constraints one must impose, associated with fixing the gauge. The number of phase space *d.o.f.* remaining is then the initial number minus the total number of constraints.

The Hamiltonian formulation of GR was developed by Arnowitt, Deser, and Misner in the 1960's [91], and I will briefly explain it here along with how one ends up with the correct number of *d.o.f.* One has to introduce so-called ADM variables

$$N = \left(-g^{00}\right)^{-\frac{1}{2}}, \qquad N_i = g_{0i}, \qquad \gamma_{ij} = g_{ij}, \tag{2.28}$$

where i, j run over $1, \ldots, D-1$; N is called the *lapse*, and it relates the passage of time between two time-slices, whilst the *shift* vector N_i relates how the spatial coördinates change between two time-slices. The Einstein-Hilbert Lagrangian then becomes

$$\mathcal{L}_{\mathrm{EH}} = \pi^{ij}\partial_t\gamma_{ij} + NR^0 + R_iN^i, \tag{2.29}$$

where the canonical momentum conjugate to γ_{ij} is

$$\pi^{ij} = \sqrt{-|g|}\left(\Gamma^0_{pq} - \gamma_{pq}\gamma^{rs}\Gamma^0_{rs}\right)\gamma^{ip}\gamma^{jq}, \tag{2.30}$$

and the two constraints are

$$R^0 = \sqrt{|\gamma|}\,{}^{(D-1)}R + \frac{1}{\sqrt{|\gamma|}}\left(\frac{1}{2}\pi^i{}_i\pi^j{}_j - \pi^{ij}\pi_{ij}\right) \tag{2.31}$$

$$R_i = 2\sqrt{|\gamma|}\,\gamma_{ij}\nabla_k\left(\frac{\pi^{jk}}{\sqrt{|\gamma|}}\right), \tag{2.32}$$

where ${}^{(D-1)}R$ is the spatial Ricci scalar, and indices are raised and lowered by the spatial metric γ_{ij} (and its inverse γ^{ij}).

One can see from (2.29) that the lapse and shift are non-dynamical, and thus there are actually only $D(D-1)$ phase space variables, in (π^{ij}, γ_{ij}). This number is reduced by the D constraints enforced by the lapse and the shift, leaving $D(D-2)$, which is then reduced again by the D coördinate conditions one is free to impose. Thus in the end there are $\frac{1}{2}D(D-3)$ pairs of phase space *d.o.f.* remaining—precisely the right amount for a massless spin-2 field.

Now use the ADM decomposition (2.28) to analyse the dRGT action (2.17); one needs to perform a similar decomposition of the fiducial metric $f_{\mu\nu}$, for which the lapse and shift will be denoted L and L_i, and its purely spatial part by ϕ_{ij}.

The Einstein-Hilbert term will be the same, but there is now a contribution from the interaction potential. One has

$$g^{\mu\lambda}f_{\lambda\nu} = \frac{1}{N^2}\begin{pmatrix} L^2 + L_k(N^k - L^k) & -L_j + N^k\phi_{kj} \\ -N^i(L^2 + L_k(N^k - L^k)) + N^2\gamma^{ik}L_k & N^i(L_j - N^k\phi_{kj}) + N^2\gamma^{ik}\phi_{kj} \end{pmatrix}, \tag{2.33}$$

and thus the resulting action will be highly non-linear in both N and N^i (though they are still non-dynamical). This means that they will no longer manifestly enforce constraints and instead it seems that their *e.o.m.* can be used to determine them; one also does not have the diff symmetry which the massless case possessed, and so it seems that one is stuck with $\frac{1}{2}D(D-1)$ *d.o.f.*—one too many for a massive spin-2 field, and this apparent extra *d.o.f.* is precisely the BD ghost.

Fortunately the set of equations for N and N^i is degenerate, and thus there exists a (non-linear) redefinition of the shift which can solve this problem [10,15], by allowing the lapse to again act as a Lagrange multiplier, thus enforcing a primary constraint. Consider the following redefinition of N^i

$$N^i = (L\delta^i_j + ND^i{}_j)n^j + L^i, \tag{2.34}$$

for some matrix D and vector n, and write

$$N\sqrt{g^{-1}f} = A + NB, \tag{2.35}$$

for some matrices A and B. Squaring this and comparing powers of N one finds expressions for the squares and commutator of A and B, which can be solved to give

$$A = \frac{1}{\sqrt{x}}\begin{pmatrix} L + L_k n^k & n^k\phi_{kj} \\ -(L + L_k n^k)(L^i + Ln^i) & -(L^i + Ln^i)n^k\phi_{kj} \end{pmatrix} \tag{2.36}$$

$$B = \sqrt{x}\begin{pmatrix} 0 & 0 \\ D^i{}_k L^k & D^i{}_j \end{pmatrix}, \tag{2.37}$$

where $x = 1 - n^i n^j \phi_{ij}$, and the matrix D is determined by

$$\sqrt{x}D = \sqrt{(\gamma^{-1} - Dnn^{\mathrm{T}}D^{\mathrm{T}})\phi}. \tag{2.38}$$

Now note that for $i > 0$, one has $(A^i B^j)^2 = [A^i B^j] A^i B^j$, and thus as a result

$$e_n\left(\sqrt{g^{-1}f}\right) = e_n\left(\frac{1}{N}A + B\right) = \frac{1}{n!}\delta^{\mu_1\cdots\mu_n}_{\nu_1\cdots\nu_n}\left(\frac{n}{N}A^{\nu_1}{}_{\mu_1} + B^{\nu_1}{}_{\mu_1}\right)B^{\nu_2}{}_{\mu_2}\cdots B^{\nu_n}{}_{\mu_n}, \tag{2.39}$$

and hence in terms of the new shift vector n^i, the Lagrangian becomes

$$\mathcal{L}_{\text{dRGT}} = \mathcal{L}_{\text{EH}} + \frac{1}{2}m^2\sqrt{|\gamma|}\sum_{n=0}^{D}\beta_n\left[\frac{1}{(n-1)!}\delta^{\mu_1\cdots\mu_n}_{\nu_1\cdots\nu_n}A^{\nu_1}{}_{\mu_1}B^{\nu_2}{}_{\mu_2}\cdots B^{\nu_n}{}_{\mu_n} + Ne_n(B)\right]. \tag{2.40}$$

Whilst this looks to be linear in the lapse, care must be taken, since the non-propagating field n^i must first be eliminated using its *e.o.m.* and, as B has implicit dependence on n^i through (2.38), it is possible for this procedure to yield higher order dependence on N.[5] Fortunately this does not happen here, and N really does enforce a constraint eliminating one phase space *d.o.f.* It is also interesting to note that one can act in reverse, and use the requirement that the action be linear in the lapse to derive the dRGT interaction terms (and demonstrate their uniqueness) [92].

As described above, one then requires this constraint to be preserved by time evolution. Note that, although there is only one constraint here, and from (2.27) it seems this should trivially yield a secondary constraint, the Poisson bracket which must vanish involves the constraint evaluated at two different points, and so one must check that this vanishes. However it turns out that it does, and thus there exists a secondary constraint [18]. Therefore the troublesome extra *d.o.f.* is completely removed, and one is left with the $\frac{1}{2}(D+1)(D-2)$ *d.o.f.* of a massive spin-2 field.

This secondary constraint is also required to be preserved by the time evolution, and in principle this could yield tertiary, quaternary, *etc.* constraints, eliminating yet more *d.o.f.* (and thus overshooting the mark). This does not happen [18], which can be understood since such an occurrence would require one of the Poisson brackets in (2.27) to vanish, yet they are already non-zero at lowest order, since there is no extra constraint in the linearised theory of massive spin-2.

[5] As a simple example of how this can happen, consider the following Lagrangian

$$p\dot{q} - \alpha^2 + \lambda(2\alpha + p); \tag{2.41}$$

it looks as if λ enforces a constraint, but using the *e.o.m.* of α to solve for it and substitute back in gives

$$p\dot{q} + \lambda^2 + \lambda p, \tag{2.42}$$

in which λ no longer appears linearly.

2.4.2 Vielbein formulation

The dRGT theory of massive gravity can also be formulated in terms of vielbeine, rather than metrics, in which form the potential looks markedly less unusual. Here I briefly recapitulate how this is done, and demonstrate its equivalence to the metric version. [19, 93–96]

The Einstein-Hilbert action becomes

$$\int \mathrm{d}^D x \sqrt{-|g|}\, R(g) \quad \to \quad \int \epsilon_{a_1 \cdots a_{D-2} bc}\, E^{a_1} \wedge \cdots \wedge E^{a_{D-2}} \wedge R^{bc}(E), \tag{2.43}$$

where E^a is a one-form vielbein and $R^{ab}(E)$ is the associated gauge curvature two-form, whilst the interaction terms become

$$\int \mathrm{d}^D x \sqrt{-|g|}\, e_n \left(\sqrt{g^{-1} f}\right) \quad \to \quad \int \epsilon_{a_1 \ldots a_{D-n} b_1 \ldots b_n} E^{a_1} \wedge \cdots \wedge E^{a_{D-n}} \wedge F^{b_1} \wedge \cdots \wedge F^{b_n}, \tag{2.44}$$

where F^a is a second one-form vielbein (distinct from E^a, and fixed). That is, the set of possible interaction terms is just given by the set of all the different combinations of E's and F's wedged together.

The Einstein-Hilbert term now manifestly respects both diffeomorphism invariance, $E^a_\mu(x) \to E^a_\nu(f(x))\partial_\mu f^\nu$, and also local Lorentz invariance, $E^a \to \Lambda^a{}_b E^b$, which, like the diff invariance, is broken by the interaction terms. Anticipating the Stückelberg trick, which will be explained in the next chapter, this can be (re)introduced to the whole theory by introducing a new field, $\Lambda^a{}_b$, via $F^a \to \Lambda^a{}_b F^b$, to make F^a transform in the same way as E^a under local Lorentz transformations.

It is clear that Λ will be an auxiliary field and, following [95], its *e.o.m.* yields

$$E^a_{[\mu|}\eta_{ab}(\Lambda F)^b_{|\nu]} = 0, \tag{2.45}$$

for any combination of interaction terms. In unitary gauge ($\Lambda^a{}_b = \delta^a_b$) this becomes the famous Deser-van-Niewenhuizen (DvN) symmetric vielbein condition. I will now show how this condition is sufficient to show the equivalence with the metric version of bi-gravity (in four dimensions it is also necessary [97]). In matrix notation (2.45) reads

$E^{\mathrm{T}}\eta(\Lambda F) = (\Lambda F)^{\mathrm{T}}\eta E$ from which one gets $(\Lambda F)E^{-1} = \eta^{-1}(E^{\mathrm{T}})^{-1}(\Lambda F)^{\mathrm{T}}\eta$, and thus

$$\begin{aligned}(E^{-1}\Lambda F)(E^{-1}\Lambda F) &= E^{-1}\eta^{-1}(E^{-1})^{\mathrm{T}}(\Lambda F)^{\mathrm{T}}\eta(\Lambda F) \\ &= (E^{\mathrm{T}}\eta E)^{-1}(F^{\mathrm{T}}\eta F) \\ &= g^{-1}f, \end{aligned} \tag{2.46}$$

where the metric $g_{\mu\nu} = E^a_\mu \eta_{ab} E^b_\nu$, and similarly for f and F. Therefore $E^{-1}(\Lambda F) = \sqrt{g^{-1}f}$ (modulo non-uniqueness of the square root), and hence

$$\begin{aligned}&\int \epsilon_{a_1\ldots a_{D-n}b_1\ldots b_n} E^{a_1}\wedge\cdots\wedge E^{a_{D-n}}\wedge(\Lambda F)^{b_1}\wedge\cdots\wedge(\Lambda F)^{b_n} \\ &\quad= \int \mathrm{d}^D x\, \epsilon_{a_1\ldots a_{D-n}b_1\ldots b_n}\epsilon^{\mu_1\ldots\mu_{D-n}\nu_1\ldots\nu_n} E^{a_1}_{\mu_1}\ldots E^{a_{D-n}}_{\mu_{D-n}}(\Lambda F)^{b_1}_{\nu_1}\ldots(\Lambda F)^{b_n}_{\nu_n} \\ &\quad= \int \mathrm{d}^D x\, |E|\, \epsilon_{\mu_1\ldots\mu_{D-n}\rho_1\ldots\rho_n}\epsilon^{\mu_1\ldots\mu_{D-n}\nu_1\ldots\nu_n}(E^{-1})^{\rho_1}_{b_1}(\Lambda F)^{b_1}_{\nu_1}\ldots(E^{-1})^{\rho_n}_{b_n}(\Lambda F)^{b_n}_{\nu_n} \\ &\quad= \int \mathrm{d}^D x\, \sqrt{-|g|}\,(D-n)!\,\delta^{\nu_1\ldots\nu_n}_{\rho_1\ldots\rho_n}\sqrt{g^{-1}f}^{\rho_1}{}_{\nu_1}\ldots\sqrt{g^{-1}f}^{\rho_n}{}_{\nu_n} \\ &\quad= n!(D-n)!\int \mathrm{d}^D x\,\sqrt{-|g|}\, e_n\left(\sqrt{g^{-1}f}\right). \end{aligned} \tag{2.47}$$

Further formulations exist, for example the chiral Plebanski formulation [98], though I shall not consider that in this thesis.

2.5 Bi- and multi-gravity

2.5.1 Bi-gravity

dRGT massive gravity can be extended to become HR bi-gravity [17,18] by declaring the fiducial metric f to be a dynamical field and adding a kinetic term for it to the dRGT action:

$$\begin{aligned}\mathcal{S}_{\mathrm{HR}} = \int \mathrm{d}^D x \Bigg[&\frac{M_g^{D-2}}{2}\sqrt{-|g|}\,R(g) + \frac{M_f^{D-2}}{2}\sqrt{-|f|}\,R(f) \\ &+ \frac{1}{2}m^2 M_{\mathrm{eff}}^{D-2}\sum_{n=0}^{D}\beta_n\sqrt{-|g|}\,e_n\left(\sqrt{g^{-1}f}\right)\Bigg], \end{aligned} \tag{2.48}$$

where $M_{\mathrm{eff}}^{-(D-2)} = M_g^{-(D-2)} + M_f^{-(D-2)}$; it might seem that the interaction term treats the two fields asymmetrically, however note (2.24). Massive gravity is recovered upon taking the limit $M_f \to \infty$, which freezes the f metric.

Considering just the sum of the two kinetic terms, there is a gauge symmetry $\text{diff}_g \times \text{diff}_f$, where diff_g acts only on g and leaves f invariant, and vice versa for diff_f; the interaction term breaks this down to just one copy of diff invariance, the diagonal subgroup in which each symmetry acts in the same way. This remaining symmetry means that of the two fields in the theory, one combination will remain massless, whilst the other gains a mass.[6] The constraint analysis for (2.48) proceeds in much the same way as in section 2.4.1, and so I shall not reproduce it here. The vielbein formulation is also arrived at in the same way as for massive gravity, and (2.45) again ensures the equivalence to the metric formulation.

2.5.2 Multi-gravity

This can be further extended to a theory of multi-gravity by introducing more dynamical metrics, giving them each a Ricci-scalar kinetic term, and adding to the action interaction terms of the same form as in (2.17) and (2.48).

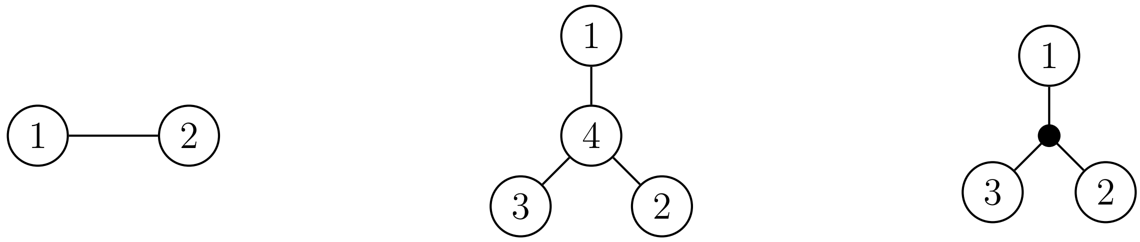
With more than two fields there also comes a qualitatively new possibility: interaction terms which directly involve more than two different fields. The possibilities are easiest to see in the vielbein version of the theory, and it turns out that wedging together any D vielbeine still yields a healthy interaction term [19], *i.e.*

$$\int \epsilon_{a_1 \cdots a_D} \, E^{a_1}_{(i_1)} \wedge \cdots \wedge E^{a_D}_{(i_D)}, \tag{2.49}$$

where i_1, i_2, *etc.* (which label the fields) are arbitrary.

The DvN symmetric vielbein condition (2.45) is now no longer automatically enforced by the *e.o.m.* of the local Lorentz invariance Stückelberg fields, however nonetheless such terms can be translated into the metric language [93]. One must enforce the DvN condition by hand for each pair $(E_{(i_l)}, E_{(j)})$ for a given $E_{(j)}$ (by choosing an appropriate Lorentz frame rotation of each $E_{(i_l)}$) and then one again has $E^{-1}_{(j)} E_{(i_l)} = \sqrt{g^{-1}_{(j)} g_{(i_l)}}$. This allows one to proceed in the same way as (2.47), except that to be truly equivalent to the theory arising from (2.49), care must be taken to ensure that the same constraints

[6]Writing the fluctuations of the two fields about Minkowski space as $g = \eta + M_g^{\frac{2-D}{2}} h$, $f = \eta + M_f^{\frac{2-D}{2}} l$, then the massless mode is $\frac{1}{\sqrt{2}}(h+l)$ and the massive one is $\frac{1}{\sqrt{2}}(h-l)$. This will be explained in more detail in section 6.2.



$$\mathcal{L}_{\text{int}} = f(g_{(1)}, g_{(2)}), \qquad f_1(g_{(1)}, g_{(4)}) + f_2(g_{(2)}, g_{(4)}) + f_3(g_{(3)}, g_{(4)}), \qquad f(g_{(1)}, g_{(2)}, g_{(3)})$$

Figure 2.1: Examples of different types of theory graphs and their corresponding interaction Lagrangians; white vertices correspond to metrics/vielbeine, and when two are present in the same interaction term the corresponding vertices are connected by an edge; black vertices indicate that the fields corresponding to the connected white vertices appear in the same, single interaction term.

are enforced; this can be done by introducing a set of auxiliary fields $L(i_l)$, and one is left with

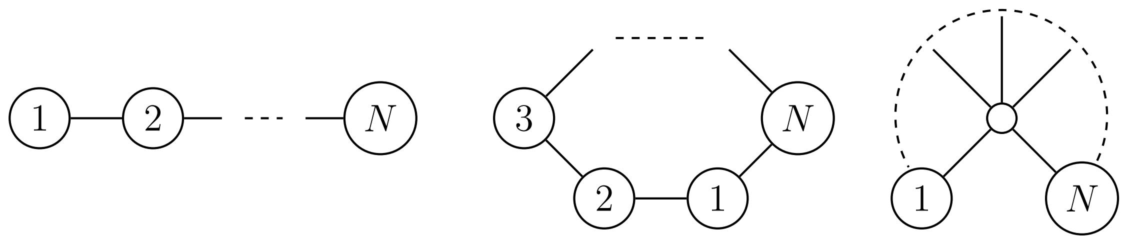
$$\int \mathrm{d}^D x \sqrt{-\left|g_{(j)}\right|}\, \delta^{\mu_1 \cdots \mu_D}_{\nu_1 \cdots \nu_D} L(i_1)^{\nu_1}{}_{\lambda_1} \sqrt{g^{-1}_{(j)} g_{(i_1)}}^{\lambda_1}{}_{\mu_1} \cdots L(i_D)^{\nu_D}{}_{\lambda_D} \sqrt{g^{-1}_{(j)} g_{(i_D)}}^{\lambda_D}{}_{\mu_D}. \tag{2.50}$$

Before considering the *d.o.f.* which a multi-gravity theory possesses, as well as equivalence between the metric and vielbein versions of it, it is appropriate to introduce a neat graphical way of describing the structure of a given theory, which will greatly simplify the discussion.

2.5.3 Theory graphs

The structure of a multi-gravity theory can be conveniently encoded in a so-called theory graph [19, 47, 77]. Each metric/vielbein in the theory is represented by a vertex, and where two fields appear in an interaction term in the action, the corresponding vertices in the theory graph are connected. If more than two fields appear in a single interaction term then all the relevant fields are connected to an extra, auxiliary vertex. Examples are given in figure 2.1, and some common graphs to which it will be useful to refer later are given in figure 2.2. Also note, the terminology used to refer to the elements of these graphs will occasionally differ from the above: vertices will sometimes be called 'sites,' and edges 'links.'

This obviously does not encode all of the information about a theory, since neither the



Path graph, P_N Cycle graph, C_N Star graph, S_N

Figure 2.2: Some common graphs to which it will be useful to refer later.[7]

specific form of the interaction terms nor their coefficients are given,[8] however it turns out that these graphs do nonetheless control many aspects of the theory, and certain general properties may be derived from them. Also, one may wonder how to deal with the cosmological constant terms, which in the metric version are $\sqrt{-|g|}\, e_0\left(\sqrt{g^{-1}f}\right)$ and $\sqrt{-|g|}\, e_D\left(\sqrt{g^{-1}f}\right)$, as these actually only involve one field. One could treat these by allowing vertices to be connected to themselves, however as they do not break diff symmetries on their own it is acceptable to ignore them when constructing the theory graph.

First let us consider the equivalence of the metric and vielbein version of a given multi-gravity theory. In the absence of cycles in the theory graph (*e.g.* there are interaction terms linking A with B, B with C, and C with A again) this equivalence will continue to hold, since each Lorentz Stückelberg field is independent and hence each pair of vielbeine joined by an interaction term will individually obey the DvN condition (2.45). In the presence of cycles however, as will be explained in section 5.1, one cannot blithely introduce Stückelberg fields in the same way because there will still be one unbroken copy of gauge invariance, and as a result the DvN condition for each pair of connected vielbeine no longer holds. In contrast to the N-vielbein interaction term discussed above, for which it was sufficient to enforce this for $N-1$ pairs of vielbeine, and which can always be done by rotating the frames of $N-1$ vielbeine relative to a final one, in the case of an N-vielbein cycle graph theory, this would have to be enforced for N pairs, which cannot

[7] The star graph is sometimes denoted by $K_{1,N-1}$, but the notation S_N is used here to avoid confusion with the kinetic matrix K, which will be important later.

[8] To include this one could generalise to a weighted graph in which, for bi-gravity type interactions, each edge carries a vector whose n-th component is equal to the coefficient of, in the metric version the $\sqrt{-|g|}\, e_n\left(\sqrt{g^{-1}f}\right)$ interaction term, and in the vielbein version the $E^{a_1}\wedge\cdots\wedge E^{a_{D-n}}\wedge F^{b_1}\wedge\cdots\wedge F^{b_n}$ interaction term, with an obvious generalisation for interaction terms involving more than two fields.

be done.[9] Thus in the presence of cycles in the theory graph, the equivalence between the metric and vielbein versions breaks down.

Now let us consider the *d.o.f.* possessed by a given theory. First consider the gauge symmetries: each connected component of the theory graph will retain one unbroken copy of diff, and so one would expect a theory graph with N fields and n connected components to contain n massless and $N-n$ massive gravitons. As each connected component of the theory graph can be treated independently, without loss of generality from here on I will consider only graphs with a single connected component.

Secondly, let us consider the constraint analysis of the theory. For each bi-metric interaction (*i.e.* each edge of the graph), one must use (2.34) to redefine the shift vector of one of the metrics involved, and to be consistent, one must apply the redefinition throughout the action. Now note that (2.34) is linear in the lapses of both of the metrics involved, and linear in the shift of the other metric, and thus when it is applied to an arbitrary term which is linear in the shift which is being redefined, the resulting term will be linear in both the lapses which are involved in the redefinition and linear in the other shift which is involved.

Consider now the theory corresponding to the star graph S_d for $d > 1$ (see figure 2.2), and call the central vertex V; for any of the outer vertices (which are of degree one) it is possible to perform the shift redefinition corresponding to the edge connected to V without causing any problem; do this for $d-1$ of the vertices adjacent to V, and the action will be linear in the lapses of those $d-1$ vertices, and linear in the lapse and shift of V; therefore, one can perform for V the shift redefinition corresponding to the edge connecting it to the remaining vertex, W, and the resulting action will be linear in all of the lapses, and linear in the shift of W.

Now imagine that W has further edges connected to it, and denote its degree d_W; if $d_W = 2$, then the above story applies and one can freely perform the shift redefinition for W corresponding to this new edge; however if $d_W > 2$, one must wait until the shifts of a further $d_W - 2$ of it adjacent vertices have been redefined (recall that the shift of V has already been redefined) before then redefining the shift of W.

[9] *e.g.* consider a tri-vielbein cycle: rotate the frames of (2) and (3) so that DvN is enforced for the pairs $(1,2)$ and $(1,3)$, but now there is no freedom to ensure that DvN is satisfied for the pair $(2,3)$.

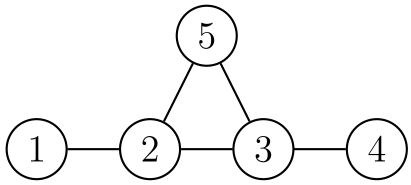


Figure 2.3: The 'shift redefinition cascade' breaks down in the presence of cycles.

The shifts of (1) and (4) can immediately be safely redefined (*i.e.* doing so leaves one with something linear in their lapses) corresponding to the edges $(1, 2)$ and $(3, 4)$; in order for the redefinition of the shift of (2) corresponding to the edge $(2, 5)$ (resp. $(2, 3)$) to be safe, the terms arising from the edge $(2, 3)$ (resp. $(2, 5)$) must be linear in the shift of (2), *i.e.* the shift of (3) (resp. (5)) must already have been redefined according to the edge $(2, 3)$ (resp. $(2, 5)$); however the equivalent is true for the vertices (3) and (5) (*i.e.* their shifts cannot safely be redefined (corresponding to either of their edges) until the shift of either (2) or (5) in the case of (3) and either (2) or (3) in the case of (5) have been redefined) and thus the redefinition cascade hits a wall—the action is only linear in the lapses of (1) and (4), and none of the shifts.

If the edge $(3, 5)$ were not present, then the shift of (5) could immediately be safely redefined (corresponding to the edge $(2, 5)$), and then one can safely either redefine (2) or (3) corresponding to the edge $(2, 3)$, and one would have an action linear in all the lapses, and linear in the shift of either (2) or (3).

Therefore one sees that it is possible to move inwards from the vertices of degree one (which form the boundary of the graph), redefining shifts as one goes (and pausing at vertices of degree greater than two as described above, if need be). The resulting action will then be linear in all the lapses, and linear in one of the shifts and thus the above reasoning about the number of *d.o.f.* is correct [19, 92].

In the presence of cycles however, this shift redefinition cascade breaks down—either the graph has no boundary from which to initiate it, or one (or more) of the redefinition 'streams' will get stuck at the entrance to a cycle subgraph, as depicted in figure 2.3. Therefore the standard constraint analysis breaks down, for the metric version of the theory, in the presence of a cycle [92, 99], and there could be additional *d.o.f.* in this case (and in fact this is what I will show in section 5.2).

2.6 Coupling to matter

Thus far the theories described have only contained gravitons, however the real world contains other fields (here generically referred to as 'matter') to which the multi-gravity needs somehow to couple, if it is to be involved in a putative description of reality. Furthermore the question of how these theories can couple to matter whilst still retaining their stability properties is interesting from a purely theoretical standpoint as one considers the space of healthy field theories. Despite this I will not examine the question in this thesis, and so instead here briefly state the known results.

One possiblity is to couple to just one metric/vielbein; this does not disrupt the constraint analysis and the theory remains ghost-free [11,15]. However one may consider full ghost-freedom to be too strong a constraint, and instead merely require that any ghost introduced by the matter coupling lies at an energy scale above the strong-coupling scale of the multi-gravity theory, as then it does not impact the low energy effective theory.

When the symmetric vielbein condition holds, the couplings which obey this requirement correspond to coupling a given matter field to an effective vielbein which is a general linear combination of those in the theory [100], *i.e.* coupling to

$$E^a_{\text{eff}} = \sum_i \alpha_i E^a_{(i)}, \tag{2.51}$$

which in terms of an effective metric is [101]

$$g^{\text{eff}}_{\mu\nu} \equiv E^a_{\text{eff}} \eta_{ab} E^b_{\text{eff}} = \sum_{i,j} \alpha_i \alpha_j \, g^{(i)}_{\mu\lambda} \sqrt{g^{-1}_{(i)} g_{(j)}}^{\lambda}{}_{\nu}. \tag{2.52}$$

When this does not hold, the most general known effective vielbein to which one may couple becomes a superposition of terms involving products of vielbeine and their inverses subject to the condition that the number of times a given field appears is equal to the number of times its inverse appears, apart from one field for which the number of inverses is one less than the number of times it appears [102]. For example, allowed terms include $E^b_{(i)\mu} \eta_{bc} E^c_{(j)\lambda} E^{-1}_{(i)d}{}^{\lambda} \eta^{ad}$.

Furthermore different matter fields may couple to different effective vielbeine/metrics (however this naturally breaks the weak equivalence principle, since not all matter will then gravitate in the same way, and so when taking this approach in the real world one must be mindful of experimental constraints). On the other hand coupling the same sector of matter separately to different metrics/vielbeine generically re-introduces the BD ghost and lowers the cutoff of the theory [101,103].

2.7 Kinetic terms

The kinetic terms for the metrics/vielbeine in the actions considered so far, (2.17) and (2.48) (and the extension to multi-gravity), have all consisted of ordinary Einstein-Hilbert terms for each field, and in particular they do no mix the different fields —this only occurs in the potential terms. The possibility of kinetic terms which mix different spin-2 fields in a consistent manner has obvious theoretical interest, especially as this would then allow the construction of Yang-Mills-like theories for spin-2, with a multiplet of fields transforming under some group.

Initial work towards constructing non-Einstein-Hilbert kinetic terms for a massive graviton was completed to cubic order in [104], and to pseudo-linear order (by which is meant, the leading order in an expansion about flat space) in [105], in which possible terms with $2d$ derivatives and n fields (with the requirement $d \leq n \leq D - d$) were found. Unfortunately it was then shown that it is not possible to non-linearly complete such terms without reintroducing the BD ghost below the cutoff scale of the effective field theory [106, 107].

The general matter couplings discussed in the previous section, although not entirely ghost free (however they are at least ghost free up to the cutoff of the effective field theory), suggest another way to investigate non-minimal kinetic terms: perform a field redefinition to make the matter coupling minimal, which will then make non-minimal the kinetic term of the relevant vielbeine, *i.e.* much the same as going from the 'Einstein' to the 'Jordan' frame in GR. Note that the equivalence of the metric and vielbein versions breaks down in such a case, since the modification of the kinetic term (such that it no longer respects the usual symmetries) will modify the *e.o.m.* for the Lorentz Stückelberg field, meaning that the DvN symmetric vielbein condition no longer holds. Thus these would not necessarily be covered by the previous no-go results, which worked in the metric picture.

This was investigated in [108], and new candidate kinetic interactions inspired by such field redefinitions were studied, though indeed they contain a ghost at some scale. Finally a no-go result in the language of vielbeine was derived in [109]—interestingly, there it is the Lorentz Stückelberg fields that cause the trouble, by becoming dynamical, as opposed to the return of the usual BD ghost (associated with the diff Stückelberg fields).[10] Thus it appears that there are no consistent deformations of the kinetic term in dRGT massive gravity, and the associated bi- and multi-gravity theories.

2.8 Quantum corrections

Given the intricate structure of the potential in (2.17) which is required to remove the BD ghost, one may wonder whether this is preserved by quantum corrections, since if it is not, then the BD ghost will reappear—and may lower the cutoff—ultimately rendering the dRGT theory no better than any other theory of massive gravity.

This question was studied in the so-called decoupling limit (which will be introduced and explained in the next chapter) in [110], and it was found that the interactions present in that limit do not get renormalised.[11] It has been studied beyond the decoupling limit in [111–113] for massive gravity and [114] for bi-gravity.

As virtual particles in the loops when calculating quantum corrections to graviton scattering amplitudes, one must consider both matter fields and gravitons themselves. When matter is minimally coupled to a single site, the effect from corrections exclusively involving matter loops is particularly simple, because it is the graviton propagators which are modified by adding a mass, but these propagators will not enter into a calculation just involving matter loops. Therefore the corrections will be the same as in GR, and only the cosmological constant receives corrections [115]. This was confirmed by explicit calculation to one-loop level in [111, 113]. The requirement that matter loops do not detune the structure of the potential is what was used to discover the more complicated matter couplings presented in section 2.6, and so the conclusion remains the same when matter is coupled in that way [100, 101, 114].

On the other hand graviton loops are potentially problematic. For massive gravity these were studied in [111–113], and I will focus in more detail on the calculation presented in [113]. Working in $D = 4$, using a dimensional regularisation renormalisation scheme, and working to one loop, it was found that at lowest order in the coupling constants the

[10] On the other hand, given that it was necessary to avoid equivalence with the metric version, in order to avoid the previous no-go results, this is perhaps unsurprising.

[11] This is related to the non-renormalisation theorems which exist for Galileons [35, 60, 61]—a special type of scalar field theory which will be discussed in section 3.3.

dRGT potential terms themselves do not spoil the structure, but that at higher order they do, as do higher order terms coming from an expansion of the Einstein-Hilbert term. Schematically, terms are found coming in as $\frac{m^4}{M_{\rm Pl}^n}h^n$, and if one then perturbs around some background $h = \bar{h} + \delta h$, one then gets a ghost with mass $\left(\frac{M_{\rm Pl}}{\bar{h}}\right)\bar{h}$, which is problematic since by taking $\bar{h} > M_{\rm Pl}$, as would be the case when the Vainshtein effect is operational,[12] the mass of the ghost can be made arbitrarily small. This neglects however the 'redressing' of terms in the action due to the Vainshtein effect, and a more careful analysis, utilising the one-loop effective action, then determines that the mass of the ghost is $\gtrsim M_{\rm Pl}$, *i.e.* beyond the range of validity of the theory, and so one does not need to worry about it. In [114] this analysis was completed for bi-gravity, and the same results were found.

Thus to one loop order the dRGT interaction terms are not detuned by quantum corrections in a dangerous way, and it would seem that even at the quantum level the dRGT theory (and its extension to multi-gravity) is healthy, at least up to its strong-coupling scale. However at higher loop order, and note that then there will be diagrams simultaneously involving matter and graviton loops, the status of the theory is currently unresolved.

2.8.1 Technical naturalness

In the above, we saw that (at one loop) the quantum corrections to the graviton potential looked like $\frac{m^4}{M_{\rm Pl}^n}h^n$, and so the mass of the graviton will get renormalised like

$$\delta m^2 \sim m^2 \left(\frac{m}{M_{\rm Pl}}\right)^2, \tag{2.53}$$

which is proportional to the mass itself. A quantity with this property is called *technically natural*, since if one sets a certain value for it, this value will not be wildly altered by quantum corrections (as opposed to, say, the mass of a scalar field, such as the Higgs).

Such behaviour is expected here due to arguments by 't Hooft and others [116, 117], which state that if the theory gains a symmetry when a certain parameter is taken to zero, then that parameter is technically natural. The canonical example of this is the electron mass, which is much smaller than the weak scale, but this is not problematic

[12]Note that this can be done whilst staying within the regime of validity of the effective theory so long as one requires $\partial\bar{h}$ to be small.

since as one sets m_e to zero the theory gains a chiral symmetry. In this case, as one sets the graviton mass(es) to zero, the theory gains diffeomorphism symmetry.

This property is useful if one wants to use massive gravity to explain the current cosmological acceleration, as one has to set $m \sim H_0$, and if the graviton mass were then wildly affected by quantum corrections, one would not have improved upon the cosmological constant problem.

2.9 Partial masslessness

Although I do not consider them in this thesis, it is worth very briefly mentioning theories of partially massless (PM) spin-2 fields. These are fields which carry a number of *d.o.f.* in between those of a massless and a massive field (of spin corresponding to its highest helicity state). In the case of spin-2 this refers to a field which carries helicity ± 2 and helicity ± 1 states, but no helicity 0 state. Such a theory exists at the linear level, describing a PM spin-2 field propagating on de Sitter space, provided the mass is related to the Hubble parameter as $m^2 = (D-2)H^2$; the removal of the helicity 0 mode occurs because of the appearance of a new gauge symmetry [118–121].

If a consistent non-linear extension of this theory existed, aside from being theoretically interesting, it would have several incredibly useful features.

- As seen in section 2.3.1, it is precisely the helicity 0 mode which causes vDVZ discontinuity in ordinary massive gravity, and so in a partially massless theory there would be no need for the Vainshtein mechanism (which, as explained in the introduction, is in some sense the cause of the low strong-coupling scale in massive gravity).

- The gauge symmetry would protect the structure of the interactions from quantum corrections, whereas in dRGT theory there is no such symmetry (though it turns out that, at least to one loop level, the structure of the dRGT potential is not perturbed—see section 2.8).

- Perhaps most pleasingly, the value of the cosmological constant is now no longer arbitrary, but is tied to the mass of the graviton.[13] Meanwhile, the mass of the graviton is a technically natural parameter, only receiving corrections which are proportional to it (see section 2.8.1).

Thus a non-linear PM gravity theory could elegantly solve the cosmological constant problem. Unfortunately it seems that such a non-linear extension does not exist [122–124]. Despite that, one might hope that it is still possible for a PM bi- or multi-gravity theory to exist [125–127], and that the theory with a single PM graviton does not exist merely because something goes wrong when taking the massive gravity limit, as discussed in [128, 129], and see the latter paper for recent work attempting to find such a theory. On the other hand, see [130] for a theorem ruling out the possibility of a Yang-Mills-like multiplet of PM spin-2 fields.

[13] And this relation is protected from quantum corrections by the gauge symmetry.

CHAPTER

3

Interacting Spin-2 Fields in The Stückelberg Picture

In this chapter I introduce the Stückelberg trick, which is a way of bestowing gauge invariance upon a theory and will be crucial in the analysis presented in the remainder of this thesis; it is discussed how to apply it to bi-gravity in section 3.2, which leads naturally to discussing Galileons in section 3.3; finally in section 3.4 the Stückelberg trick is applied to multi-gravity, and it is explained how this differs from the case of massive and bi-gravity. This chapter is largely based on [77].

3.1 The Stückelberg trick

The essence of the Stückelberg trick [131] is that gauge symmetry is a sham! It is not a 'real' symmetry but merely a redundancy of description which means that the true number of *d.o.f.* in a theory is less than the naïve number due to the fields appearing in the theory. The Stückelberg trick is a way of exploiting this to introduce a given gauge symmetry into a theory which does not possess it, at the expense of introducing extra fields. Before using it to study multi-gravity, I will quickly give a simple example to demonstrate how it works.

Consider a non-Abelian gauge field $A_\mu = A^i_\mu t^i$, where t^i, $i = 1, \ldots, N$, are the generators of a group G, in four dimensional Minkowski space, and consider adding a mass term to its Lagrangian, which becomes:

$$-\frac{1}{4}F^{i\mu\nu}F^i_{\mu\nu} - \frac{1}{2}m^2 A^{i\mu}A^i_\mu, \tag{3.1}$$

where the field strength is $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu + ig[A^\mu, A^\nu]$. When $m = 0$ this Lagrangian is invariant under $A^\mu \to U^{-1}\left(A_\mu + ig^{-1}\partial_\mu\right)U$, with $U \in G$; this gauge invariance eliminates $2N$ of the apparent $4N$ *d.o.f.* in A_μ.

When $m \neq 0$ the gauge invariance of the Lagrangian is lost, and it describes N massive vector fields (for $3N$ *d.o.f.* in total); the gauge invariance can be (re)introduced however by *introducing a new field* $U = \exp(i\frac{g}{m}\phi^i t^i)$ in a manner patterned after the gauge transformation, and the new Lagrangian is

$$-\frac{1}{4}F^{i\mu\nu}F^i_{\mu\nu} - \frac{1}{2}m^2 A^{i\mu}A^i_\mu - \frac{1}{2}\partial^\mu\phi^i\partial_\mu\phi^i + mA^i_\mu\partial^\mu\phi^i - \frac{1}{2}gf^{ijk}A^i_\mu\phi^j\partial^\mu\phi^k + \cdots, \tag{3.2}$$

where f^{ijk} are the structure constants of the Lie algebra of G, and the ellipsis represents terms higher order in ϕ. The whole Lagrangian is now invariant under the following gauge transformation:

$$\begin{aligned} A^i_\mu &\to A^i_\mu - \frac{1}{g}\partial_\mu\alpha^i + f^{ijk}A^j_\mu\alpha^k + \cdots \\ \phi^i &\to \phi^i - \frac{m}{g}\alpha^i + \frac{m}{2g}f^{ijk}\phi^j\alpha^k + \cdots, \end{aligned} \tag{3.3}$$

which means that of the $5N$ *d.o.f.* naïvely present in A^μ and ϕ, only $3N$ propagate.

Now one may ask, why is this useful? After all, adding extra fields would seem to make things more complicated and less transparent.

3.1.1 The Goldstone boson equivalence theorem

The answer lies in the Goldstone boson equivalence theorem [132–135], which states that in a spontaneously broken gauge theory, at energies far greater than the mass of the gauge boson, in matrix elements, external longitudinally polarised gauge bosons can be replaced by the Goldstone boson (which the gauge boson 'eats' to become massive), up to corrections of order m/E.

Now here we are dealing not with a spontaneously broken, but explicitly broken gauge theory, however using the Stückelberg trick the same logic applies [47]—for example, we can think of (3.2) as being a low energy approximation to a conventional spontaneously broken gauge theory

$$-\frac{1}{2}\mathrm{tr}\left[(F_{\mu\nu})^2\right] - \mathrm{tr}\left[|D_\mu U|^2\right] - V(|U|^2), \tag{3.4}$$

where $U = \rho \exp(i\phi)$, and the potential fixes ρ, whose fluctuations we ignore. Thus one can think of ϕ as being the Goldstone boson eaten by A to become massive, and so in the limit that m is very small (compared to the energies) then in a very real sense, ϕ carries the longitudinal polarisation of the gauge boson.

This makes many aspects of the theory much easier to understand [47], in particular it is much simpler to discover the cutoff of the effective theory, controlled by the interactions of longitudinally polarised gauge bosons, by examining the dynamics of ϕ. It is this property which will prove most useful to me here, though there are additional reasons to favour the Stückelberg approach, which I briefly mention below.

The gauge invariance makes it easy to see what kind of higher order terms will be generated by quantum corrections, for example in the above there would be generated terms such as $\mathrm{tr}|D^2U|^2$, which in unitary gauge, $U = 1$, gives a non-gauge invariant kinetic term for A_μ of the form $g^2\mathrm{tr}(\partial A)^2$, however, since the original Lagrangian (3.1) was not gauge invariant it was not immediately obvious why such terms should not be generated at leading order [47].

Finally, this proceedure may make it easier to discover viable UV completions for the theory. For example a theory consisting of one massless and many massive spin-1 fields arises in the context of dimensional deconstruction, and each of the U fields introduced there may be understood as arising from a condensation of two fermions linking the gauge fields via a set of new, strongly-interacting gauge fields, as described in [46].

3.2 Bi-gravity

I will now review how the Stückelberg trick can be applied to bi-gravity, as has been studied for example in [77, 136]. As described in section 2.5.1, in a bi-metric theory with two metrics g and f, the interaction term breaks the combined gauge invariance of the kinetic terms, $\mathrm{diff}_g \times \mathrm{diff}_f$, down to the diagonal subgroup.

For the reasons described in the previous section, one might want to return to a theory with the two copies of diff invariance. The way this is done is to turn one of the metrics into an object which transforms under the gauge group of the other, *e.g.*

$$f_{\mu\nu}(x) \to F_{\mu\nu}(x) = \partial_\mu Y^\alpha \partial_\nu Y^\beta f_{\alpha\beta}\left(Y(x)\right), \tag{3.5}$$

where new fields Y^μ have been introduced, which have the following transformation properties:[1]

$$\begin{aligned} \mathrm{diff}_g &: Y^\mu(x) \to Y^\mu\left(\phi(x)\right) \\ \mathrm{diff}_f &: Y^\mu(x) \to \phi^{-1\mu}\left(Y(x)\right), \end{aligned} \tag{3.6}$$

where $x^\mu \to \phi^\mu(x)$ for objects which transform under that symmetry. These ensure that $F_{\mu\nu}$ is a *tensor* under diff_g:

$$F_{\mu\nu}(x) \to \partial_\mu Y^\alpha(\phi(x))\partial_\nu Y^\beta(\phi(x)) f_{\alpha\beta}\left(Y(\phi(x))\right) = \partial_\mu \phi^\alpha \partial_\nu \phi^\beta F_{\alpha\beta}\left(\phi(x)\right), \tag{3.7}$$

and *invariant* under diff_f:

$$\begin{aligned} F_{\mu\nu}(x) \to\ & \partial_\mu\left(\phi^{-1\alpha}(Y)\right)\partial_\nu\left(\phi^{-1\beta}(Y)\right)\frac{\partial\phi^\lambda\left(\phi^{-1}(Y)\right)}{\partial\left(\phi^{-1\alpha}(Y)\right)}\frac{\partial\phi^\rho\left(\phi^{-1}(Y)\right)}{\partial\left(\phi^{-1\beta}(Y)\right)} f_{\lambda\rho}\left(\phi\left(\phi^{-1}(Y)\right)\right) \\ &= F_{\mu\nu}(x), \end{aligned} \tag{3.8}$$

and hence the interaction term $\int \mathrm{d}^D x \sqrt{-|g|}\, e_n\left(\sqrt{g^{-1}F}\right)$ contructed using g and F is invariant under both transformations.[2]

[1] Note that despite appearances, Y^μ does not transform as a vector.

[2] In bi-gravity it does not matter in which 'direction' one does this, *i.e.* make f transform as g or vice versa, however in massive gravity it does, since the reference metric f does not naturally transform under anything, and so if we map g to transform as f the resulting object is *not* a tensor, but in fact is *invariant* under diff. Whilst in principle there is nothing wrong with this, it does complicate the analysis since objects now do not transform as their indices indicate, which can lead to drawing some erroneous conclusions, as in [137].

	field	diff_g:	diff_f:
Before	g	tensor	invariant
	f	invariant	tensor
After	g	tensor	invariant
	$F = f \circ Y$	tensor	invariant

Table 3.1: The Stückelberg trick for an interaction term coupling g and f.

This can also all be succinctly stated using functional composition notation:

$$\begin{aligned} & F = f \circ Y \\ \mathrm{diff}_g : Y \to Y \circ \phi \qquad & \Rightarrow \qquad F \to f \circ Y \circ \phi = F \circ \phi \\ \mathrm{diff}_f : Y \to \phi^{-1} \circ Y \qquad & \Rightarrow \qquad F \to f \circ \phi \circ \phi^{-1} \circ Y = F, \end{aligned} \tag{3.9}$$

and is summarised in table 3.1, and can be described using theory graphs as

$$g \xleftarrow[Y_{(f,g)}]{} f$$

Also, I will frequently refer to 'mapping' one site to another: this just means introducing a Stückelberg field for one site to make it transform under the gauge symmetry of the other, *e.g.* the diagram directly above would be described by 'mapping f to g.'

3.2.1 The decoupling limit

As in the spin-1 example, it is useful to expand the Stückelberg fields about the identity

$$Y^\mu(x) = x^\mu + A^\mu, \tag{3.10}$$

as well as to expand the metrics about Minkowski space:[3] $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, and $f_{\mu\nu} = \eta_{\mu\nu} + l_{\mu\nu}$. Note that since we are expanding both metrics in the same way, from now on the Minkowski metric will always be used to raise and lower indices.

According to the discussion in section 3.1.1, in the limit of small m, A^μ will carry the $D-1$ longitudinal polarisations of the massive spin-2 field—the number of *d.o.f.* of a

[3] Of course one could expand about other backgrounds (which are solutions to the field equations of the theory), in particular (anti-)de Sitter space, however Minkowski is chosen here for simplicity, and because this is the arena in which micro-physics plays out.

massive spin-1 field—and so it is useful to further introduce a $U(1)$ gauge symmetry via

$$A^\mu \to A^\mu + \partial^\mu \pi, \tag{3.11}$$

where now π carries the longitudinal polarisation of A (which now carries the $D-2$ *d.o.f.* of a massless spin-1 field). Thus the full Stückelberg replacement to do is

$$\begin{aligned} f_{\mu\nu} \to F_{\mu\nu} &= \partial_\mu \left(x^\alpha + A^\alpha + \partial^\alpha \pi\right) \partial_\nu \left(x^\beta + A^\beta + \partial^\beta \pi\right) f_{\alpha\beta}\left(x + A + \partial\pi\right) \\ &= \eta_{\mu\nu} + 2\partial_{(\mu} A_{\nu)} + 2\pi_{\mu\nu} + \partial_\mu A^\alpha \partial_\nu A_\alpha + 2\partial_{(\mu} A^\alpha \pi_{\nu)\alpha} + \pi_\mu^\alpha \pi_{\alpha\nu} \\ &\quad + l_{\mu\nu} + 2\partial_{(\mu} A^\alpha l_{\nu)\alpha} + 2\pi^\alpha_{(\mu} l_{\nu)\alpha} + \left(A^\alpha + \pi^\alpha\right) \partial_\alpha l_{\mu\nu} + \cdots , \end{aligned} \tag{3.12}$$

where I have used the shorthand notation $\partial_\mu \partial_\nu \cdots \pi = \pi_{\mu\nu\cdots}$. In the case of massive gravity with a flat reference metric ($f_{\mu\nu} = \eta_{\mu\nu}$) the terms on the final line are not present.

Before substituting this into the Lagrangian in (2.48), let us consider first the expansion of that Lagrangian just up to quadratic order in the tensor fields. One has

$$\begin{aligned} &-\frac{1}{4} M_g^{D-2} h^\alpha_\mu \mathcal{E}^{\mu\ \ \nu}_{\ \alpha\ \beta} h^\beta_\nu - \frac{1}{4} M_f^{D-2} l^\alpha_\mu \mathcal{E}^{\mu\ \ \nu}_{\ \alpha\ \beta} l^\beta_\nu \\ &+ \frac{m^2 M_{\rm eff}^{D-2}}{4} \left\{ \frac{\alpha_2}{4} \left[(l^\mu_\mu - h^\mu_\mu)(l^\nu_\nu - h^\nu_\nu) - (l_{\mu\nu} - h_{\mu\nu})(l^{\mu\nu} - h^{\mu\nu}) \right] \right. \\ &+ \alpha_1 \left[(l^\mu_\mu - h^\mu_\mu) + \frac{1}{2}(l^\mu_\mu - h^\mu_\mu) h^\nu_\nu - (l_{\mu\nu} - h_{\mu\nu}) h^{\mu\nu} - \frac{1}{4}(l_{\mu\nu} - h_{\mu\nu})(l^{\mu\nu} - h^{\mu\nu}) \right] \\ &\left. + \alpha_0 \left(2 + h^\mu_\mu + \frac{1}{4} h^\mu_\mu h^\nu_\nu - \frac{1}{2} h_{\mu\nu} h^{\mu\nu} \right) \right\}, \end{aligned} \tag{3.13}$$

and the spin-2 kinetic operator is defined in (2.2). The second line is a mass term for $l_{\mu\nu} - h_{\mu\nu}$ which is of Fierz-Pauli form—precisely the right form to ensure that the first two lines just propagate one massless and one massive spin-2 *d.o.f.* (see section 2.1)—whilst the third and fourth lines contain quadratic terms which perturb this structure along with tadpoles. As discussed in section 2.4, the appearance of these indicates that we have expanded around the wrong background, and thus to ensure consistency one should set $\alpha_0 = \alpha_1 = 0$. Though for the sake of completeness they will be retained for the remainder of this chapter.

Now substituting (3.12) into (3.13) and ignoring total derivatives, one has the following pieces involving the Stückelberg fields up to quadratic order:

$$\frac{m^2M_{\rm eff}^{D-2}}{4}(\alpha_2-\alpha_1)\left\{-\frac{1}{4}G_{\mu\nu}G^{\mu\nu}+(l_{\mu\nu}-h_{\mu\nu})\left[\eta^{\mu\nu}(\partial_\lambda A^\lambda+\Box\pi)-(\partial^\mu A^\nu+\pi^{\mu\nu})\right]\right\} \tag{3.14}$$

where $G_{\mu\nu}=\partial_\mu A_\nu-\partial_\nu A_\mu$ is the field strength tensor for the Stückelberg vector. One notes that this field has an ordinary kinetic term, whilst the Stückelberg scalar, on the other hand, does not. Before discussing how one can be introduced, let us consider the form of the terms involving just the Stückelberg scalars. These are easily derived by noting that $F^\mu_\nu|_{l=A=0}=[(1+\partial^2\pi)^2]^\mu_\nu$, and so $\left(\sqrt{g^{-1}F}-I_D\right)^\mu_\nu\Big|_{(h)^0,(l)^0,(A)^0}=\pi^\mu_\nu$. Thus they take the form

$$\mathcal{L}_{\partial^2\pi}=\frac{1}{2}m^2M_{\rm eff}^{D-2}\sum_{n=0}^{D}\alpha_n\mathcal{L}^{\rm TD}_{(n)}(\pi) \tag{3.15}$$

where

$$\mathcal{L}^{\rm TD}_{(n)}(\pi)=\frac{1}{n!}\delta^{\mu_1\cdots\mu_n}_{\nu_1\cdots\nu_n}\pi^{\nu_1}_{\mu_1}\cdots\pi^{\nu_n}_{\mu_n}=\frac{1}{n!}\partial_{\mu_1}\left(\pi^{\nu_1}\delta^{\mu_1\cdots\mu_n}_{\nu_1\cdots\nu_n}\pi^{\nu_2}_{\mu_2}\cdots\pi^{\nu_n}_{\mu_n}\right) \tag{3.16}$$

is the unique combination of n copies of $\partial^2\pi$ which is a total derivative.

Meanwhile the terms involving just scalars and one Stückelberg vector are a little more complicated to derive. One has

$$\begin{aligned}&\left(\sqrt{g^{-1}F}-\eta\right)_{\mu\nu}\Big|_{(h)^0,(l)^0,(A)^1}\\&=2\sum_{i=1}^{\infty}{}^{\frac{1}{2}}C_i\sum_{j=0}^{i-1}\left[(2\partial^2\pi+(\partial^2\pi)^2)^j\right]^\alpha_\mu\partial_{(\alpha|}A_\beta(\delta^\beta_{|\gamma)}+\pi^\beta_{|\gamma)})\left[(2\partial^2\pi+(\partial^2\pi)^2)^{i-1-j}\right]^\gamma_\nu,\end{aligned} \tag{3.17}$$

and by using the cyclicality of the trace,

$$\begin{aligned}\mathcal{L}_{\partial A\partial^2\pi}&=\frac{1}{2}m^2M_{\rm eff}^{D-2}\sum_{n=0}^{D}\alpha_n\left(\sqrt{g^{-1}F}-\eta\right)_{\mu\nu}\Big|_{(h)^0,(l)^0,(A)^1}X^{\mu\nu}_{(n-1)}(\pi)\\&=m^2M_{\rm eff}^{D-2}\sum_{n=0}^{D}\alpha_n\sum_{i=1}^{\infty}i\,{}^{\frac{1}{2}}C_i\partial_{(\mu|}A_\alpha(\delta^\alpha_\beta+\pi^\alpha_\beta)\left[(2\partial^2\pi+(\partial^2\pi)^2)^{i-1}\right]^\beta_{|\nu)}X^{\mu\nu}_{(n-1)}(\pi)\\&=\frac{1}{2}m^2M_{\rm eff}^{D-2}\sum_{n=0}^{D}\alpha_n\partial_\mu A_\nu X^{\mu\nu}_{(n-1)}(\pi)\end{aligned} \tag{3.18}$$

where

$$X^{\mu\nu}_{(n)}(\pi)=\frac{1}{n!}\eta^{\nu\lambda}\delta^{\mu\mu_1\cdots\mu_n}_{\lambda\nu_1\cdots\nu_n}\pi^{\nu_1}_{\mu_1}\cdots\pi^{\nu_n}_{\mu_n}=\frac{\delta}{\delta\pi_{\mu\nu}}\mathcal{L}^{\rm TD}_{(n+1)}(\pi). \tag{3.19}$$

From their definition once can see that these X-tensors are symmetric and are conserved ($\partial_\mu X^{\mu\nu}_{(n)}=0$); this latter property means that (3.18) is a total derivative.

3.2.1.1 Ostrogradsky's theorem

The fact that both (3.15) and (3.18) vanish up to total derivatives is very important, and is indeed central to the avoidance of the BD ghost in dRGT theory. This is because of Ostrogradsky's theorem [83, 138], which states that a non-degenerate (classical) system which contains higher than second (time) derivatives in the equations of motion inevitably possesses a runaway instability. Thus for a theory to be healthy at the classical level it must satisfy this restriction on the number of time derivatives, which means that, barring lucky cancellations when deriving these from the action—a point which will prove to be important in section 3.3—it must be possible to bring the action into a form in which at most first order derivatives act on the fields. A requirement that the pure scalar interactions in the present case clearly would not satisfy were it not for the fact that they vanish up to total derivatives.

The equations of motion being higher (than second) order means that more than two initial condition per field are required in order to solve them, and thus the system contains more *d.o.f.* than it would naïvely seem. The extra *d.o.f.* are called ghosts, and another way to appreciate their appearance (and understand their name) is to start with a higher derivative theory and reformulate it thus

$$\int \mathrm{d}^D x \left[-\frac{1}{2}(\partial\pi)^2 + \frac{1}{M^2}\left(\Box\pi\right)^2\right] \tag{3.20}$$
$$= \int \mathrm{d}^D x \left[-\frac{1}{2}(\partial\pi)^2 + \sigma\Box\pi - \frac{1}{2}M^2\sigma^2\right]$$
$$= \int \mathrm{d}^D x \left[-\frac{1}{2}(\partial\pi)^2 + \frac{1}{2}(\partial\sigma)^2 - \frac{1}{2}M^2\sigma^2\right], \tag{3.21}$$

where to get to the second line an auxiliary field σ was introduced, followed by a field redefinition $\pi \to \pi - \sigma$ to get to the third line. Thus (3.20) actually contains two *d.o.f.*, and one of them has the wrong sign for its kinetic term. Such a field, with a wrong sign kinetic term, is called a ghost, and quantum mechanically leads to disaster [83, 84]! Either one is forced to admit states of negative norm (and thus can no longer do quantum mechanics), or one is stuck with a Hamiltonian that is not bounded from below (which means that there are no stable states as it is possible to continuously decay by producing particles of increasingly negative energy).[4]

[4] Of course in (3.21) the healthy π field and ghostly σ field do not interact, and so we can independently

There have however been some attempts to deal with this quantisation of higher-derivative theories, for example see [139] for one recent approach.

3.2.1.2 Fierz-Pauli massive gravity

It is perhaps illuminating at this point to briefly cast one's mind back to the linear theory of massive spin-2 presented in section 2.1. Applying the Stückelberg trick to this, the replacement one must make is $h_{\mu\nu} \to h_{\mu\nu} + 2\partial_{(\mu}A_{\nu)} + 2\pi_{\mu\nu}$, and thus for a general mass term $h_{\mu\nu}h^{\mu\nu} - a(h^\mu_\mu)^2$, the contribution from $\pi_{\mu\nu}$ would be $(a-1)(\Box\pi)^2$—yielding a ghost, which only vanishes when Fierz-Pauli tuning ($a = 1$) is enforced.

If one were to non-linearly complete the kinetic term to the Einstein-Hilbert Lagrangian for $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, then the Stückelberg replacement would have to become non-linear (3.12); if the mass term were kept as Fierz-Pauli: $h_{\mu\nu}h_{\alpha\beta}\left(\eta^{\mu\alpha}\eta^{\nu\beta} - \eta^{\mu\nu}\eta^{\alpha\beta}\right)$, then it is easy to see that the terms from the replacement which are non-linear in $\pi_{\mu\nu}$ will lead to expressions such as $\pi^\mu_{[\nu}\pi^\lambda_{\mu]}\pi^\nu_\lambda$, which are not total derivatives, and hence will generate ghosts.

3.2.1.3 Scalar-tensor mixing

Returning now to non-linear bi-gravity, the other set of interaction terms which will prove to be very important are those involving one copy of a tensor field and several copies of the Stückelberg scalar. The terms involving $h_{\mu\nu}$ take the form

$$\begin{aligned}
\mathcal{L}_{h\partial^2\pi} &= \frac{1}{4}m^2 M_{\text{eff}}^{D-2} h_{\mu\nu} \sum_{n=0}^{D} \left(\sum_{i=n}^{D} \beta_i{}^{D-n-1}C_{i-n} \right) X^{\mu\nu}_{(n)}(\pi) \\
&= \frac{1}{4}m^2 M_{\text{eff}}^{D-2} h_{\mu\nu} \sum_{n=0}^{D} \left(\alpha_n - \alpha_{n+1}\right) X^{\mu\nu}_{(n)}(\pi).
\end{aligned} \tag{3.22}$$

The terms involving $l_{\mu\nu}$ can be derived elegantly by using the symmetry of the dRGT interaction terms (2.24), along with the gauge invariance of the action after applying the

treat each of them consistently and have no problems, however the problems mentioned above will appear as soon as the two sectors couple.

Stückelberg trick [78]:

$$
\begin{aligned}
&\int \mathrm{d}^D x \sqrt{-|g|}\, e_n \left(\sqrt{g^{-1}(f \circ Y)}\right) \\
&= \int \mathrm{d}^D x \sqrt{-|(f \circ Y)|}\, e_{D-n} \left(\sqrt{(f \circ Y)^{-1} g}\right) \\
&= \int \mathrm{d}^D x \sqrt{-|f|}\, e_{D-n} \left(\sqrt{f^{-1}\left(g \circ Y^{-1}\right)}\right),
\end{aligned} \tag{3.23}
$$

and thus all the interactions between f and the Stückelberg fields, coming from the $\sqrt{-|g|}\, e_n \left(\sqrt{g^{-1} f}\right)$ term, take exactly the same form as those involving g coming from the $\sqrt{-|g|}\, e_{D-n} \left(\sqrt{g^{-1} f}\right)$ term, except that the Stückelberg fields are replaced by their *duals*, which are defined using Y^{-1} instead of Y. The dual fields are described in more detail in the next section, and chapter 4. Thus, denoting by ϕ the dual to the Stückelberg scalar π, the interaction terms involving $l_{\mu\nu}$ look like

$$
\mathcal{L}_{l\partial^2\pi} = \frac{1}{4} m^2 M_{\text{eff}}^{D-2} l_{\mu\nu} \sum_{n=0}^{D} \left(\sum_{i=0}^{D-i} \beta_i{}^{D-n-1}C_{D-i-n} \right) X^{\mu\nu}_{(n)}(\phi). \tag{3.24}
$$

This expression can also be derived in a more involved way [77], which I include in appendix A for the reader's amusement.

The symmetry of the dRGT interaction terms does not remain in quite so simple a form when expressed in terms of $\sqrt{g^{-1}f} - I_D$, but nonetheless an equivalent expression can be derived[5] which leads one to[6]

$$
\mathcal{L}_{l\partial^2\pi} = \frac{1}{4} m^2 M_{\text{eff}}^{D-2} l_{\mu\nu} \sum_{n=0}^{D} \left(\sum_{i=0}^{n} \alpha_{i+1} (-1)^i \, {}^nC_i \right) X^{\mu\nu}_{(n)}(\phi). \tag{3.26}
$$

As mentioned above, the Stückelberg scalar does not have its own kinetic term at this stage, and also it mixes with the tensor fields at quadratic order (*i.e.* there is a $h\partial^2\pi$ term in the action). To proceed with the analysis we must diagonalise the action at quadratic order, *i.e.* we must remove the scalar-tensor mixing and in doing so introduce a kinetic term for π.

[5] For completeness, it is

$$
\sqrt{-|g|}\, e_n \left(\sqrt{g^{-1}f} - I_D\right) = (-1)^n \sum_{k=n}^{D} {}^kC_n \sqrt{-|f|}\, e_k \left(\sqrt{f^{-1}g} - I_D\right). \tag{3.25}
$$

[6] Alternatively this can be derived directly from (3.24) by relating β_n and α_n.

This can be done through the following field redefinitions:

$$
\begin{aligned}
h_{\mu\nu} &\to h_{\mu\nu} + \frac{c}{2(D-2)} m^2 \left(\frac{M_{\text{eff}}}{M_g}\right)^{D-2} \pi\eta_{\mu\nu} \\
l_{\mu\nu} &\to l_{\mu\nu} + \frac{c}{2(D-2)} m^2 \left(\frac{M_{\text{eff}}}{M_f}\right)^{D-2} \phi\eta_{\mu\nu},
\end{aligned} \tag{3.27}
$$

where

$$
c = \sum_{n=0}^{D} \beta_n {}^{D-2}C_{n-1} = \alpha_1 - \alpha_2; \tag{3.28}
$$

this works because the spin-2 kinetic operator (2.2) satisfies

$$
\mathcal{E}^{\mu\nu\alpha\beta}\pi\eta_{\alpha\beta} = \frac{D-2}{2} X^{\mu\nu}_{(1)}(\pi). \tag{3.29}
$$

The result is that the scalar-tensor mixing is removed at lowest order and the scalar gains new self-interaction terms[7]

$$
\begin{aligned}
\mathcal{L}_{\pi\partial^2\pi} = &\frac{\left(\sum_{n=0}^{D} \beta_n {}^{D-2}C_{n-1}\right)}{8(D-2)} m^4 M_{\text{eff}}^{D-2} \times \\
&\sum_{n=0}^{D}(D-n)\Bigg\{ \left(\frac{M_{\text{eff}}}{M_g}\right)^{D-2} \left[\sum_{i=n}^{D} \beta_i {}^{D-n-1}C_{i-n}\right] \pi\mathcal{L}^{\text{TD}}_{(n)}(\pi) \\
&\qquad + \left(\frac{M_{\text{eff}}}{M_f}\right)^{D-2} \left[\sum_{i=0}^{D-n} \beta_i {}^{D-n-1}C_{D-i-n}\right] \phi\mathcal{L}^{\text{TD}}_{(n)}(\phi) \Bigg\}
\end{aligned} \tag{3.30}
$$

$$
\begin{aligned}
= &\frac{\alpha_1-\alpha_2}{8(D-2)} m^4 M_{\text{eff}}^{D-2} \sum_{n=0}^{D}(D-n)\Bigg\{ \left(\frac{M_{\text{eff}}}{M_g}\right)^{D-2} (\alpha_n - \alpha_{n+1})\pi\mathcal{L}^{\text{TD}}_{(n)}(\pi) \\
&\qquad + \left(\frac{M_{\text{eff}}}{M_f}\right)^{D-2} \left[\sum_{i=0}^{n} \alpha_{i+1}(-1)^i {}^nC_i\right] \phi\mathcal{L}^{\text{TD}}_{(n)}(\phi) \Bigg\}.
\end{aligned} \tag{3.31}
$$

In particular, it gains a kinetic term[8]

$$
-\frac{D-1}{D-2}\left(\frac{c}{4}\right)^2 m^4 M_{\text{eff}}^{D-2} \left[\left(\frac{M_{\text{eff}}}{M_g}\right)^{D-2} (\partial\pi)^2 + \left(\frac{M_{\text{eff}}}{M_f}\right)^{D-2} (\partial\phi)^2\right]. \tag{3.32}
$$

Now that the fields are de-mixed at the quadratic order it is possible to canonically normalise them (which must be done if one wishes to, for example, read off the strong-coupling scale of the theory from its higher order terms); this takes the form

$$
h_{\mu\nu} \to \frac{2h_{\mu\nu}}{M_g^{\frac{D-2}{2}}}, \quad l_{\mu\nu} \to \frac{2l_{\mu\nu}}{M_f^{\frac{D-2}{2}}}, \quad A_\mu \to \frac{2}{\sqrt{\alpha_2}} \frac{A_\mu}{m M_{\text{eff}}^{\frac{D-2}{2}}}, \quad \pi \to c_\pi \frac{\pi}{m^2 M_{\text{eff}}^{\frac{D-2}{2}}}, \tag{3.33}
$$

[7]There will of course also be other new interactions arising from terms higher order in the tensor fields, however for reasons described below, those can be ignored.

[8]If $\alpha_1 \neq 0$ and there is a ϕ tadpole (arising from an $l_{\mu\nu}$ tadpole) then there will be an extra contribution from the expansion of that in terms of π (3.49).

with $c_\pi = \frac{2\sqrt{2}(D-2)}{\alpha_2(D-1)}\left[\left(\frac{M_{\rm eff}}{M_g}\right)^{D-2} + \left(\frac{M_{\rm eff}}{M_f}\right)^{D-2}\right]^{-\frac{1}{2}}$, and I have now set $\alpha_1 = 0$. Also, from now on I will set $M_g = M_f = M$ (although there is of course lots of interesting physics that arises when there is a hierarchy between them, *e.g.* see [42, 43, 140]).

After canonical normalisation a term arising from the interaction, which, before performing the field redefinitions (3.27) required for scalar-tensor de-mixing, consisted of n_h tensor fields, n_A vectors, and n_π scalars will have a dimensional pre-factor given by $m^{2-n_A-2n_\pi}M^{\frac{D-2}{2}(2-n_h-n_A-n_\pi)}$, or, in other words, it is suppressed by an energy scale $\Lambda_\lambda^{\frac{1}{2}((D+2)n_\pi+Dn_A+(D-2)n_h-2D)}$, where

$$\Lambda_\lambda = \left(m^{\lambda-1}M^{\frac{D-2}{2}}\right)^{\frac{1}{\lambda+\frac{D-4}{2}}}, \qquad \text{and} \qquad \lambda = \frac{3n_\pi + 3n_A + n_h - 4}{n_\pi + n_A + n_h - 2}. \tag{3.34}$$

Larger λ means smaller Λ_λ, and so let us now investigate which terms will be suppressed by the lowest scale, *i.e.* which ones have the largest λ. Terms such as those in (3.22) and (3.24), *i.e.* with $n_h = 1$, $n_A = 0$, have $\lambda = 3$, and a little algebra reveals that any terms suppressed by a lower scale must have $n_h = 0$, $n_A = 0, 1$; these are simply the terms in (3.15) and (3.18), which vanish. Thus the lowest scale in the theory is the famous $\Lambda_3 = \left(m^2 M^{\frac{D-2}{2}}\right)^{\frac{2}{D+2}}$; this scale is then the cutoff of the effective field theory (EFT).[9] Terms with $n_h = 0$ and $n_A = 2$ are also suppressed by this scale, however they prove not be as interesting or as important for my analysis as those with $n_A = 0$ and so I do not give them here, but the interested reader can find them in equation (5.26).

Therefore the Stückelberg vector does not appear linearly—a property which holds provided the theory graph is acyclic; if a minimal coupling to matter were included, then A_μ would couple to matter like $\int \mathrm{d}^D x\, \partial_\mu A_\nu T^{\mu\nu}$, which vanishes (up to a total derivative) if the stress energy tensor is conserved. These two facts means that classically one can consistently set the Stückelberg vector to zero (though quantum mechanically it can still be relevant).

Note that the de-mixing field redefinition (3.27) does not change the suppression scale, and so the terms in (3.31) also are suppressed by Λ_3. Further, this means that one does not have to worry about the new scalar interactions which arise from higher orders in the

[9] Of course, because there is nothing pathological about the terms suppressed by this scale, strictly speaking it is just the strong-coupling scale of the theory (*i.e.* the scale at which perturbative unitarity breaks down), and, as mentioned in the introduction, the scale at which new physics enters could be higher. However, I will occasionally abuse terminology by referring to this as the cutoff of the EFT.

tensor field as these will be suppressed by a higher scale. Also note that one can rewrite λ as $\lambda = 3 - \frac{n_A+2n_h-2}{n_\pi+n_A+n_h-2}$, from which one sees that, for terms with $n_A + 2n_h > 2$, as n_π tends towards infinity, the suppression scale tends towards Λ_3. In an effective theory these terms can be safely neglected, since they only come in at infinitely high order in an expansion in terms of E/Λ_3.

It is also worthwhile now to take stock of how gauge transformations act on the canonically normalised fields in the theory, which can be derived from (3.6) and (3.11). There are three gauge invariances:

$$\mathrm{diff}_g : \begin{cases} \delta h_{\mu\nu} = \partial_\mu \xi_\nu + \partial_\nu \xi_\mu + \frac{2}{M^{\frac{D-2}{2}}} \mathcal{L}_\xi h_{\mu\nu} \\ \delta l_{\mu\nu} = 0 \\ \delta A^\mu = m\sqrt{\frac{\alpha_2}{2}} \xi^\mu + \frac{2}{M^{\frac{D-2}{2}}} \xi^\nu \partial_\nu A^\mu \\ \delta\pi = 0, \end{cases} \tag{3.35}$$

$$\mathrm{diff}_f : \begin{cases} \delta h_{\mu\nu} = 0 \\ \delta l_{\mu\nu} = \partial_\mu \xi_\nu + \partial_\nu \xi_\mu + \frac{2}{M^{\frac{D-2}{2}}} \mathcal{L}_\xi l_{\mu\nu} \\ \delta A^\mu = -m\sqrt{\frac{\alpha_2}{2}} \xi^\mu - \frac{2}{M^{\frac{D-2}{2}}} A^\nu \partial_\nu \xi^\mu - \sqrt{\frac{2}{\alpha_2}} \frac{2}{\sqrt{2} m M^{D-2}} A^\nu A^\lambda \partial_{\nu\lambda} \xi^\mu + \cdots \\ \delta\pi = 0, \end{cases} \tag{3.36}$$

$$U(1) : \begin{cases} \delta h_{\mu\nu} = -m\frac{\sqrt{2\alpha_2}}{(D-2)} \eta_{\mu\nu} \chi \\ \delta l_{\mu\nu} = -m\frac{\sqrt{2\alpha_2}}{(D-2)} \eta_{\mu\nu} (-\chi + \cdots) \\ \delta A^\mu = \partial^\mu \chi \\ \delta\pi = -m\sqrt{\frac{\alpha_2}{2}} \frac{D-1}{D-2} \chi, \end{cases} \tag{3.37}$$

where $\mathcal{L}$ here indicates the Lie derivative; the transformation of the tensor fields under the $U(1)$ arises from the field redefinition (3.27) used to de-mix them from the scalar at lowest order, and the higher order terms in $\delta l_{\mu\nu}$ come from the expansion of ϕ in terms of π given in (3.49).

Those interaction terms which are suppressed by the smallest energy scale are (generally) those which will have the largest effect on the low energy physics, and thus it is useful to focus on these. This can be done by taking the so-called *decoupling limit*:

$$m \to 0, \qquad M \to \infty, \qquad \Lambda_3 \text{ fixed.} \tag{3.38}$$

All terms suppressed by a scale larger than Λ_3 vanish in this limit;[10] also note that sending $m \to 0$ is just what is needed for the Goldstone boson equivalence theorem to be applicable, and so in the decoupling limit it really is the case that A and π carry the helicity-1 and helicity-0 components of the massive graviton.

One may wonder to what extent the decoupling limit is ever a good approximation physically. In order for this to be the case, there must be some range of energies which are simultaneously large with respect to the graviton mass (so that the GBET can be applied) and small with respect to Λ_3 (so that the neglected terms are not important). To get an idea of what range of m this leaves, consider in four dimensions the ratio $\Lambda_3/m = (M_{\rm Pl}/m)^{1/3}$; requiring this to be *at least* a few orders of magnitude (say 10^4 so as to have $m/E \sim E/\Lambda_3 \sim 1\%$ in a small window) then means that we require $m < 10^{-12} M_{\rm Pl} \sim 10^6\,\text{GeV}$.

3.2.2 Bi-vielbein theory and the Stückelberg trick

Before discussing the very rich structure of the terms in (3.31), it makes sense to briefly mention how everything described above proceeds in the bi-vielbein theory. The specific way in which the calculation proceeds will be covered in section 5.3.1, however one immediately knows that the results must of course be identical to the bi-metric theory, due to the equivalence of the two as demonstrated in section 2.4.2.

The key difference is that the dRGT interaction terms break not just the diff invariances of the kinetic terms, but also the local Lorentz invariances, and thus for those we should also introduce Stückelberg fields, so that the total Stückelberg replacement is

$$F^a_\mu(x) \to \Lambda^a{}_b \partial_\mu Y^\nu F^b_\nu \left(Y(x)\right). \tag{3.39}$$

The Lorentz Stückelberg field, $\Lambda^a{}_b$, will clearly be an auxiliary field, and it ends up enforcing as a constraint the DvN symmetric vielbein condition, as described in section 2.4.2. In order to know its behaviour upon taking the decoupling limit, the fluctuation

[10]In particular, the mixing between the tensor and vector fields in (3.14) is removed in this limit.

of the Lorentz Stückelberg field about the identity must be appropriately normalised, however as it lacks a kinetic term this is a little subtle, as shall be seen in section 5.3.2.

3.3 Galileons and dualities

The pure scalar interaction terms sitting at the scale Λ_3 (3.31) are an example of so-called *Galileons* [61],[11] which have a very rich structure that I shall explore in this section.

In flat space the n-th Galileon term can be written as:

$$\mathcal{L}^{\text{Gal}}_{(n)} = \frac{1}{\Lambda^{3(n-1)}} \pi \mathcal{L}^{\text{TD}}_{(n)}(\pi), \tag{3.40}$$

where Λ is some energy scale, and in D dimensions there are clearly D of them (ignoring the tadpole term). This Lagrangian contains second order derivatives (contained in $\mathcal{L}^{\text{TD}}_{(n)}$), and so one might expect that it should lead to higher order equations of motion and via Ostrogradsky's theorem to an instability; however calculating the equations of motion from (3.40) yields

$$\begin{aligned}\partial_{\mu\nu}\left(\frac{\partial \mathcal{L}^{\text{Gal}}_{(n)}}{\partial \pi_{\mu\nu}}\right) + \frac{\partial \mathcal{L}^{\text{Gal}}_{(n)}}{\partial \pi} &= \frac{1}{\Lambda^{3(n-1)}}\left[\partial_{\mu\nu}\left(\pi X^{\mu\nu}_{(n-1)}(\pi)\right) + \mathcal{L}^{\text{TD}}_{(n)}(\pi)\right] \\ &= \frac{n+1}{\Lambda^{3(n-1)}} \mathcal{L}^{\text{TD}}_{(n)}(\pi) = 0,\end{aligned} \tag{3.41}$$

where $\partial_\mu X^{\mu\nu}_{(n)} = 0$ has been used. Thus the special antisymmetric structure of the Galileon Lagrangian ensures that the equations of motion are still only second order, which means that they avoid the pathologies associated with higher derivative theories at the classical and quantum level.

Galileons also posses a number of other remarkable properties. First of all, their name is due to that fact that their action is invariant under the following transformation

$$\pi(x) \to \pi(x) + c + b_\mu x^\mu, \tag{3.42}$$

where c and b_μ are respectively a constant scalar and vector; this generalised shift symmetry is of course reminiscent of Galilean transformations in classical mechanics. As mentioned in the introduction, Galileons arise in the DGP model, as well as naturally from theories of probe branes in higher dimensional spaces, in which the symmetry (3.42)

[11] Nicolis *et al.* [61] actually independently rediscovered the earlier work of [141, 142].

results from higher dimensional Poincaré invariance [62] (and in fact this leads to many ways to generalise them: see [143] for a short review). The fact that the symmetry (3.42) is non-linearly realised also points to another way of understanding Galileons, as parameterising the coset $Gal(D,1)/SO(D-1,1)$—where the Galileon group $Gal(D,1)$ consists of the generators of (3.42) along with the Poincaré group (and the coset has this form because the Lorentz group is still linearly realised)—see for example [144] for details.

The second very useful property is that there exist theorems which state that the Galileon terms themselves are not renormalised [35, 60, 61] (though quantum corrections will generate additional terms). Consider an n-point interaction built from vertices of the form in (3.40); if all of the external legs are ones with two derivatives acting on the fields, then this will clearly generate interactions of the form $(\partial^2\pi)^n$, *i.e.* not of the form (3.40). If one of the external legs is one with no derivatives acting on the field then the momentum structure of the vertex to which that external leg, with momentum p, is attached looks like

$$\delta^{\mu\mu_1\cdots}_{\nu\nu_1\cdots}\left(p_\mu - \sum_i k^{(i)}_\mu\right)\left(p^\nu - \sum_i k^\nu_{(i)}\right)\prod_i k^{(i)}_{\mu_i} k^{\nu_i}_{(i)} \tag{3.43}$$

which, due to the special antisymmetric structure in (3.40), is equal to

$$\delta^{\mu\mu_1\cdots}_{\nu\nu_1\cdots} p_\mu p^\nu \prod_i k^{(i)}_{\mu_i} k^{\nu_i}_{(i)}. \tag{3.44}$$

In position space this would correspond to two derivatives acting on the external leg, and so again this will generate interactions of the form $(\partial^2\pi)^n$, not of the form (3.40). Since the same antisymmetric structure is present in the interactions of the helicity-2 and helicity-0 modes in the bi-gravity decoupling limit—(3.22) and (3.24)[12]—these terms will also be protected from quantum corrections (as in fact will those involving the helicity-1 mode) [110].

Thus the first set of terms in (3.31) are explicitly healthy; whilst the second set, on the other hand, do take the Galileon form, they do this in terms of ϕ rather than π, and since we know that there is really only one scalar field, *i.e.* $\phi = \phi(\pi)$, it remains to show that these are, in fact, healthy.

[12]Of course for the l-ϕ terms, really one needs to re-express them in terms of π, however this can be easily done using the techniques introduced in the next chapter, and one sees that the antisymmetric structure is explicitly preserved (in particular, see (4.13)).

3.3.0.1 Explicit expression for $\phi = \phi(\pi)$

One way this can be done is by explicitly finding an expression for the dual Stückelberg scalar ϕ as a function of the Stückelberg scalar π [78]. They are related by

$$Y(x) = x + \partial\pi = (x + \partial\phi)^{-1} = \left(Y^{-1}(x)\right)^{-1}, \tag{3.45}$$

which can be written succinctly as $Y^{-1}(Y(x))^\mu = x^\mu$. Expanding Y and Y^{-1} in terms of π and ϕ respectively one then has

$$\begin{aligned} & (x+\partial\pi)^\mu + \frac{\partial}{\partial(x+\partial\pi)_\mu}\phi(x+\partial\pi) = x^\mu, \\ \implies \qquad & \partial^\mu\phi = -\partial^\mu\pi - \sum_{n=1}^{\infty} D_{(n)}\partial^\mu\phi, \qquad \text{where} \qquad D_{(n)} = \frac{1}{n!}\pi^{\nu_1}\cdots\pi^{\nu_n}\partial_{\nu_1\cdots\nu_n}. \end{aligned} \tag{3.46}$$

This final form allows one to recursively solve for $\partial^\mu\phi$ in terms of $\partial^\mu\pi$ as

$$\partial^\mu\phi = \partial^\mu\left(-\pi + \frac{1}{2}\pi^\nu\pi_\nu + \ldots\right). \tag{3.47}$$

Note that the first terms in (3.47) are such that $\partial^\mu\phi$ is a total derivative; it is certainly not obvious that this continues to all orders, since $[D,\partial] \neq \partial(\cdots)$, but nonetheless it does, and so it makes sense to talk of ϕ rather than just $\partial^\mu\phi$. After some algebra one then reaches an expression for the n-th order piece in terms of lower orders:

$$\phi\big|_{\pi^n} = -\sum_{i=1}^{n-1} D_{(n-i)}\left(\phi\big|_{\pi^i}\right), \tag{3.48}$$

which can be solved to give

$$\phi = -\pi + \sum_{n=2}^{\infty}\frac{1}{2(n-1)!}\sum_{i=0}^{n-2}{}^{n-2}C_i(-1)^i\tilde{D}_{(i)}\left(\pi^\mu\pi_\mu\mathcal{L}^{\mathrm{TD}}_{(n-2-i)}(\pi)\right), \tag{3.49}$$

where $\tilde{D}_{(n)}(X) = \partial_{\nu_1\cdots\nu_n}\left(\pi^{\nu_1}\cdots\pi^{\nu_n}X\right)$ for some Lorentz scalar X. We have now explicitly related the dual field ϕ to π, and what remains is to show that a Galileon constructed from ϕ remains of Galileon form when this expression is used.

Not a little algebra reveals that, up to total derivatives,

$$\phi\mathcal{L}^{\mathrm{TD}}_{(n)}(\phi) = (-1)^{n+1}(n+1)\sum_{i=n}^{D}\frac{{}^iC_n}{i+1}\pi\mathcal{L}^{\mathrm{TD}}_{(i)}(\pi) \tag{3.50}$$

and so indeed the the field redefinition (3.49), or equivalently the relation (3.45), maps one

$$\bigcirc \xleftarrow{\;Y = x + \partial\pi\;} \bigcirc \qquad \sim \qquad \bigcirc \xrightarrow{\;Y^{-1} = x + \partial\phi\;} \bigcirc$$

Figure 3.1: The Galileon duality is equivalent to changing the direction of the Stückelberg link field.

Galileon theory into another.[13] This is the *Galileon duality* [73] and from this discussion one sees that it is equivalent to changing the direction of the Stückelberg link field [78,136], as shown in figure 3.1, *i.e.* it is essentially a result of gauge invariance.

3.4 Multi-gravity

In this section I will describe in general some of the differences, compared to bi-gravity, that one encounters when using the Stückelberg trick to analyse a multi-gravity theory.

3.4.1 Different ways to apply the Stückelberg trick

The first difference that one encounters is that there are ways of applying the Stückelberg trick that are seemingly different, yet ultimately equivalent [77]. They are described here, and presented graphically in figure 3.2. For concreteness they will be illustrated by the following tri-metric action whose interactions have the structure

$$\mathcal{S}_{\rm int} = \mathcal{S}_1[g_{(1)}, g_{(2)}] + \mathcal{S}_2[g_{(2)}, g_{(3)}]. \tag{3.55}$$

[13]For completeness, this means that the scalar self interaction terms sitting at the cutoff (before canonical normalisation), *i.e.* (3.31) when expressed solely in terms of π, are

$$\mathcal{L}_{\pi\partial^2\pi} = \frac{\left(\sum_{n=0}^{D} \beta_n {}^{D-2}C_{n-1}\right)}{8(D-2)} m^4 M_{\rm eff}^{D-2} \sum_{n=0}^{D}\sum_{i=0}^{D} \beta_i \gamma_{i,n}(D, M_g, M_f)^{D-n-1} C_{i-n}(D-n)\pi\mathcal{L}^{\rm TD}_{(n)}(\pi) \tag{3.51}$$

$$= \frac{\alpha_1 - \alpha_2}{8(D-2)} m^4 M_{\rm eff}^{D-2} \sum_{n=0}^{D} \tilde{\alpha}_n(D, M_g, M_f)\pi\mathcal{L}^{\rm TD}_{(n)}(\pi) \tag{3.52}$$

with

$$\gamma_{i,n}(D, M_g, M_f) = \left(\frac{M_{\rm eff}}{M_g}\right)^{D-2} + \left(\frac{M_{\rm eff}}{M_f}\right)^{D-2} \frac{i(i-n)(nD - [(n+1)i - (n-1)])}{(n+1)(D-i)} \tag{3.53}$$

$$\begin{aligned}\tilde{\alpha}_n(D, M_g, M_f) = {} & \left(\frac{M_{\rm eff}}{M_g}\right)^{D-2} (\alpha_n - \alpha_{n+1})(D-n) \\ & + \left(\frac{M_{\rm eff}}{M_f}\right)^{D-2} \frac{\alpha_{n-1} n(n-1) + \alpha_n n(D-2n) - \alpha_{n+1}(n+1)(D-n)}{n+1}.\end{aligned} \tag{3.54}$$

The astute reader will notice that this does not give the same kinetic term as (3.32) when $\phi = -\pi + \cdots$ is used: this is because the contribution to $(\partial\pi)^2$ coming from the ϕ tadpole is not included in (3.32), and which vanishes when $\alpha_1 = 0$.

1. Treat every site identically, *i.e.* introduce a Stückelberg field for each site, essentially mapping them to some external site.[14]

$$\mathcal{S}_{\text{int}} \to \mathcal{S}_{\text{int},1} = \mathcal{S}_1[g_{(1)} \circ Y_{(1,x)}, g_{(2)} \circ Y_{(2,x)}] + \mathcal{S}_2[g_{(2)} \circ Y_{(2,x)}, g_{(3)} \circ Y_{(3,x)}]. \tag{3.56}$$

   This has the advantage of symmetry between all the sites, but it has the disadvantage noted previously that the resulting objects are not tensors (but instead are invariant), which complicates the analysis.

2. Pick one site, and map every other site to that, *e.g.*

$$\mathcal{S}_{\text{int}} \to \mathcal{S}_{\text{int},2} = \mathcal{S}_1[g_{(1)}, g_{(2)} \circ Y_{(2,1)}] + \mathcal{S}_2[g_{(2)} \circ Y_{(2,1)}, g_{(3)} \circ Y_{(3,1)}]. \tag{3.57}$$

   This has the advantage of perhaps being the easiest to understand, since it results in every $g \circ Y$ in the action transforming under the same symmetry; it has the disadvantage that terms like $\mathcal{S}[g_{(2)} \circ Y_{(2,1)}, g_{(3)} \circ Y_{(3,1)}]$ contain contributions which look like they introduce a ghost, but are just gauge artefacts (see section 5.2.1 for an example of the terms which appear, though in that case they are not just gauge artefacts).

3. Treat every interaction term independently, and for each, pick one site to which to map all the other fields in that term *e.g.*

$$\mathcal{S}_{\text{int}} \to \mathcal{S}_{\text{int},3} = \mathcal{S}_1[g_{(1)}, g_{(2)} \circ Y_{(2,1)}] + \mathcal{S}_2[g_{(2)} \circ Y_{(2,3)}, g_{(3)}]. \tag{3.58}$$

   This at first seems that it must be an incorrect thing to do, since a given field is not being mapped consistently, *e.g.* $g_{(2)}$ is made to transform under $\text{diff}_{(1)}$ in the first term, but under $\text{diff}_{(3)}$ in the second, note however that in the first it is $g_{(2)} \circ Y_{(2,1)}$ that transforms under $\text{diff}_{(1)}$ and $g_{(2)} \circ Y_{(2,3)}$ that transforms under $\text{diff}_{(3)}$, whilst $g_{(2)}$ itself still transforms under $\text{diff}_{(2)}$. This method has the advantage of yielding the simplest Lagrangians in the end, and so is that which will be used throughout the rest of this thesis (with the exception of sections 5.2.1 and 5.3.2.1 in the context of cycles).

[14]Technically this is not new to multi-gravity, as one could do the same in bi-gravity.

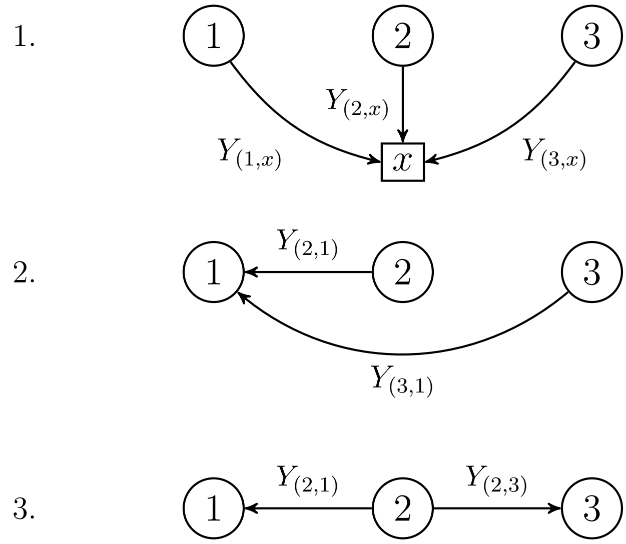


Figure 3.2: Different ways of applying the Stückelberg trick in a multi-gravity theory: 1. map all the sites to an external node; 2. map all the sites to one node; 3. treat each interaction term independently.

These approaches are actually all equivalent and related by gauge transformations, which, in the end, they must be, as they all describe the same physics. First consider the relations between the fields presented, which can easily be understood by considering how fields transform:

$$\begin{aligned} &Y_{(2,x)} \circ Y^{-1}_{(1,x)} = Y_{(2,1)} \\ &Y_{(2,x)} \circ Y^{-1}_{(3,x)} = Y_{(2,3)} \\ &Y_{(3,x)} \circ Y^{-1}_{(1,x)} = Y_{(3,1)} = Y^{-1}_{(2,3)} \circ Y_{(2,1)}. \end{aligned} \tag{3.59}$$

That all three $Y_{(i,x)}$ can be expressed in terms of the two $Y_{(i,j)}$ is a direct manifestation of the fact that there are really only two Stückelberg field degrees of freedom here (*i.e.* in the first approach one too many sets of would-be Goldstone bosons were introduced). In this sense equation (3.59) can be seen as providing the constraint eliminating one of the $Y_{(i,x)}$. We can now explicitly relate interaction terms (3.56)–(3.58) to each other. Consider (3.57)

$$\begin{aligned} \mathcal{S}_{\text{int},2} &= \mathcal{S}_1\left(g_{(1)}, g_{(2)} \circ Y_{(2,1)}\right) + \mathcal{S}_2\left(g_{(2)} \circ Y_{(2,1)}, g_{(3)} \circ Y_{(3,1)}\right) \\ &= \mathcal{S}_1\left(g_{(1)}, g_{(2)} \circ Y_{(2,1)}\right) + \mathcal{S}_2\left(g_{(2)} \circ Y_{(2,1)}, g_{(3)} \circ Y^{-1}_{(2,3)} \circ Y_{(2,1)}\right) \\ &= \mathcal{S}_1\left(g_{(1)}, g_{(2)} \circ Y_{(2,1)}\right) + \mathcal{S}_2\left(g_{(2)}, g_{(3)} \circ Y^{-1}_{(2,3)}\right) \\ &= \mathcal{S}_1\left(g_{(1)}, g_{(2)} \circ Y_{(2,1)}\right) + \mathcal{S}_2\left(g_{(2)} \circ Y_{(2,3)}, g_{(3)}\right) = \mathcal{S}_{\text{int},3} \end{aligned} \tag{3.60}$$

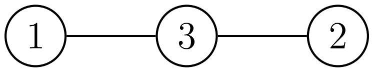


Figure 3.3: The path graph on three vertices, P_3, and the theory graph for (3.62)

In the second line I have used (3.59) to substitute for $Y_{(3,1)}$; in the third line I have made a diff$_{(1)}$ transformation with parameter $Y^{-1}_{(2,1)}$—note that $\mathcal{S}_1\left(g_{(1)}, g_{(2)} \circ Y_{(2,1)}\right)$ is gauge-invariant under diff$_{(i)}$, allowing one to explicitly apply a 'diffeomorphism' just to the second term; in the fourth line I have made a diff$_{(2)}$ transformation with parameter $Y_{(2,3)}$—the same comment as before applies here. Similarly one has

$$\begin{aligned}
\mathcal{S}_{\text{int},1} &= \mathcal{S}_1\left(g_{(1)} \circ Y_{(1,x)}, g_{(2)} \circ Y_{(2,x)}\right) + \mathcal{S}_2\left(g_{(2)} \circ Y_{(2,x)}, g_{(3)} \circ Y_{(3,x)}\right) \\
&= \mathcal{S}_1\left(g_{(1)} \circ Y_{(1,x)}, g_{(2)} \circ Y_{(2,1)} \circ Y_{(1,x)}\right) + \mathcal{S}_2\left(g_{(2)} \circ Y_{(2,1)} \circ Y_{(1,x)}, g_{(3)} \circ Y_{(3,1)} \circ Y_{(1,x)}\right) \\
&= \mathcal{S}_1\left(g_{(1)}, g_{(2)} \circ Y_{(2,1)}\right) + \mathcal{S}_2\left(g_{(2)} \circ Y_{(2,1)}, g_{(3)} \circ Y_{(3,1)}\right) = \mathcal{S}_{\text{int},2} = \mathcal{S}_{\text{int},3}. \qquad (3.61)
\end{aligned}$$

3.4.2 Scalar de-mixing

The second key difference is that in multi-gravity, the field redefinitions (3.27) used to remove the lowest order scalar-tensor mixing, generically lead to mixing of the scalars at quadratic order. This mixing must be removed (*i.e.* the quadratic action of the Stückelberg scalars must be diagonalised) in order to be dealing with the true propagating degrees of freedom in the theory (rather than linear combinations), and it turns out that this can have a large impact on the strong coupling scale of the theory, as discussed in chapter 6.

The necessity of, and the procedure for, de-mixing of the scalar modes is best illustrated here with an example, and it will be discussed in more detail and generality in later chapters. Consider a tri-metric theory in which two of the fields (labelled 1 and 2) interact only with a third (labelled 3), and whose interaction Lagrangian is (also see figure 3.3)

$$\mathcal{L}_{\text{int}} = \frac{1}{4} m^2 M_{\text{eff}}^{D-2} \sum_{n=0}^{D} \left[\beta_n^{(1)} \sqrt{-\left|g_{(3)}\right|}\, e_n\left(\sqrt{g^{-1}_{(3)} g_{(1)}}\right) + \beta_n^{(2)} \sqrt{-\left|g_{(3)}\right|}\, e_n\left(\sqrt{g^{-1}_{(3)} g_{(2)}}\right)\right] \qquad (3.62)$$

Now introduce Stückelberg fields $Y_{(1)}$ and $Y_{(2)}$ mapping $g_{(1)}$ and $g_{(2)}$ to site 3 in each term as appropriate. For the remainder of this subsection I set $m = M = 1$ in order to focus on the interaction structure of the theory, since the explicit mass scales are the same as in bi-gravity. To cubic order in the fields the scalar-tensor mixing then looks like

$$\begin{aligned}\mathcal{L}_{h\pi} =& h^{(1)}_{\mu\nu}\left[a_1 X^{\mu\nu}_{(1)}(\phi_{(1)}) + b_{1,R} X^{\mu\nu}_{(2)}(\phi_{(1)})\right] + h^{(2)}_{\mu\nu}\left[a_2 X^{\mu\nu}_{(1)}(\phi_{(2)}) + b_{2,R} X^{\mu\nu}_{(2)}(\phi_{(2)})\right] \\ &+ h^{(3)}_{\mu\nu}\left[a_1 X^{\mu\nu}_{(1)}(\pi_{(1)}) + b_{1,L} X^{\mu\nu}_{(2)}(\pi_{(1)}) + a_2 X^{\mu\nu}_{(1)}(\pi_{(2)}) + b_{2,L} X^{\mu\nu}_{(2)}(\pi_{(2)})\right], \end{aligned} \tag{3.63}$$

where the coefficients are

$$a_i = \sum_{n=1}^{D-1} \beta^{(i)}_n {}^{D-2}C_{n-1} = \alpha^{(i)}_1 - \alpha^{(i)}_2, \tag{3.64}$$

$$b_{i,L} = \sum_{n=2}^{D-1} \beta^{(i)}_n {}^{D-3}C_{n-2} = \alpha^{(i)}_2 - \alpha^{(i)}_3, \qquad b_{i,R} = \sum_{n=1}^{D-2} \beta^{(i)}_n {}^{D-3}C_{n-1} = \alpha^{(i)}_1 - 2\alpha^{(i)}_2 + \alpha^{(i)}_3.$$

One can de-mix this through the following field redefinition:

$$\begin{aligned} h^{(i)}_{\mu\nu} &\to h^{(i)}_{\mu\nu} + \frac{1}{4(D-2)} a_i \phi_{(i)} \eta_{\mu\nu} \qquad \text{for } i = 1, 2 \\ h^{(3)}_{\mu\nu} &\to h^{(3)}_{\mu\nu} + \frac{1}{4(D-2)} (a_1 \pi_{(1)} + a_2 \pi_{(2)}) \eta_{\mu\nu}, \end{aligned} \tag{3.65}$$

and it is clear that this will introduce terms quadratic (and higher order) in the fields which mix (1) fields with (2) fields. To examine the quadratic and cubic terms we must pick one from each pair of a Stückelberg field and its dual, and rewrite the other in terms of that via (3.49)

$$\phi = -\pi + \frac{1}{2}(\partial\pi)^2 + \cdots, \tag{3.66}$$

(or equivalently for π in terms of ϕ); call the so-chosen fields $\rho_{(i)}$ and write $\tilde{\rho}_{(i)} = a_i \rho_{(i)}$. Provided that $\alpha^{(i)}_1 = 0$, *i.e.* there are no tadpoles for $h^{(i)}_{\mu\nu}$,[15] one has the following kinetic terms

$$\mathcal{L}_{\rho\Box\rho} \propto \begin{pmatrix} \tilde{\rho}_{(1)} & \tilde{\rho}_{(2)} \end{pmatrix} \begin{pmatrix} 2 & \sigma \\ \sigma & 2 \end{pmatrix} \begin{pmatrix} \Box\tilde{\rho}_{(1)} \\ \Box\tilde{\rho}_{(2)} \end{pmatrix} = \tilde{\rho}^{\mathrm{T}} K \Box \tilde{\rho}, \tag{3.67}$$

where $\sigma = 1$ if $\{\rho_{(i)}\} = \{\pi_{(1)}, \pi_{(2)}\}$ or $\{\phi_{(1)}, \phi_{(2)}\}$, *i.e.* either $g_{(1)}$ is being always mapped to the other sites, or $g_{(2),(3)}$ are both being mapped to site 1; and $\sigma = -1$ otherwise, *i.e.* $g_{(1)}$ is mapped to another site in one interaction term, but not in the other, which corresponds to $\{\rho_{(i)}\} = \{\pi_{(1)}, \phi_{(2)}\}$ or $\{\phi_{(1)}, \pi_{(2)}\}$. See figure 3.4.

At cubic order there are two types of terms: those arising from $\rho_{(i)} \mathcal{L}^{\mathrm{TD}}_{(2)}(\rho_{(j)})$, as appear

[15] One may worry about the appropriateness of this extra restriction, but if there is a tadpole term, then it just means that the expansion has been performed about the wrong background.

$$1 \longrightarrow 3 \longleftarrow 2 \qquad\qquad 1 \longleftarrow 3 \longrightarrow 2$$

Figure 3.4: Stückelberg mapping directions allowed by $\sigma = 1$ in (3.67).

$\{\rho_{(1)}, \rho_{(2)}\}$	C^A	C^B
$\{\pi_{(1)}, \pi_{(2)}\}$	$\begin{pmatrix} a_1^{-2}(b_{1,L}-b_{1,R}) & -a_2^{-2}b_{2,R} \\ -a_1^{-2}b_{1,R} & a_2^{-2}(b_{2,L}-b_{2,R}) \end{pmatrix}$	$\begin{pmatrix} a_1^{-1} & 0 \\ 0 & a_2^{-1} \end{pmatrix}$
$\{\phi_{(1)}, \phi_{(2)}\}$	$\begin{pmatrix} a_1^{-2}(b_{1,R}-b_{1,L}) & a_2^{-2}b_{2,R} \\ a_1^{-2}b_{1,R} & a_2^{-2}(b_{2,R}-b_{2,L}) \end{pmatrix}$	$\begin{pmatrix} a_1^{-1} & a_2^{-1} \\ a_1^{-1} & a_2^{-1} \end{pmatrix}$
$\{\pi_{(1)}, \phi_{(2)}\}$	$\begin{pmatrix} a_1^{-2}(b_{1,L}-b_{1,R}) & -a_2^{-2}b_{2,R} \\ a_1^{-2}b_{1,R} & a_2^{-2}(b_{2,R}-b_{2,L}) \end{pmatrix}$	$\begin{pmatrix} a_1^{-1} & -a_2^{-1} \\ 0 & a_2^{-1} \end{pmatrix}$
$\{\phi_{(1)}, \pi_{(2)}\}$	$\begin{pmatrix} a_1^{-2}(b_{1,R}-b_{1,L}) & a_2^{-2}b_{2,R} \\ -a_1^{-2}b_{1,R} & a_2^{-2}(b_{2,L}-b_{2,R}) \end{pmatrix}$	$\begin{pmatrix} a_1^{-1} & 0 \\ -a_1^{-1} & a_2^{-1} \end{pmatrix}$

Table 3.2: The matrices of coefficients of cubic terms (before quadratic demixing and canonical normalisation of the modes), for the different possible orientations of the links.

in massive gravity, and those arising from the second order expansion of one of the fields in a nominally quadratic term, *e.g.* $\phi_{(1)}\Box\phi_{(2)} \to \frac{1}{2}(\partial\rho_{(1)})^2\Box\rho_{(2)}$. I call these respectively, A-, and B-type terms, and write

$$\mathcal{L}_A \propto 2\frac{D-2}{D-1}\begin{pmatrix}\tilde{\rho}_{(1)} & \tilde{\rho}_{(2)}\end{pmatrix} C^A \begin{pmatrix}\mathcal{L}^{\mathrm{TD}}_{(2)}(\tilde{\rho}_{(1)}) \\ \mathcal{L}^{\mathrm{TD}}_{(2)}(\tilde{\rho}_{(2)})\end{pmatrix} \tag{3.68}$$

$$\mathcal{L}_B \propto 2\begin{pmatrix}\Box\tilde{\rho}_{(1)} & \Box\tilde{\rho}_{(2)}\end{pmatrix} C^B \begin{pmatrix}(\partial\tilde{\rho}_{(1)})^2 \\ (\partial\tilde{\rho}_{(2)})^2\end{pmatrix}, \tag{3.69}$$

where the coefficient matrices $C^{A,B}$ depend on how the links are oriented, and are displayed in table 3.2, and the constants of proportionality are the same as in the kinetic term. Note that there is a degeneracy between the diagonal elements of C^A and C^B, since $(\partial\pi)^2\Box\pi = -\frac{4}{3}\pi\mathcal{L}^{\mathrm{TD}}_{(2)}\pi + \frac{1}{3}\partial_\mu\left((\partial\pi)^2\pi^\mu\right)$.

As the scalars are mixed at the quadratic level, in order to ask questions concerning when interactions become strongly coupled one must find the actual propagating modes, *i.e.* one must diagonalise the kinetic matrix K. One also needs to diagonalise their mass matrix,[16] however when $\alpha_0 = \alpha_1 = 0$ this is automatically achieved by the diagonalisation of the kinetic terms, see section 6.1.4 for more discussion of this. Transforming to the eigenbasis of K involves $\tilde{\rho} \to U\chi$, where U is the matrix whose columns are the normalised eigenvectors of K; the kinetic terms will then look like $\sum_{i=1}^{2}\lambda_i\chi_{(i)}\Box\chi_{(i)}$, and so finally

[16]The mass terms obviously go to zero when the decoupling limit is taken, and so one may wonder whether diagonalising it is really needed, but clearly one wants to be dealing with propagating modes which are correct even away from the decoupling limit.

one must canonically normalise the fields, $\chi_{(i)} \to \chi_{(i)}/\sqrt{\lambda_i}$.

Once this is done the cubic terms will look like

$$\mathcal{L}_A \propto 2\frac{D-2}{D-1}\sum_{ijk}\frac{\tilde{C}^A_{ijk}}{\Lambda_3^{\frac{D+2}{2}}}\chi_{(i)}\mathcal{L}^{\mathrm{TD}}_{(2)}\left(\chi_{(j)},\chi_{(k)}\right) \tag{3.70}$$

$$\mathcal{L}_B \propto 2\sum_{ijk}\frac{\tilde{C}^B_{ijk}}{\Lambda_3^{\frac{D+2}{2}}}\Box\chi_{(i)}\partial_\mu\chi_{(j)}\partial^\mu\chi_{(k)}, \tag{3.71}$$

where $\mathcal{L}^{\mathrm{TD}}_{(2)}(\chi_{(j)},\chi_{(k)}) = \frac{1}{2}\left(\Box\chi_{(j)}\Box\chi_{(k)} - \partial_{\mu\nu}\chi_{(j)}\partial^{\mu\nu}\chi_{(k)}\right)$, and

$$\tilde{C}_{ijk} = \frac{1}{\sqrt{\lambda_i\lambda_j\lambda_k}}\sum_{l,m}C_{lm}U_{li}U_{mj}U_{mk}, \tag{3.72}$$

and the explicit mass scale dependence via $\Lambda_3^{\frac{D+2}{2}}$ has been reintroduced. The least suppressed cubic interaction is the one with the largest (in magnitude) element of $\tilde{C}$, call it $\tilde{C}_{\mathrm{max}}$, and hence the effective strong coupling scale is $\Lambda_3/\left(\tilde{C}_{\mathrm{max}}\right)^{\frac{2}{D+2}}$.

The scalar de-mixing thus shifts the strong coupling scale from the massive/bi-gravity value. Let us now do this explicitly for the tri-metric theory described above and compare with the bi-gravity result; I will specialise to $D=4$, and the case of the first row in table 3.2, as in that case C^B is diagonal and can thus be completely absorbed into C^A. Up to cubic order, the Lagrangian for the scalars is

$$\begin{aligned}\mathcal{L}_{\mathrm{tri}} \propto & \begin{pmatrix}\pi_{(1)} & \pi_{(2)}\end{pmatrix}\begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}\begin{pmatrix}\Box\pi_{(1)}\\ \Box\pi_{(2)}\end{pmatrix}\\ &+\frac{4}{3\alpha_2^{(1)}}\begin{pmatrix}\pi_{(1)} & \pi_{(2)}\end{pmatrix}\begin{pmatrix}1-2\nu^{(1)} & \mu\left(1-\nu^{(2)}\right)\\ 1-\nu^{(1)} & \mu\left(1-2\nu^{(2)}\right)\end{pmatrix}\begin{pmatrix}\mathcal{L}^{\mathrm{TD}}_{(2)}(\pi_{(1)})\\ \mathcal{L}^{\mathrm{TD}}_{(2)}(\pi_{(2)})\end{pmatrix},\end{aligned} \tag{3.73}$$

where $\nu^{(i)} = \frac{\alpha_3^{(i)}}{\alpha_2^{(i)}}$ and $\mu = \frac{\alpha_2^{(1)}}{\alpha_2^{(2)}}$. Although I have set all the mass parameters to unity, by keeping all the $\alpha_j^{(i)}$ arbitrary the interaction mass parameters of the two terms, $m^2_{(i,3)}$ can be absorbed into an overall scaling of $\alpha^{(i)}$, and thus μ can be thought of as $\frac{m^2_{(1,3)}}{m^2_{(2,3)}}$.[17]

The kinetic term in (3.73) is diagonalised and canonically normalised (up to an overall

[17] Unsurprisingly, the graviton masses are a function of μ. The graviton mass matrix is given by (see section 6.2)

$$\mathcal{L}_{h^2} \propto \begin{pmatrix}h^{(1)} & h^{(2)} & h^{(3)}\end{pmatrix}\begin{pmatrix}\mu & 0 & -\mu\\ 0 & 1 & -1\\ -\mu & -1 & 1+\mu\end{pmatrix}\begin{pmatrix}h^{(1)}\\ h^{(2)}\\ h^{(3)}\end{pmatrix}, \tag{3.74}$$

where the tensor structure (which is of Fierz-Pauli form) has been suppressed. This has eigenvalues

$$0, \quad 1+\mu\pm\sqrt{1-\mu+\mu^2}, \tag{3.75}$$

factor) by

$$\begin{pmatrix} \pi_{(1)} \\ \pi_{(2)} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \chi_{(1)} \\ \chi_{(2)} \end{pmatrix}, \tag{3.76}$$

and the new cubic terms are (up to an overall factor)

$$\begin{aligned} \frac{4}{3\alpha_2^{(1)}} \frac{1}{2^{3/2}} \Bigg\{ & \left[\frac{1}{\sqrt{3}} \left[(2 - 3\nu^{(1)}) - \mu(2 - 3\nu^{(1)}) \right] \chi_{(1)} - \left[\nu^{(1)} + \mu\nu^{(2)} \right] \chi_{(2)} \right] \frac{1}{\sqrt{3}} \mathcal{L}^{\mathrm{TD}}_{(2)}(\chi_{(1)}, \chi_{(2)}) \\ & + \left[\frac{1}{\sqrt{3}} \left[(2 - 3\nu^{(1)}) + \mu(2 - 3\nu^{(1)}) \right] \chi_{(1)} - \left[\nu^{(1)} - \mu\nu^{(2)} \right] \chi_{(2)} \right] \left[\frac{1}{3} \mathcal{L}^{\mathrm{TD}}_{(2)}(\chi_{(1)}) + \mathcal{L}^{\mathrm{TD}}_{(2)}(\chi_{(2)}) \right] \Bigg\}. \end{aligned} \tag{3.77}$$

Now consider the bi-metric theory corresponding to either of the two edges alone; this has cubic scalar action

$$\mathcal{L}^{(i)}_{\mathrm{bi}} \propto 2\pi_{(i)} \Box \pi_{(i)} + \frac{4}{3\alpha_2^{(1)}} \mu \left(1 - 2\nu^{(i)} \right) \pi_{(i)} \mathcal{L}^{\mathrm{TD}}_{(2)}(\pi_{(i)}), \tag{3.78}$$

and thus after canonical normalisation, the coefficient of the cubic term, up to the same overall factor in (3.77), is

$$\frac{4}{3\alpha_2^{(1)}} \frac{1}{2^{3/2}} \mu \left(1 - 2\nu^{(i)} \right) \pi_{(i)} \mathcal{L}^{\mathrm{TD}}_{(2)}(\pi_{(i)}). \tag{3.79}$$

As explained above, the cutoff of the effective theory broadly is governed by the largest (in magnitude) coefficient in front of a suppressed term, and thus it is instructive to compare the largest coefficient in (3.77) to the larger of (3.79) for each of $i = 1, 2$. The inverse of this ratio, *i.e.* $\frac{\max(C_{\mathrm{bi}})}{\max(C_{\mathrm{tri}})}$, is shown in figure 3.5 as a function of $\nu^{(1)}$ and $\nu^{(2)}$, for $\mu = 1$ (the behaviour for other values of μ is qualitatively similar), and one sees that it *is* possible to choose coefficients such that the tri-metric theory has a higher cutoff than an equivalent bi-metric one obtained by removing one of the edges in the theory graph. Tuning $\nu^{(1)}$ and $\nu^{(2)}$, the largest value of $\frac{\max(C_{\mathrm{bi}})}{\max(C_{\mathrm{tri}})}$ one can obtain for a given μ (which, without loss of generality, I have taken to be ≤ 1, and also note that one requires $\mu \geq 0$ to avoid the appearance of tachyonic gravitons) is

$$1 + \frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{6} \left(\frac{1 - \mu}{1 + \mu} \right). \tag{3.80}$$

Thus the maximum factor by which the cutoff can be raised is ≈ 1.23 when $\mu = 1$.[18]

to which the graviton masses are proportional.

[18]The fact that this maximum value decreases as the ratio of the interaction term mass parameters

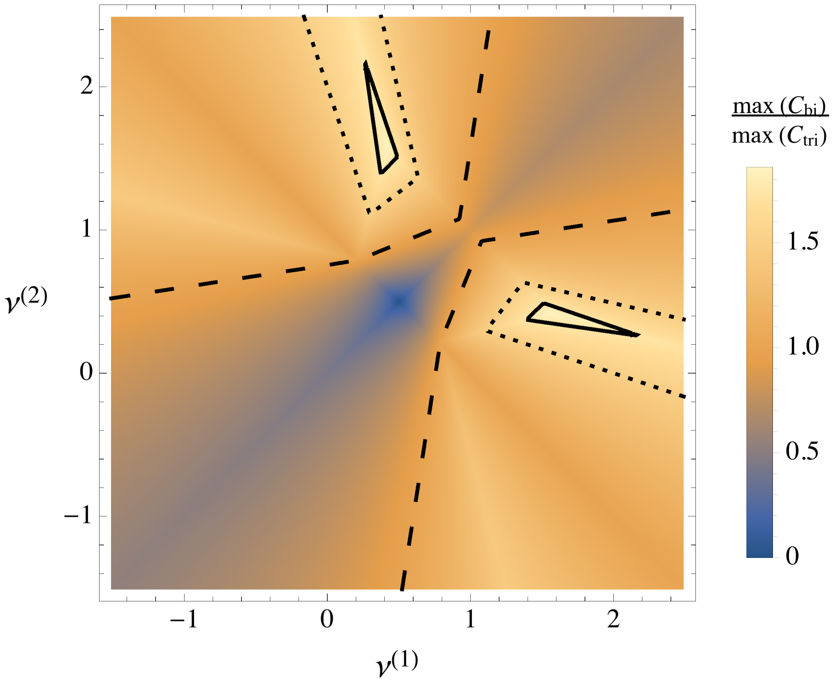


Figure 3.5: The ratio of the maximum coefficient of cubic scalar terms in a tri-gravity theory and the corresponding bi-gravity theory with one edge removed from the theory graph, as function of the interaction term parameters, with idential interaction mass scales; values larger than one indicate that the cutoff in the tri-gravity theory is higher than in corresponding bi-gravity theory. The contour lines are: 1 (dashed); 1.5 (dotted); 1.75 (solid).

Whilst in the tri-metric case this shift is not terribly large, when the number of fields involved is large the effect can be significant; unfortunately however, for larger theory graphs the effect tends to work in the opposite direction and the cutoff is actually *lowered*, as is shown in chapter 6. The reader may wonder whether higher order terms (quartic in the scalar and above) could lead to an effective strong coupling scale which differs more than in the example above. For tri-gravity this may be possible, since the numbers are still $\mathcal{O}(1)$, however as we shall explain in section 6.1.3.1, one would expect higher order terms to be less relevant than cubic for larger numbers of interacting spin-2 fields.

I emphasise that this requirement to de-mix the scalar fields, and the resulting shift in the strong coupling scale, is a qualitatively different feature of multi-gravity, which does not appear in bi- or massive- gravity.

deviates from unity can be understood as due to the fact that the link with the smaller mass parameter will tend to dominate the dynamics—see the next section for more on this.

3.4.3 Multiple decoupling limits

Finally, a further new feature of multi-gravity theories is that the decoupling limit is no longer unique—a fact that was obscured slightly in the previous example, as the mass scales were all set to unity. Consider a multi-gravity theory with general Planck masses, $M_{(i)}$, for each site, and mass parameters $m_{(i,j)}$ for each interaction term, and set $D = 4$ to avoid clutter. A scalar-tensor mixing term then schematically looks like

$$M^2_{(i,j)} m^2_{(i,j)} h^{(i)}_{\mu\nu} X^{\mu\nu}_{(n)}(\pi_{(i,j)}), \tag{3.81}$$

where $M_{(i,j)}$ is defined as M_{eff} for $M_{(i)}$ and $M_{(j)}$. De-mixing this at the lowest order and then performing mass-dimension canonical normalisation[19] leads to interaction terms of the form

$$\frac{1}{\Lambda^{3(n-1)}_{(i,j)}} \pi_{(i,k)} \mathcal{L}^{\text{TD}}_{(n)}(\pi_{(i,j)}), \qquad \text{and} \qquad \frac{1}{\Lambda^{3(n-1)}_{(i,j)}} \frac{M_{(i,j)}}{M_{(i)}} h^{(i)}_{\mu\nu} X^{\mu\nu}_{(n)}(\pi_{(i,j)}), \tag{3.82}$$

which are suppressed by scales

$$\Lambda^3_{(i,j)} = m^2_{(i,j)} M_{(i,j)} = \frac{m^2_{(i,j)} M_{(i)}}{\sqrt{1 + \left(\frac{M_{(i)}}{M_{(j)}}\right)^2}}. \tag{3.83}$$

The decoupling limits which are interesting here involve $M_{(i)} \to \infty$ and $m_{(i,j)} \to 0$, however from (3.83) one sees that within those conditions there are different possibilities for the behaviour of $\Lambda_{(i,j)}$.

Requiring that none vanish, then one has that for each link $m^2_{(i,j)} \min(M_{(i)}, M_{(j)})$ must not go to zero, and $\Lambda_{(i,j)}$ is essentially controlled by the smaller of the two Planck masses for that link. From the scalar-tensor interactions one has that $\frac{M_{(i,j)}}{M_{(i)}}$ must not go to infinity as the decoupling limit is taken, however this will be automatically satisfied, since if there is a hierarchy between $M_{(i)}$ and $M_{(j)}$, then $M_{(i,j)}$ tends towards the smaller one. This of course all agrees with what one expects if the massive gravity limit is taken (one Planck mass goes to infinity with respect to the other).

Thus one set of possible decoupling limits involves sending some Planck masses to infinity quicker than others—essentially giving a 'multi-massive gravity' limit in which

[19] One can perform the mass-dimension canonical normalisation separately, and before, the scalar de-mixing discussed in the previous section, which is ignored here.

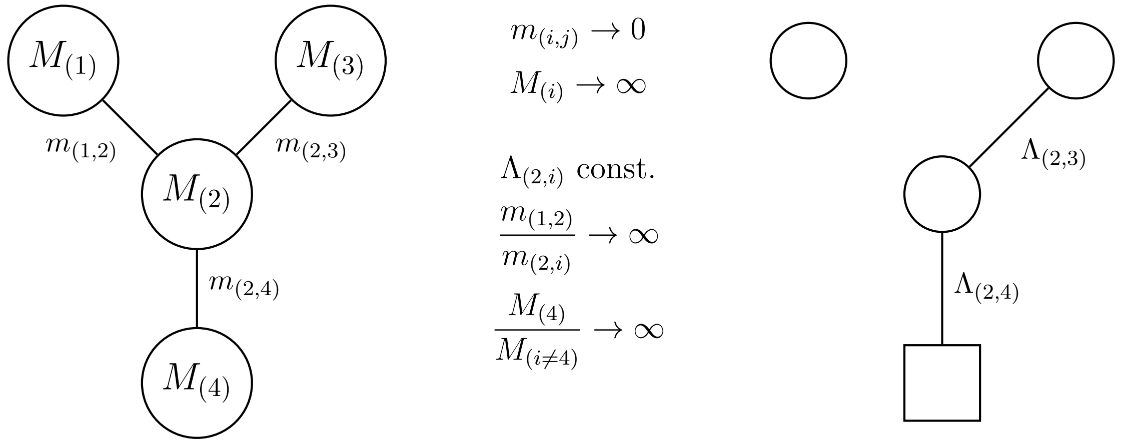


Figure 3.6: Due to the existence of different mass parameters, there are different decoupling limits one can take. Sending an interaction mass parameter to zero slower than others which share a vertex, *viz.* $m_{(1,2)}$ in this case, eliminates those interactions/eliminates that edge relative to others; sending a Planck mass to infinity quicker than others adjacent to it, *viz.* $M_{(4)}$ in this case, freezes the fluctuations of that site relative to those adjacent, making it behave as in massive gravity—here denoted by a square vertex.

the fluctuations of some fields are frozen out. Another thing one could do is to send some interaction mass parameters, $m_{(i,j)}$, to zero slower than others, which essentially removes those edges from the theory graph as the decoupling limit is taken. This is explained graphically in figure 3.6. Note however that the decoupling limit does not change the number of *d.o.f.* in a theory, since the kinetic terms of all the fields, after canonical normalisation, are unaffected by taking it. Thus saying above that some fields are frozen out, or that some edges are removed from the theory graph, is technically incorrect, since they still remain in the theory, and are dynamical, but just are completely decoupled from the rest of the fields in the theory.

Whilst the possibility of taking different decoupling limits is intriguing, and could prove useful either if one wants to isolate and investigate different parts of the theory graph, or if the theory one is considering already has hierarchies between the parameters, in the remainder of this thesis I will consider only the 'ordinary' decoupling limit, in which all Planck masses are taken to infinity at the same rate, and all interaction mass parameters are taken to zero at the same rate.

CHAPTER

4

The Decoupling Limit of Multi-Gravity

In the previous chapter we saw that the decoupling limit of bi-gravity contains Galileons, and so in this chapter I investigate how, and to what extent, this structure extends to multi-gravity. First I shall briefly consider precisely what terms appear and relate this to the graph describing the multi-gravity theory; then in section 4.2 I discuss a more sophisticated way of examining the duality between Galileon theories that was encountered in section 3.3, and in section 4.3 use this to examine the terms which appear in multi-gravity, showing that in general the decoupling limit contains multi-Galileons, although this form is not always manifest in the action; finally I will give a tri-gravity example in section 4.4 in order to demonstrate in a more concrete way the ideas which have been discussed. Throughout this chapter, in order to simplify the discussion, I will ignore explicit mass scales, though it should be obvious how to reintroduce them if the reader so desires. This chapter is largely based on [78].

4.1 Structure of terms in the decoupling limit of a general multi-gravity theory

In the example at the end of the previous chapter we saw that in a tri-gravity theory, in the decoupling limit, the Stückelberg scalar sector contains not just Galileon terms of the form $\pi^{(i)}\mathcal{L}^{\mathrm{TD}}_{(n)}(\pi^{(i)})$ and $\phi^{(i)}\mathcal{L}^{\mathrm{TD}}_{(n)}(\phi^{(i)})$, but also terms mixing the two Stückelberg scalars, *viz.* $\phi^{(1)}\mathcal{L}^{\mathrm{TD}}_{(n)}(\phi^{(2)})$, and $\phi^{(2)}\mathcal{L}^{\mathrm{TD}}_{(n)}(\phi^{(1)})$, which arose from the field redefinitions required to remove the lowest order scalar-tensor mixing and give the scalars a kinetic term.

For a general theory graph the scalar-tensor de-mixing requires the field redefinitions

$$h^{(i)}_{\mu\nu} \to h^{(i)}_{\mu\nu} - \frac{1}{4(D-2)}\sum_{\langle i,j\rangle}\alpha_2^{(i,j)}\pi^{(i,j)}\eta_{\mu\nu}, \tag{4.1}$$

where the sum is over all sites j adjacent to the site i, and the corresponding Stückelberg scalar is $\pi^{(i,j)}$ and its dual $\pi^{(j,i)}$ (note the index ordering). Thus, considering only bi-gravity-type interaction terms, the Stückelberg scalar terms which will appear are

$$\sum_i\sum_{\langle i,j\rangle}\sum_{\langle i,k\rangle}\pi^{(i,j)}\mathcal{L}^{\mathrm{TD}}_{(n)}\left(\pi^{(i,k)}\right), \tag{4.2}$$

and fields on adjacent edges (but no further away) are mixed. The 'direction' of the Stückelberg mapping for the edges connected to site i, by which I mean, for a given interaction term $\sqrt{-\left|g_{(i)}\right|}\,e_n\left(\sqrt{g^{-1}_{(i)}g_{(j)}}\right)$, whether $g^{(j)}$ is made to transform as $g^{(i)}$ or vice versa—or equivalently, of the two fields $\pi^{(i,j)}$ and $\pi^{(j,i)}$ which appear in the action, which will be re-expressed in terms of the other—controls whether the contribution to (4.2) from site i involves pairs of 'fundamental' fields, pairs of fields which should be re-expressed, or mixtures of the two types. Note that when the theory graph does not contain cycles of odd length, it is possible to consistently orient the directions so that such mixtures do not appear—this fact will turn out to be important in chapter 6.

4.1.1 Multi-Galileons

The terms (4.2) have a structure which is identical to the Galileons (3.40), except that multiple fields are involved. In fact they are related to a generalisation of the Galileons—

the aptly named multi-Galileons [145–147]. A general multi-Galileon term looks like

$$\pi^{(i_1)}\delta^{\mu_1\cdots\mu_{n-1}}_{\nu_1\cdots\nu_{n-1}}\pi^{(i_2)\nu_1}{}_{\mu_1}\cdots\pi^{(i_n)\nu_{n-1}}{}_{\mu_{n-1}}, \tag{4.3}$$

and calculating the equations of motion arising from this, one indeed sees that they are of second order, thus preserving that nice property of the Galileons; similarly, they also obey a non-renormalisation theorem [147]. One should also note that whilst there are many different ways to express a Galileon term, related by partial integration, this number is reduced for the multi-Galileon, at least if one wants to maintain explicitly second order equations of motion. For example whilst $\pi^{(i_1)}_{\mu}\pi^{(i_2)\nu}\delta^{\mu\mu_1\cdots}_{\nu\nu_1\cdots}\pi^{(i_3)\nu_1}{}_{\mu_1}\cdots$ still only has second order equations of motion, $\pi^{(i_1)}_{\mu}\pi^{(i_2)\mu}\delta^{\mu_1\cdots}_{\nu_1\cdots}\pi^{(i_3)\nu_1}{}_{\mu_1}\cdots$ only does when $i_1 = i_2 = i_3 = \cdots$.

The property of second order *e.o.m.* is shared by the terms in (4.2), provided one takes each field to be independent. As mentioned above, however, some of the fields will need to be expanded in terms of their duals, in order that one is being consistent when considering propagating modes in the action. One could apply (3.49) appropriately to recast the action without dual fields, and then compare the result to the multi-Galileon form (4.3), however it turns out that this analysis is greatly simplified by considering a more sophisticated way to interpret the relation (3.45) between a field and its dual.

4.2 The Galileon duality map

In section 3.3.0.1 it was shown, by explicitly expanding the dual Stückelberg field ϕ in terms of π, that a Galileon constructed from ϕ is still a Galileon when expressed in terms of π, *i.e.* the relation (3.45) between them defines a duality between different Galileon theories; this can be explored in a more sophisticated manner by introducing the duality map [74]. Also see [148] for discussion of this duality in terms of different parametrisations of the coset $Gal(D,1)/SO(D-1,1)$, as mentioned in section 3.3.

Recall that to get from $\sqrt{-\left|(g\circ Y^{-1})\right|}\,e_n\left(\sqrt{(g\circ Y^{-1})^{-1}f}\right)$ to $\sqrt{-|g|}\,e_n\left(\sqrt{g^{-1}(f\circ Y)}\right)$ one must perform a diffeomorphism

$$\mathcal{D}_\pi:\; x^\mu \to \tilde{x}^\mu = Y^\mu(x) = x^\mu + \partial^\mu\pi; \tag{4.4}$$

this acts in the usual way on all fields, *except* for ϕ/π, whose transformation properties can be derived by noting that the inverse mapping to (4.4) is

$$\mathcal{D}_\pi^{-1} = \mathcal{D}_\phi : \tilde{x}^\mu \to x^\mu = Y^{-1\mu}(\tilde{x}) = \tilde{x}^\mu + \tilde{\partial}^\mu \phi, \tag{4.5}$$

from which one sees that $\partial\phi$ and $\partial\pi$ are related by $\partial^\mu\pi = -\tilde{\partial}^\mu\phi$, *i.e.* it behaves as a scalar (up to a sign). This can be solved to give $\phi(\tilde{x}) = -\pi(x) - \frac{1}{2}(\partial\pi)^2$. Thus in total one has

$$\mathcal{D}_\pi : \begin{cases} \pi(x) \to \phi(\tilde{x}) = -\pi(x) - \frac{1}{2}(\partial\pi)^2 \\ \partial^\mu\pi \to \tilde{\partial}^\mu\phi = -\partial^\mu\pi \\ \Pi^{\mu\nu}(x) \to \Phi^{\mu\nu}(\tilde{x}) = -\left(\delta^\lambda_\mu + \Pi^\lambda_\mu\right)^{-1}\Pi^{\lambda\nu}, \end{cases} \tag{4.6}$$

where $\Pi^{\mu\nu} = \partial^{\mu\nu}\pi$; note that since $\partial^\mu\pi$ transforms as a scalar, $\partial^{\mu\nu}\pi$ transforms as a vector, and I have used $\tilde{\partial}x = 1 + \Phi = (1+\Pi)^{-1} = (\partial\tilde{x})^{-1}$. Any object which does not depend on π behaves in the usual way under coördinate transformations:

$$T^{\mu\cdots}{}_{\nu\cdots}(x) \to \tilde{T}^{\mu\cdots}{}_{\nu\cdots}(\tilde{x}) = (\delta^\mu_\lambda + \Pi^\mu_\lambda)\cdots\left(\delta^\nu_\rho + \Pi^\nu_\rho\right)^{-1}\cdots T^{\lambda\cdots}{}_{\rho\cdots}(x). \tag{4.7}$$

With this machinery it is now possible to easily re-derive (3.50):

$$\begin{aligned}
&\int \mathrm{d}^D\tilde{x}\,\phi(\tilde{x})\mathcal{L}^{\mathrm{TD}}_{(n)}(\phi) \\
&= \frac{(-1)^{n+1}}{n!}\int \mathrm{d}^D x\,|1+\Pi|\left(\pi + \frac{1}{2}(\partial\pi)^2\right)\delta^{\mu_1\cdots\mu_n}_{\nu_1\cdots\nu_n}\left[(1+\Pi)^{-1}\Pi\right]^{\nu_1}_{\mu_1}\cdots\left[(1+\Pi)^{-1}\Pi\right]^{\nu_n}_{\mu_n} \\
&= \frac{(-1)^{n+1}}{n!(D-n)!}\int \mathrm{d}^D x\,|1+\Pi|\left(\pi + \frac{1}{2}(\partial\pi)^2\right)\epsilon^{\mu_1\cdots\mu_D}\epsilon_{\nu_1\cdots\nu_D}\delta^{\nu_{n+1}}_{\mu_{n+1}}\cdots\delta^{\nu_D}_{\mu_D} \\
&\qquad\qquad \times\left[(1+\Pi)^{-1}\Pi\right]^{\nu_1}_{\mu_1}\cdots\left[(1+\Pi)^{-1}\Pi\right]^{\nu_n}_{\mu_n} \\
&= \frac{(-1)^{n+1}}{n!(D-n)!}\int \mathrm{d}^D x\,|1+\Pi|\left(\pi + \frac{1}{2}(\partial\pi)^2\right)\left|(1+\Pi)^{-1}\right|\epsilon^{\lambda_1\cdots\lambda_D}\epsilon_{\nu_1\cdots\nu_D} \\
&\qquad\qquad \times\Pi^{\mu_1}_{\lambda_1}\Pi^{\nu_2}_{\lambda_2}\cdots\Pi^{\nu_n}_{\lambda_n}(1+\Pi)^{\nu_{n+1}}_{\lambda_{n+1}}\cdots(1+\Pi)^{\nu_D}_{\lambda_D} \\
&= \int \mathrm{d}^D x\,(-1)^{n+1}\left(\pi + \frac{1}{2}(\partial\pi)^2\right)\sum_{i=n}^{D}{}^iC_n\mathcal{L}^{\mathrm{TD}}_{(i)}(\pi) \\
&= \int \mathrm{d}^D x\,(-1)^{n+1}\sum_{i=n}^{D}{}^iC_n\left[\pi\mathcal{L}^{\mathrm{TD}}_{(i)}(\pi) - \frac{i+1}{i+2}\pi\mathcal{L}^{\mathrm{TD}}_{(i+1)}(\pi)\right] \\
&= \int \mathrm{d}^D x\,(-1)^{n+1}\left[\pi\mathcal{L}^{\mathrm{TD}}_{(n)}(\pi) + \sum_{i=n+1}^{D}\left({}^iC_n - \frac{i}{i+1}{}^{i-1}C_n\right)\pi\mathcal{L}^{\mathrm{TD}}_{(i)}(\pi)\right] \\
&= \int \mathrm{d}^D x\,(-1)^{n+1}(n+1)\sum_{i=n}^{D}\frac{{}^iC_n}{i+1}\pi\mathcal{L}^{\mathrm{TD}}_{(i)}(\pi),
\end{aligned} \tag{4.8}$$

where in going from the second to the third equality the Kronecker deltas have all been replaced by $[(1+\Pi)^{-1}]^{\lambda_i}_{\mu_i}(1+\Pi)^{\nu_i}_{\lambda_i}$ and use has been made of the identity $\epsilon^{\mu_1\cdots\mu_D}M^{\lambda_1}_{\mu_1}\dots M^{\lambda_D}_{\mu_D} = |M|\epsilon^{\lambda_1\cdots\lambda_D}$; in going from the fourth to fifth equality the identity $(\partial\pi)^2\mathcal{L}^{\mathrm{TD}}_{(n-1)}(\pi) = -\frac{2n}{n+1}\pi\mathcal{L}^{\mathrm{TD}}_{(n)}(\pi) + \frac{1}{n+1}\partial_\mu\left(\pi^\lambda\pi_\lambda\pi_\nu X^{\mu\nu}_{(n-2)}\right)$ has been used.

4.3 Multi-gravity and multi-Galileons

The technique developed in the previous section can clearly be extended to the types of terms (4.2) appearing in multi-gravity theories: one now has multiple duality diffeomorphisms (4.4)–(4.6)—one for each Stückelberg scalar, each of which behaves as an ordinary field under the transformation associated with a different Stückelberg field.

With only bi-gravity-type individual interactions, there are four general types of term which can appear. Consider two Stückelberg scalars π and σ, and their associated duals ϕ and ρ, and imagine that we always want to map from ϕ to π and σ to ρ, then the distinct possibilities which mix the fields are

$$\text{no mapping:} \qquad \int \mathrm{d}^D x\, \pi\mathcal{L}^{\mathrm{TD}}_{(n)}(\sigma) \tag{4.9}$$

$$\text{mapping the field within } \mathcal{L}^{\mathrm{TD}}_{(n)}: \qquad \int \mathrm{d}^D x\, \pi\mathcal{L}^{\mathrm{TD}}_{(n)}(\rho) \tag{4.10}$$

$$\text{mapping the field outside of } \mathcal{L}^{\mathrm{TD}}_{(n)}: \qquad \int \mathrm{d}^D x\, \phi\mathcal{L}^{\mathrm{TD}}_{(n)}(\sigma) \tag{4.11}$$

$$\text{mapping both fields:} \qquad \int \mathrm{d}^D x\, \phi\mathcal{L}^{\mathrm{TD}}_{(n)}(\rho) \tag{4.12}$$

The first of these is already manifestly a bi-Galileon and needs no further consideration. I will now investigate the remaining three in relation to the multi-Galileon form (4.3) above; first by applying the dualities at the level of the action, and then at the level of the equations of motion.

4.3.1 Mapping the action

The second, (4.10), can be dealt with in a similar manner to (4.8):

$$\begin{aligned}
&\int \mathrm{d}^D\tilde{x}\,\pi(\tilde{x})\mathcal{L}^{\mathrm{TD}}_{(n)}(\rho) \\
&\quad= \frac{(-1)^n}{n!}\int \mathrm{d}^D x\,|1+\Sigma|\pi(x)\delta^{\mu_1\cdots\mu_n}_{\nu_1\cdots\nu_n}\left[(1+\Sigma)^{-1}\Sigma\right]^{\nu_1}_{\mu_1}\cdots\left[(1+\Sigma)^{-1}\Sigma\right]^{\nu_n}_{\mu_n} \\
&\quad= \int \mathrm{d}^D x\,(-1)^n\sum_{i=n}^{D}{}^iC_n\pi\mathcal{L}^{\mathrm{TD}}_{(i)}(\sigma),
\end{aligned}\tag{4.13}$$

which is now manifestly of bi-Galileon form. In the same way, the fourth expression above, (4.12), can be brought into a form identical to the third, (4.11), which thus remains the final one to examine.

Under the mapping which takes $\phi(x)$ to $\pi(x)$, $\Sigma_{\mu\nu}=\partial_{\mu\nu}\sigma$ transforms as a tensor

$$\begin{aligned}
\Sigma^\nu_\mu(\tilde{x}) = \tilde{\partial}^\nu_\mu\sigma &= \left[(1+\Pi)^{-1}\right]^\lambda_\mu(1+\Pi)^\nu_\rho\nabla_\lambda\nabla^\rho\sigma \\
&= \left[(1+\Pi)^{-1}\right]^\lambda_\mu(1+\Pi)^\nu_\rho\partial^\rho_\lambda\sigma + \left[(1+\Pi)^{-1}\right]^\lambda_\mu\pi^\nu_{\lambda\rho}\partial^\rho\sigma,
\end{aligned}\tag{4.14}$$

using that the Christoffel symbols in the new coördinate system, $\tilde{x}^\mu = x^\mu + \partial^\mu\pi$, are

$$\Gamma^\mu_{\alpha\beta} = \left[(1+\Pi)^{-1}\right]^\mu_\lambda\pi^\lambda_{\alpha\beta}.\tag{4.15}$$

Alternatively (4.14) may be derived just by a direct coördinate substitution

$$\begin{aligned}
\Sigma^\nu_\mu(\tilde{x}) = \eta^{\nu\rho}\tilde{\partial}_{\mu\rho}\sigma &= \left[(1+\Pi)^{-1}\right]^\lambda_\mu\partial_\lambda\left((1+\Pi)^\nu_\alpha(1+\Pi)^\rho_\beta g^{\alpha\beta}\left[(1+\Pi)^{-1}\right]^\gamma_\rho\partial_\gamma\sigma\right) \\
&= \left[(1+\Pi)^{-1}\right]^\lambda_\mu\partial_\lambda\left((1+\Pi)^\nu_\rho\partial^\rho\sigma\right),
\end{aligned}\tag{4.16}$$

using that the metric in the new coördinate system is

$$g_{\mu\nu} = (1+\Pi)^\alpha_\mu(1+\Pi)^\beta_\nu\eta_{\alpha\beta}.\tag{4.17}$$

Thus (4.11) can be re-expressed

$$\begin{aligned}
&\int \mathrm{d}^D\tilde{x}\,\phi(\tilde{x})\mathcal{L}^{\mathrm{TD}}_{(n)}(\sigma) \\
&\quad= \frac{-1}{n!}\int \mathrm{d}^D x\,|1+\Pi|\left(\pi+\frac{1}{2}(\partial\pi)^2\right)\delta^{\mu_1\cdots\mu_n}_{\nu_1\cdots\nu_n}\left[(1+\Pi)^{-1}\Sigma(\nabla)(1+\Pi)\right]^{\nu_1}_{\mu_1}\cdots \\
&\quad= -\int \mathrm{d}^D x\left(\pi+\frac{1}{2}(\partial\pi)^2\right)\sum_{i=0}^{D-n}\mathcal{L}^{\mathrm{TD}}_{(n,i)}\left(\Sigma(\nabla)(1+\Pi),\Pi\right),
\end{aligned}\tag{4.18}$$

Where $\Sigma(\nabla)^{\mu}_{\nu} = \nabla^{\mu}_{\nu}\sigma$ (note that as the metric is still flat, the covariant derivatives commute, and so a similar notation as used with partial derivatives is adopted), and it is written in this form for later convenience. Note that there is no ambiguity about whether it is covariant or partial derivatives in Π, since $\partial\pi$ transforms as a scalar. Also $\mathcal{L}^{\mathrm{TD}}_{(i,j)}(A, B)$ (and $X^{\mu\nu}_{(i,j)}(A, B)$ which will be required shortly), where A and B are tensors, are defined in a similar way to their single-label counterparts, (3.16) and (3.19) respectively, but with i copies of A and j copies of B:

$$\mathcal{L}^{\mathrm{TD}}_{(i,j)}(A, B) = \frac{1}{i!j!}\delta^{\mu_1\cdots\mu_i\lambda_1\cdots\lambda_j}_{\nu_1\cdots\nu_i\rho_1\cdots\rho_j} A^{\nu_1}_{\mu_1}\cdots A^{\nu_i}_{\mu_i} B^{\rho_1}_{\lambda_1}\cdots B^{\rho_j}_{\lambda_j} \tag{4.19}$$

$$X_{(i,j)}{}^{\alpha}_{\beta}(A, B) = \frac{1}{i!j!}\delta^{\alpha\mu_1\cdots\mu_i\lambda_1\cdots\lambda_j}_{\beta\nu_1\cdots\nu_i\rho_1\cdots\rho_j} A^{\nu_1}_{\mu_1}\cdots A^{\nu_i}_{\mu_i} B^{\rho_1}_{\lambda_1}\cdots B^{\rho_j}_{\lambda_j}. \tag{4.20}$$

(4.18) contains terms which have third order derivatives acting on π, and so when the equations of motion are calculated, there will necessarily be contributions which are higher than second order.[1] On the face of it, this is confusing, since, as discussed in section 2.5.2, in the absence of cycles in the theory graph, the constraint structure ensures that the correct number of *d.o.f.* propagate—yet the cycle-free constraint is clearly not enough to remove terms such as (4.11) and (4.12) above. Thus this apparently higher order system actually describes fewer *d.o.f.* than it would seem; this can be analysed by examining the *e.o.m.* directly.

4.3.1.1 Higher-derivative equations of motion

Previously, in section 3.2.1.1, it was explained how, usually, higher-derivative equations of motion yield theories which propagate more *d.o.f.* than expected and contain instabilities. This, however, is not always true—it is possible for additional constraints to exist which allow higher-derivative equations of motion to nonetheless still yield a theory which propagates the correct number of *d.o.f.*

Before discussing this in the context of (4.18), I shall give a brief example, lifted from [11]. Consider the Lagrangian

$$\mathcal{L} = \frac{1}{2}\pi\Box\pi + \frac{1}{2}\phi\Box\phi + \frac{1}{\Lambda^5}\Box\pi\partial_{\mu\nu}\phi\partial^{\mu\nu}\phi + \frac{1}{2\Lambda^{10}}\partial_{\mu\nu}\phi\partial^{\mu\nu}\phi\Box(\partial_{\lambda\rho}\phi\partial^{\lambda\rho}\phi), \tag{4.21}$$

[1] Note also that even if the $\partial^3\pi$ terms were to vanish, there are terms $(\partial\pi)^2\delta^{\mu_1\cdots}_{\nu_1\cdots}\Sigma^{\nu_1}_{\mu_1}\cdots\Pi^{\nu_i}_{\mu_i}$ which will generate higher order *e.o.m.*, as explained in section 4.1.1.

the equations of motion arising from which are

$$\mathcal{E}_\phi = \Box\phi + \frac{2}{\Lambda^5}\partial_{\mu\nu}\left(\partial^{\mu\nu}\phi\left[\Box\pi + \frac{1}{\Lambda^5}\Box\left(\partial_{\lambda\rho}\phi\partial^{\lambda\rho}\phi\right)\right]\right) = 0 \tag{4.22}$$

$$\mathcal{E}_\pi = \Box\pi + \frac{1}{\Lambda^5}\Box(\partial_{\mu\nu}\phi\partial^{\mu\nu}\phi) = 0, \tag{4.23}$$

which appear to be higher order in derivatives. However, note the explicit dependence of the first *e.o.m.* on the second, which is the way in which the extra constraint manifests itself here. By writing $\hat{\mathcal{E}}_\phi = \mathcal{E}_\phi - \frac{2}{\Lambda^5}\partial_{\mu\nu}\left(\partial^{\mu\nu}\phi\mathcal{E}_\pi\right)$ and $\hat{\mathcal{E}}_\pi = \mathcal{E}_\pi$, one can instead write the *e.o.m.* for ϕ and π as

$$\hat{\mathcal{E}}_\phi\left[\partial^2\phi\right] = \Box\phi = 0 \tag{4.24}$$

$$\hat{\mathcal{E}}_\pi\left[\partial^2\pi, \partial^2\phi, \partial^3\phi, \partial^4\phi\right] = \Box\pi + \frac{1}{\Lambda^5}\Box(\partial_{\mu\nu}\phi\partial^{\mu\nu}\phi) = 0, \tag{4.25}$$

where I have made the dependence on derivatives explicit, *i.e.* $\hat{\mathcal{E}}_\pi\left[\partial^2\pi, \ldots\right]$ denotes that this *e.o.m.* depends on second derivatives acting on π and so on. It is then clear that a solution for ϕ can be found in terms of just two initial conditions. $\hat{\mathcal{E}}_\pi$ is higher-derivative in nature, but the higher-derivative dependence is restricted to ϕ, for which one has already solved. So a solution for π can also be found in terms of two initial conditions and the whole system is actually second-order with a well-defined Cauchy problem. Schematically one derives a final set of *e.o.m.*, $\tilde{\mathcal{E}}_\pi, \tilde{\mathcal{E}}_\phi$, where $\tilde{\mathcal{E}}_\phi = \hat{\mathcal{E}}_\phi$ and $\tilde{\mathcal{E}}_\pi$ is obtained by substituting the solution of $\hat{\mathcal{E}}_\phi$ into $\hat{\mathcal{E}}_\pi$. Consequently one has two final *e.o.m.* $\tilde{\mathcal{E}}_\phi\left[\partial^2\phi\right], \tilde{\mathcal{E}}_\pi\left[\partial^2\pi, \partial^2\phi\right]$. The same conclusion could have been reached by performing an explicit constraint analysis for this model in the Hamiltonian picture. In general, if there is any linear combination of the *e.o.m.* and derivatives of the *e.o.m.* of a system that explicitly allows solutions to be found for all of the fields in terms of two initial conditions for each field (here we are assuming we are dealing with scalars, obviously this count is modified for other tensors), then one has a second-order system free from Ostrogradsky ghosts. Note that the models proposed in [149] as well as the recently proposed 'beyond Horndeski' theories [150, 151] are precisely of this type as well (albeit in the case of 'beyond Horndeski' theories being theories of a scalar and a metric tensor, instead of multiple scalars as discussed here).

Now consider the *e.o.m.* coming from (4.18):

$$\mathcal{E}_\sigma = \nabla^\nu_\mu \left[\left(\pi + \frac{1}{2}(\partial\pi)^2 \right) (1+\Pi)^\lambda_\nu \sum_{i=0}^{D-n} X_{(n-1,i)\lambda}{}^{\mu} \left(\Sigma(\nabla)(1+\Pi), \Pi\right) \right] = 0 \tag{4.26}$$

$$\begin{aligned}\mathcal{E}_\pi = \nabla^\nu_\mu \Bigg[\left(\pi + \frac{1}{2}(\partial\pi)^2 \right) \Bigg(& \Sigma(\nabla)^\mu_\lambda \sum_{i=0}^{D-n} X_{(n-1,i)\nu}{}^{\lambda} \left(\Sigma(\nabla)(1+\Pi), \Pi\right) \\ & + \sum_{i=0}^{D-n-1} X_{(n,i)\nu}{}^{\lambda} \left(\Sigma(\nabla)(1+\Pi), \Pi\right) \Bigg) \Bigg] \\ & + (1 - \Box\pi - \partial^\mu\pi\nabla_\mu) \sum_{i=0}^{D-n} \mathcal{L}^{\mathrm{TD}}_{(n,i)} \left(\Sigma(\nabla)(1+\Pi), \Pi\right) = 0;\end{aligned} \tag{4.27}$$

each contains terms with up to fourth order derivatives acting on π, and up to third order derivatives acting on σ. In principle one should be able to perform a similar procedure to that described above, *i.e.* finding a combination of these equations which eliminates the derivatives higher than second order on one of the fields, allowing one to solve for the other with just two initial conditions, which solution can then be substituted into the other equation in order to solve for the remaining field.

4.3.2 Mapping the equations of motion

One could also imagine calculating the equations of motion arising from the action *before* applying the duality map, which one then applies to the *e.o.m.* directly. The equations one has will then of course not be identical to those which arise from applying the duality map at the level of the action, but they will be equivalent, in as much as they describe the same dynamics.

Before explaining how to show that (4.11) propagates the correct number of *d.o.f.*, it is helpful to demonstrate first how this works for (4.10), which we already know to be manifestly a bi-Galileon after mapping at the level of the action.

The *e.o.m.* arising from (4.10) before one applies the duality map are

$$\tilde{\mathcal{E}}_\rho = \mathcal{L}^{\mathrm{TD}}_{(1,n-1)}(\pi, \rho) = 0 \tag{4.28}$$

$$\tilde{\mathcal{E}}_\pi = \mathcal{L}^{\mathrm{TD}}_{(n)}(\rho) = 0, \tag{4.29}$$

which are then mapped to

$$\hat{\mathcal{E}}_\sigma = (-1)^{n-1}\mathcal{L}^{\mathrm{TD}}_{(1,n-1)}\left((1+\Sigma)^{-1}\Pi(\nabla)(1+\Sigma),(1+\Sigma)^{-1}\Sigma\right) = 0 \tag{4.30}$$

$$\hat{\mathcal{E}}_\pi = (-1)^{n}\mathcal{L}^{\mathrm{TD}}_{(n)}\left((1+\Sigma)^{-1}\Sigma\right) = 0. \tag{4.31}$$

One sees that $\hat{\mathcal{E}}_\sigma$ has up to second order derivatives acting on π, and up to third order acting on σ, whilst $\hat{\mathcal{E}}_\pi$ only has second order derivatives acting on σ. Thus one could take the derivative of $\hat{\mathcal{E}}_\pi$ to arrive at, on shell, a relation between $\partial^3\sigma$ and $\partial^2\sigma$, and this relation can then be substituted into $\hat{\mathcal{E}}_\sigma$ to generate a new *e.o.m.* which explicitly only depends on up to second order derivatives. In this way one sees that (4.30) and (4.31) indeed only describe two *d.o.f.*—in full agreement with the result (4.13) of mapping at the level of the action.[2] Furthermore, one can show that the resulting equations do not involve lower than second order derivatives, and so, like the *e.o.m.* arising from (4.13), they are manifestly invariant under the Galileon symmetry (3.42).

Now return to considering (4.11); the *e.o.m.* arising from it, before one applies the duality map, are

$$\tilde{\mathcal{E}}_\sigma = \mathcal{L}^{\mathrm{TD}}_{(1,n-1)}(\phi,\sigma) = 0 \tag{4.34}$$

$$\tilde{\mathcal{E}}_\phi = \mathcal{L}^{\mathrm{TD}}_{(n)}(\sigma) = 0, \tag{4.35}$$

which are then mapped to

$$\hat{\mathcal{E}}_\sigma = -\mathcal{L}^{\mathrm{TD}}_{(1,n-1)}\left((1+\Pi)^{-1}\Pi,(1+\Pi)^{-1}\Sigma(\nabla)(1+\Pi)\right) = 0 \tag{4.36}$$

$$\hat{\mathcal{E}}_\pi = \mathcal{L}^{\mathrm{TD}}_{(n)}\left((1+\Pi)^{-1}\Sigma(\nabla)(1+\Pi)\right) = 0. \tag{4.37}$$

[2]For completeness, I will briefly show explicitly how this works.

$$\begin{aligned}\partial^\mu\hat{\mathcal{E}}_\pi =&(-1)^\mu\partial^\mu\left[(1+\Sigma)^{-1}\Sigma\right]_{\alpha\beta}X^{\alpha\beta}_{(n-1)}\left((1+\Sigma)^{-1}\Sigma\right)\\ =&(-1)^n\left[(1+\Sigma)^{-1}\right]^\lambda_\alpha\sigma^\mu_{\lambda\gamma}\left(\delta^\gamma_\beta-\left[(1+\Sigma)^{-1}\right]^\gamma_\delta\Sigma^\delta_\beta\right)X^{\alpha\beta}_{(n-1)}\left((1+\Sigma)^{-1}\Sigma\right)\\ =&(-1)^n\left[(1+\Sigma)^{-1}\right]^\lambda_\alpha\sigma^\mu_{\lambda\gamma}\left[(1+\Sigma)^{-1}\right]^\gamma_\beta X^{\alpha\beta}_{(n-1)}\left((1+\Sigma)^{-1}\Sigma\right);\end{aligned} \tag{4.32}$$

setting this to zero on-shell then implies

$$\begin{aligned}\hat{\mathcal{E}}_\sigma =&(-1)^{n-1}\left[(1+\Sigma)^{-1}\right]^\lambda_\alpha\nabla_{\lambda\gamma}\pi\left[(1+\Sigma)^{-1}\right]^\gamma_\beta X^{\alpha\beta}_{(n-1)}\left((1+\Sigma)^{-1}\Sigma\right)\\ =&(-1)^{n-1}\left[(1+\Sigma)^{-1}\right]^\lambda_\alpha\left(\partial_{\lambda\gamma}\pi-\left[(1+\Sigma)^{-1}\right]^\nu_\mu\sigma^\mu_{\lambda\gamma}\partial_\nu\pi\right)\left[(1+\Sigma)^{-1}\right]^\gamma_\beta X^{\alpha\beta}_{(n-1)}\left((1+\Sigma)^{-1}\Sigma\right)\\ =&(-1)^{n-1}\left[(1+\Sigma)^{-1}\right]^\lambda_\alpha\partial_{\lambda\gamma}\pi\left[(1+\Sigma)^{-1}\right]^\gamma_\beta X^{\alpha\beta}_{(n-1)}\left((1+\Sigma)^{-1}\Sigma\right),\end{aligned} \tag{4.33}$$

which is now explicitly second order in derivatives acting on σ and π.

Again, in principle one should be able to remove this $\partial^3\pi$ dependence from one of the equations, though this is not quite as simple as the previous case, because the equations now both have up to third order derivatives acting on π and second order acting on σ, and so simply taking the derivative of one of them will not help.

However there is an alternative way to show that this system can be solved just with two initial conditions for each field. This idea is to use (4.34), the σ *e.o.m.* from before performing the mapping, to re-express Σ in terms of Φ, and then to substitute this expression into the (4.35), the ϕ *e.o.m.* from before performing the mapping, to obtain a new *e.o.m.* $\mathcal{L}^{\mathrm{TD}}_{(n)}(\Sigma(\Phi))$. On performing the mapping, this will then have second order derivatives acting on π, and one is now in the same position as in the previous example—one can take a derivative of this and eliminate the third order derivatives (on π) that are present in (4.36).

4.3.3 Coupling to matter

So far, I have only considered the effect of the duality maps on the graviton and Stückelberg scalar terms, thus it behoves me to briefly discuss their effect on matter fields to which the multi-gravity theory may couple. The possible matter couplings are described in section 2.6, and let us start by minimally coupling matter to just one of the metrics,

$$\mathcal{S}_{\mathrm{matter}}[\Phi_i, \{g_{(i)}\}] = \mathcal{S}_{\mathrm{matter}}[\Phi_i, g^{(1)}_{\mu\nu}], \tag{4.38}$$

chosen to be labelled (1). If the matter metric only couples to one other metric, *i.e.* the matter metric is a vertex of degree one in the theory graph, then the matter action will only be sensitive to a single Stückelberg scalar ϕ. Expanding around a flat background space-time $\eta_{\mu\nu}$ and focusing on the interactions between that single Stückelberg *d.o.f.* and the matter *d.o.f.*, from now on for simplicity taken to be described by a single scalar *d.o.f.* χ (no important feature depends on this assumption, however), one has

$$\mathcal{S}_{\mathrm{matter}} = \int \mathrm{d}^D x \, \mathcal{L}_{\mathrm{matter}}\left[\chi, \partial_\mu\chi, \phi, \partial_\mu\phi, \partial^\nu_\mu\phi\right], \tag{4.39}$$

where we have assumed that there is no dependence on second derivatives or higher of the matter fields (which is straightforward to generalise via (4.7), however). Under the duality transformation $\mathcal{D}_\pi$ this maps to

$$\int \mathrm{d}^D x\, |1+\Pi| \mathcal{L}_{\rm matter}\Big[\chi, [(1+\Pi)^{-1}]_\mu^\nu \partial_\nu \chi, -\left(\pi + \frac{1}{2}(\partial\pi)^2\right), -\partial_\mu \pi, -(1+\Pi)^{-1}\Pi\Big]. \quad (4.40)$$

Unlike for single-field Galileons the form of matter interactions is therefore not invariant under the duality (as was also not the case for some of the bi-Galileon interactions considered above). However, both ways of writing the interaction are of course physically equivalent due to the nature of the duality transformation, and the associated *e.o.m.* will consequently remain secretly second-order in the sense discussed in the previous sections (also see [74]).

Staying with the case where the matter metric only couples to one other metric (carrying a label (2)), as is the case in standard bi- and massive gravity, I now consider the matter coupling in the decoupling limit. Note that in order for this to survive the decoupling limit procedure, one must perform, in addition to (3.38), a scaling of the energy-momentum tensor that keeps $\frac{1}{M}T^{\mu\nu}$ fixed. The field redefinition of $h^{(1)}_{\mu\nu}$ required to remove the lowest order scalar-tensor mixing will induce a matter coupling which survives taking the decoupling limit

$$\mathcal{S}_{\rm matter,dec} = \int \mathrm{d}^D x\, \phi(x) T^\mu_\mu \left[\chi(x), \partial\chi(x)\right], \quad (4.41)$$

where T^μ_μ is the trace of the stress-energy tensor defined with respect to the flat background metric. This is straightforward to see, if one chooses to introduce Stückelberg fields solely through the metric(s) to which matter does not couple, since the only dependence of the matter coupling on helicity-0 modes then comes in through the field redefinition required for scalar-tensor de-mixing. Under the duality transformation $\mathcal{D}_\pi$ this decoupling limit contribution maps to

$$-\int \mathrm{d}^D x\, |1+\Pi| \left(\pi + \frac{1}{2}(\partial\pi)^2\right) T^\alpha_\alpha \left[\chi(x), [(1+\Pi)^{-1}]_\mu^\nu \partial_\nu \chi(x)\right], \quad (4.42)$$

and one notes that, even though the decoupling limit contribution only depends on ϕ and not its derivatives, in the dual picture there is explicit dependence on $\pi, \partial\pi, \partial^2\pi$.

If the metric to which matter couples has direct interactions with N other metrics (*i.e.* it is a vertex of degree N in the theory graph), then expanding around a flat background space-time $\eta_{\mu\nu}$ and focusing on the interactions between the Stückelberg scalars and the matter *d.o.f.* as before, one finds the following schematic action

$$\mathcal{S}_{\text{matter}} = \int \mathrm{d}^D x\, \mathcal{L}_{\text{matter}}\Big[\chi, \partial_\mu \chi, \phi_{(1)}, \partial_\mu \phi_{(1)}, \partial_\mu^\nu \phi_{(1)}, \ldots, \phi_{(N)}, \partial_\mu \phi_{(N)}, \partial_\mu^\nu \phi_{(N)}\Big]. \tag{4.43}$$

The de-mixing procedure for the matter coupling vertex now takes on the following form

$$h^{(1)}_{\mu\nu} \to h^{(1)}_{\mu\nu} + \sum_{i=1}^{N} c_{(1,i+1)} \phi_{(i)} \eta_{\mu\nu}, \tag{4.44}$$

leading to a decoupling limit action

$$\begin{aligned} \mathcal{S}_{\text{matter,dec}} &\sim \int \mathrm{d}^D x \sum_{n=1}^{N} \phi_{(n)} T^\mu_\mu \left[\chi(x), \partial\chi(x)\right] \\ &\xrightarrow{\mathcal{D}_{\pi_{(1)}}} \int \mathrm{d}^D x\, |1+\Pi| \left(- \left(\pi_{(1)} + \frac{1}{2}(\partial\pi_{(1)})^2\right) + \sum_{n=2}^{N} \phi_{(n)} \right) T^\alpha_\alpha \left[\chi, [(1+\Pi_{(1)})^{-1}]^\nu_\mu \partial_\nu \chi\right], \end{aligned} \tag{4.45}$$

In the same fashion one can now iteratively apply the remaining $\mathcal{D}_{\pi(i)}$ as appropriate to fully map all of the $\phi_{(i)}$ into $\pi_{(i)}$. Once again the duality guarantees that even if the resulting *e.o.m.* are higher order, the initial value problem is well-defined in terms of two conditions for each scalar field and hence no extra (Ostrogradsky) *d.o.f.* propagate.

Finally, let me remark on the more complicated matter coupling mentioned in section 2.6, for example

$$\mathcal{S}_{\text{matter}}[\Phi_i, \{g_{(i)}\}] = \mathcal{S}_{\text{matter}}[\Phi_i, \tilde{g}^{\text{matter}}_{\mu\nu}[g_{(1)}, \ldots, g_{(N)}]]. \tag{4.46}$$

When expanding around a flat background space-time $\eta_{\mu\nu}$ and focusing on the interactions between the Stückelberg scalar and the matter *d.o.f.* as before, this takes on the form

$$\mathcal{S}_{\text{matter}} = \int \mathrm{d}^D x\, \mathcal{L}_{\text{matter}}\Big[\chi, \partial_\mu \chi, \phi_{(1)}, \partial_\mu \phi_{(1)}, \partial_\mu^\nu \phi_{(1)}, \ldots, \phi_{(N)}, \partial_\mu \phi_{(N)}, \partial_\mu^\nu \phi_{(N)}\Big]. \tag{4.47}$$

Investigating duality mappings in the decoupling limit for these new matter couplings is left for future work. However, one may reasonably expect that these mappings will be rather different from the cases considered above in one important respect. A primary reason why the matter coupling (4.38) was not invariant under the duality map was that (4.38) explicitly breaks the symmetry between metrics present at the level of the potential self-interactions. Introducing Stückelberg fields via one metric or the other in a bi-gravity-like interaction, while physically equivalent, no longer leaves the form of interactions invariant at the level of the matter action. The many metric coupling (4.46), especially when considering maximally symmetric matter couplings in multi-gravity theories, can restore this symmetry and hence potentially some of the duality invariance as was encountered for single-field Galileons.

4.4 Tri-gravity example

Before concluding this chapter, it is perhaps useful to give an example of the dualities in action in a tri-gravity theory. As at the end of the previous chapter, consider two sites, 1 and 2, which only directly interact with a third site, labelled 3—see figure 3.3.

Overall there are essentially three different[3] ways to introduce Stückelberg fields into each interaction term mapping sites to transform as others (modulo the discussion in section 3.4.1).

1. Map 1 and 2 each to site 3.

$$\begin{tikzcd} 1 \arrow[r, "Y_{(1,3)}"'] & 3 & 2 \arrow[l, "Y_{(2,3)}"] \end{tikzcd}$$

   Let us call the fields that are taken to be 'fundamental' $\pi_{(i,j)}$, and their duals $\phi_{(i,j)}$, and consider the pure scalar interactions which are present in the decoupling limit, after performing scalar-tensor de-mixing. In the first case these are of the form

$$\phi_{(1,3)}\mathcal{L}^{\rm TD}_{(n)}(\phi_{(1,3)}) + \phi_{(2,3)}\mathcal{L}^{\rm TD}_{(n)}(\phi_{(2,3)}) + \left(\pi_{(1,3)} + \pi_{(2,3)}\right)\left[\mathcal{L}^{\rm TD}_{(n)}(\pi_{(1,3)}) + \mathcal{L}^{\rm TD}_{(n)}(\pi_{(2,3)})\right], \tag{4.48}$$

   with the first two terms coming from the field redefinitions of the outer vertices ($h^{(1)}_{\mu\nu}$ and $h^{(2)}_{\mu\nu}$), and the last term from the field redefinition of the inner vertex ($h^{(3)}_{\mu\nu}$). The first two terms are of the same form as arises in bi-gravity and as in (4.8), which were shown to be of Galileon form, whilst the last term is of the form (4.9), $\pi\mathcal{L}^{\rm TD}_{(n)}(\pi)$, which again is manifestly of (bi-)Galileon form. Thus with this choice of Stückelberg fields the helicity-0 interactions are explicitly bi-Galileons.

2. Map 1 to site 3 and 3 to site 2.

[3]Though recall that these are of course all related by gauge transformations and so are all physically equivalent.

$$1 \xrightarrow{Y_{(1,3)}} 3 \xrightarrow{Y_{(3,2)}} 2$$

In this case one has

$$\phi_{(1,3)}\mathcal{L}^{\mathrm{TD}}_{(n)}(\phi_{(1,3)}) + \pi_{(3,2)}\mathcal{L}^{\mathrm{TD}}_{(n)}(\pi_{(3,2)}) + \left(\pi_{(1,3)} + \phi_{(3,2)}\right)\left[\mathcal{L}^{\mathrm{TD}}_{(n)}(\pi_{(1,3)}) + \mathcal{L}^{\mathrm{TD}}_{(n)}(\phi_{(3,2)})\right], \tag{4.49}$$

where now the last term contains pieces which are of the bi-gravity form, as well as a part of the form (4.10), $\pi\mathcal{L}^{\mathrm{TD}}_{(n)}(\phi)$, which was shown in (4.13) to transform into bi-Galileon form; it also contains a part of the form (4.11), $\phi\mathcal{L}^{\mathrm{TD}}_{(n)}(\pi)$—it is not possible to directly show that this is of bi-Galileon form, but one sees here how it is gauge-related to a bi-Galileon term. Also note that due to the diffeomorphism invariance of the action after the Stückelberg fields have been introduced, it is acceptable to apply different Galileon duality maps, which act like a diffeomorphisms, to the separate interaction terms individually.

3. Map 3 to sites 1 and 2.

$$1 \xleftarrow{Y_{(3,1)}} 3 \xrightarrow{Y_{(3,2)}} 2$$

In this case one has

$$\pi_{(3,1)}\mathcal{L}^{\mathrm{TD}}_{(n)}(\pi_{(3,1)}) + \pi_{(3,2)}\mathcal{L}^{\mathrm{TD}}_{(n)}(\pi_{(3,2)}) + \left(\phi_{(3,1)} + \phi_{(3,2)}\right)\left[\mathcal{L}^{\mathrm{TD}}_{(n)}(\phi_{(3,1)}) + \mathcal{L}^{\mathrm{TD}}_{(n)}(\phi_{(3,2)})\right], \tag{4.50}$$

where now the last term contains, in addition to the usual pieces of bi-gravity form, two pieces of the form (4.12), $\phi\mathcal{L}^{\mathrm{TD}}_{(n)}(\phi')$, which can be transformed into terms of the form (4.11), and as above one sees how they are related to bi-Galileon terms via a gauge transformation.

4.5 Conclusions

In this chapter it has been shown how the Galileon structure which arose for the interactions of the helicity-0 mode of the massive graviton in the decoupling limit of massive and bi-gravity, and the associated duality between various Galileon theories makes itself felt in the decoupling limit of multi-gravity theories. The Galileon duality map is essentially a gauge transformation, and thus exists due to the fact that after applying the Stückelberg trick, the action is gauge invariant.

It can be used to show that some of the interaction terms for the helicity-0 modes which are present in the decoupling limit of multi-gravity take the form of multi-Galileons. Other terms do not explicitly take this form, however they can be shown to comprise a system of *e.o.m.* which are of second order. This can be done by applying the duality map at the level of the *e.o.m.* (rather than the action), and performing appropriate manipulations either before or after doing so; it can also be understood by the fact that the apparently higher order system is related via a gauge transformation to an explicitly second order one.

The theories considered in this chapter all only contained interaction terms which directly involve two fields, and it would be interesting to explore the decoupling limit of the intrinsically multi-gravity interaction terms mentioned in section 3.4, which directly involve more than two fields. It is tempting to conjecture that their decoupling limit will again contain multi-Galileon terms for the helicity-0 modes, and in particular that an interaction term which directly involves more than three fields will yield multi-Galileons which involve more than two Stückelberg scalar fields. Investigation of this is left to future work.

It has also been shown how these dualities affect the matter action, when matter is minimally coupled to one metric. As in the case of some of the terms mentioned above, one is generally not left with an action whose *e.o.m.* are manifestly second order, although, since the duality map is invertible, no extra *d.o.f.* are introduced by it.

Understanding the decoupling limit of multi-gravity aids one in investigating the physics of these theories and their (low energy) interactions. An obvious task for the future will be to complete the work started here and derive the full decoupling limit interactions for multi-Gravity (*i.e.* for all ghost-free such models and for all helicity *d.o.f.*) and to use this to understand the cosmological phenomenology and relevance of such theories. The dualities uncovered in the process are also of interest from a purely field-theoretic point of view, establishing the equivalence of seemingly unrelated field theories via invertible, non-local field re-definitions that would have been very hard to discover in another way. Realising that these dualities exist can be extremely useful in investigating multi-scalar field models, *e.g.* for cases where the duality relates strongly and weakly coupled theories and hence makes the physics in seemingly strongly-coupled regimes calculable (this is also the case for the single Galileon duality). It also alerts one to the existence of a rich space of higher-derivative theories, which are nevertheless healthy and free of any ghost-like degrees of freedom. Investigating theories of interacting spin-2 fields appears to be an especially fruitful avenue to discovering healthy theories of this kind.

CHAPTER

5

Cycles of Interactions

In this chapter I investigate the consequences of theory graphs which are no longer trees but instead contain cycles, such as in figure 2.2. The possible importance of cycles in the theory graph when it comes to the ghost-freedom of a theory of multiple interacting spin-2 fields was briefly discussed earlier, in section 2.5.3, and in the literature was first mentioned in [19]. In that paper the authors note that the equivalence between the vielbein and metric formulations of multi-gravity breaks down in the presence of a cycle, and whilst demonstrating the health of the vielbein theory go on to conjecture that the metric version will contain a Boulware-Deser ghost. The authors of [99] (see also [92]) then showed how the standard constraint analysis, which is used to prove the ghost-freedom of multi-metric theories with a tree-graph structure, breaks down in the presence of a cycle, again suggesting the presence of a ghost. For a related analysis in 3D see [152].

The work in this chapter, which is based on the paper [79], confirms this suspicion, by explicitly demonstrating in the Stückelberg formulation the presence of higher derivative terms which will lead to a ghost. It is structured in the following way: section 5.1

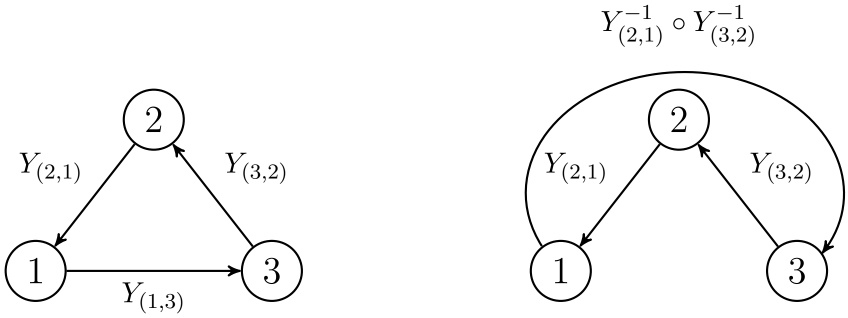


Figure 5.1: Left: introducing a Stückelberg field for every link the case of a cycle leads to an overall constraint. Right: the constraint eliminates one Stückelberg field, replacing it with a plaquette formed from the other fields in the cycle.

investigates a crucial way in which theories with cycles of interactions differ from purely tree-like interactions. That this difference will lead to ghosts in the metric version of the theory is then demonstrated in two different ways in section 5.2; in section 5.3 I review the decoupling limit of a bi-vielbein theory, before investigating the structure of interactions in the presence of a cycle, and argue why the same ghost is not present there. Finally section 5.4 discusses in more detail the link with dimensional deconstruction, before concluding in section 5.5. In this chapter, and the next, I will specialise to $D = 4$, though none of the conclusions essentially depend on this, and it is simple to generalise to arbitrary dimensions.

5.1 Plaquettes

As noted in section 2.5.3, one key difference between theory graphs with a cycle and those without (tree graphs) is that in the presence of a cycle there are now more links than broken copies of diffeomorphism invariance (since the diagonal subgroup remains unbroken). Hence if one introduces a Stückelberg field for every link as in the tree case, then one will end up with a set of fields which are not in fact independent, but satisfy some constraint. For example in the case of a tri-gravity cycle as depicted in figure 5.1, the Stückelberg fields satisfy

$$Y_{(1,3)} \circ Y_{(3,2)} \circ Y_{(2,1)} = \text{id}. \tag{5.1}$$

One way of dealing with this is to use the constraint to re-express one Stückelberg field in the cycle in terms of the others, *i.e.* in the tri-gravity case

$$Y_{(1,3)} = Y^{-1}_{(2,1)} \circ Y^{-1}_{(3,2)}. \tag{5.2}$$

Following [47] such a construction is called a *plaquette*. Now site 1 is being mapped to site 3 all the way around the cycle. This has the advantage that all the fields are now explicitly independent, however it does break the symmetry of the cycle by forcing one to pick a Stückelberg field to eliminate, as well as introducing interactions between *all* the remaining Stückelberg fields through the plaquette. Thus one may wonder whether this is truly necessary, and in appendix B I show that if one introduces more Stückelberg fields than broken copies of diff invariance and treats them all independently, one encounters fields which are infinitely strongly coupled, hindering the analysis.

5.1.1 Plaquettes beyond the linear level

The condition (5.2) yields for the Stückelberg scalars at linear order: $\pi_{(1,3)} = -(\pi_{(2,1)} + \pi_{(3,2)})$, however at higher order we can no longer look at the scalars and vectors separately. In fact, writing

$$Y^{\mu}_{(1,3)}(x) = x^{\mu} + A^{\mu}_{(1,3)} + \partial^{\mu}\pi_{(1,3)}, \tag{5.3}$$

one has

$$Y^{\mu}_{(1,3)}(x) = Y^{-1\ \mu}_{(3,2)}\left(Y^{-1}_{(2,1)}(x)\right) = x^{\mu} + \tilde{Z}^{\mu}_{(2,1)} + \tilde{Z}^{\mu}_{(3,2)} + \sum_{n=1}^{\infty} \frac{1}{n!} \tilde{Z}^{\nu_1}_{(2,1)} \dots \tilde{Z}^{\nu_n}_{(2,1)} \tilde{Z}^{\mu}_{(3,2),\nu_1\dots\nu_n}, \tag{5.4}$$

where a comma denotes partial differentiation, and $Y^{-1\mu}(x) = x^{\mu} + \tilde{Z}^{\mu} = x^{\mu} + B^{\mu} + \partial^{\mu}\phi$, and B^{μ} and ϕ are the dual fields associated with the Stückelberg vector A^{μ} and scalar π. This was already discussed in detail for the scalar in the previous chapter, and the duality can be extended to include the Stückelberg vector as well, in which case

$$x + B + \partial\phi = x + \tilde{Z} = Y^{-1} = (x+Z)^{-1} = (x + A + \partial\pi)^{-1}. \tag{5.5}$$

Solving this, and disentangling the vector and scalar parts one finds

$$\begin{aligned} \phi &= \sum_{n=1}^{\infty} \phi_n, \qquad \text{with} \qquad \phi_n = -\sum_{i=1}^{n-1} \frac{1}{i!} Z^{\nu_1} \dots Z^{\nu_i} \partial_{\nu_1\dots\nu_i} \phi_{n-i}, \\ B^{\mu} &= \sum_{n=1}^{\infty} B^{\mu}_n, \qquad \text{with} \qquad B^{\mu}_n = -\sum_{i=1}^{n-1} \frac{1}{i!} Z^{\nu_1} \dots Z^{\nu_i} \partial_{\nu_1\dots\nu_i} B^{\mu}_{n-i}, \end{aligned} \tag{5.6}$$

with the following initial values for the recursion relations:

$$\phi_1 = -\pi, \qquad \phi_2 = \frac{1}{2}\pi^{,\mu}\pi_{,\mu},$$
$$B_1^\mu = -A^\mu, \quad B_2^\mu = Z^\nu\partial_\nu A^\mu + A^\nu\partial_\nu^\mu\pi. \tag{5.7}$$

Using these one can rewrite (5.4) as

$$x^\mu + \partial^\mu\left(\phi_{(2,1)} + \phi_{(3,2)} + \sum_{n=1}^{\infty}\frac{1}{n!}\tilde{Z}^{\nu_1}_{(2,1)}\dots\tilde{Z}^{\nu_n}_{(2,1)}\phi_{(3,2),\nu_1\dots\nu_n}\right)$$
$$+ B^\mu_{(2,1)} + \sum_{n=0}^{\infty}\frac{1}{n!}\tilde{Z}^{\nu_1}_{(2,1)}\dots\tilde{Z}^{\nu_n}_{(2,1)}\left(B^\mu_{(3,2),\nu_1\dots\nu_n} - \tilde{Z}^{\lambda,\mu}_{(2,1)}\phi_{(3,2),\lambda\nu_1\dots\nu_n}\right), \tag{5.8}$$

from which the expressions for $A^\mu_{(1,3)}$ and $\pi_{(1,3)}$ can be read. And so one sees that through $\tilde{Z}$ each receives contributions from both the vectors and the scalars; in particular even if $A^\mu_{(2,1)}$ and $A^\mu_{(3,2)}$ are set to zero one still has $A^\mu_{(1,3)} \neq 0$. For example

$$\pi_{(1,3)} = -\left(\pi_{(2,1)} + \pi_{(3,2)}\right) + \frac{1}{2}(\pi_{(2,1)} + \pi_{(3,2)})^{,\mu}(\pi_{(2,1)} + \pi_{(3,2)})_{,\mu} + A^\mu_{(2,1)}\pi_{(3,2),\mu} + \dots \tag{5.9}$$
$$\begin{aligned} A^\mu_{(1,3)} =& - A^\mu_{(2,1)} - A^\mu_{(3,2)} + A^\nu_{(2,1)}A^\mu_{(2,1),\nu} + A^\nu_{(3,2)}A^\mu_{(3,2),\nu} + A^\nu_{(2,1)}A^\mu_{(3,2),\nu} \\ &+ (\pi_{(2,1)} + \pi_{(3,2)})^{,\nu}(A^\mu_{(2,1)} + A^\mu_{(3,2)})_{,\nu} - \pi_{(3,2),\nu}(A^{\mu,\nu}_{(2,1)} + A^{\nu,\mu}_{(2,1)}) \\ &+ A^\nu_{(2,1)}\pi^{,\mu}_{(2,1),\nu} + A^\nu_{(3,2)}\pi^{,\mu}_{(3,2),\nu} - \pi^{,\mu\nu}_{(2,1)}\pi_{(3,2),\nu} + \dots . \end{aligned} \tag{5.10}$$

The final term of (5.10) will turn out to have important consequences.

It is also worth mentioning at this point that there will be introduced quadratic mixing between the vectors through the kinetic term for the plaquette vector:

$$\partial_{[\mu}A_{(1,3)\nu]}\partial^{[\mu}A^{\nu]}_{(1,3)} \supset \partial_{[\mu}A_{(2,1)\nu]}\partial^{[\mu}A^{\nu]}_{(2,1)} + 2\partial_{[\mu}A_{(2,1)\nu]}\partial^{[\mu}A^{\nu]}_{(3,2)} + \partial_{[\mu}A_{(3,2)\nu]}\partial^{[\mu}A^{\nu]}_{(3,2)}. \tag{5.11}$$

This is a qualitatively new feature, as in the absence of a plaquette only the tensors and scalars will be mixed quadratically (and of course in removing the scalar-tensor mixing one leaves the scalars mixed).

5.2 Ghosts in multi-metric theories

Now I shall show how the cycle leads to the introduction of a ghost at an energy scale *below* Λ_3, the cutoff of the effective theory in the absense of a cycle (or equivalently one can think of this as leading to a lowering of the cutoff). For simplicity a tri-metric cycle is considered, but in section 5.4.1 I will consider larger cycles in the context of deconstructing dimensions; also all of the interactions strengths and Planck masses are set equal. For related work on tri-metric cycle theories see [99, 153]

The key point is that it is the $(2,1)$ and $(3,2)$ fields which are canonically normalised: $A^{\mu}_{(2,1)} \to \frac{1}{\Lambda_2^2} A^{\mu}_{(2,1)}$, $\pi_{(2,1)} \to \frac{1}{\Lambda_3^3}\pi_{(2,1)}$, and similarly for $(3,2)$. Thus the $(1,3)$ fields do not have the overall normalisation one would expect. In fact

$$\pi_{(1,3)} \to \sum_{n=0,m=1}^{\infty} \frac{1}{\Lambda_2^{2n}\Lambda_3^{3m}} A^n \pi^m, \tag{5.12}$$

$$A^{\mu}_{(1,3)} \to \sum_{n=1,m=0}^{\infty} \frac{1}{\Lambda_2^{2n}\Lambda_3^{3m}} A^n \pi^m + \sum_{n=0,m=2}^{\infty} \frac{1}{\Lambda_2^{2n}\Lambda_3^{3m}} A^n \pi^m. \tag{5.13}$$

For the scalar this is not an issue, since $\Lambda_2 > \Lambda_3$, and so $\Lambda_2^{2n}\Lambda_3^{3m} \geq \Lambda_3^{2n+3m}$, and thus any terms from the plaquette will sit at or above Λ_3. On the other hand $\Lambda_2^{2n}\Lambda_3^{3m} < \Lambda_2^{2n+3m}$ for $m > 0$, and so these terms from the vector will come in below Λ_3 (since $A^{\mu} \sim \frac{1}{\Lambda_2}$ is what is required for the terms involving the vector to sit precisely at Λ_3).

More precisely, recalling that the interaction Lagrangian has an overall pre-factor m^2M^2, one has

$$\begin{aligned} m^2M^2\partial_{[\mu}A_{(1,3)\nu]}\partial^{[\mu}A^{\nu]}_{(1,3)} \supset\ & \frac{1}{\Lambda_4^4}\pi_{(2,1),\lambda[\mu}\pi^{,\lambda}_{(3,2),\nu]}\left(\partial^{[\mu}A^{\nu]}_{(2,1)} + \partial^{[\mu}A^{\nu]}_{(3,2)}\right) \\ & - \frac{1}{\Lambda_4^8}\pi_{(2,1),\lambda[\mu}\pi^{,\lambda}_{(3,2),\nu]}\pi^{,\rho[\mu}_{(2,1)}\pi^{,\nu]}_{(3,2),\rho}, \end{aligned} \tag{5.14}$$

which involves higher derivatives, but *not* of such a form as discussed in chapter 4 which eliminates higher order equations of motion. Thus this theory contains a ghost associated with an energy scale $\Lambda_4 < \Lambda_3$.

There will be (an infinite number of) other, potentially dangerous, terms at energy scales between Λ_4 and Λ_3, however it is just the lowest energy scale which concerns us here, since this gives the new cutoff of the theory. Also one need not worry that this is just an artefact of some sort of truncation since (5.14) are the *only* terms at Λ_4.

Finally note that the first term in (5.14) involves the Stückelberg vector *linearly*—a qualitatively new feature, which means that it cannot classically be set to zero and ignored as it can in the absence of a cycle.

5.2.1 Without plaquettes

The presence of these dangerous terms can also be demonstrated using a different method, in which one does not introduce $Y_{(1,3)}$ in the first place (and hence does not introduce a plaquette). This necessitates a slightly different approach to introducing the Stückelberg fields, as explained in section 3.4.1: one treats the action as a whole, picking one site onto which one maps all of the other fields. For a tri-metric cycle this means

$$\begin{aligned}\mathcal{S}_{\text{int}} &= \mathcal{S}[g_{(1)}, g_{(2)}] + \mathcal{S}[g_{(2)}, g_{(3)}] + \mathcal{S}[g_{(3)}, g_{(1)}] \\ &\to \mathcal{S}[g_{(1)} \circ Y_{(1,2)}, g_{(2)}] + \mathcal{S}[g_{(2)}, g_{(3)} \circ Y_{(3,2)}] + \mathcal{S}[g_{(3)} \circ Y_{(3,2)}, g_{(1)} \circ Y_{(1,2)}]; \end{aligned} \tag{5.15}$$

note the final term, which is different to all those considered previously, as it involves Stückelberg fields applied to both the metrics involved.

It turns out that for the pure scalar part of the action coming from this term one finds (*e.g.* for an interaction term consisting of just the first symmetric polynomial)

$$\begin{aligned}&M\left[3\mathcal{L}^{\text{TD}}_{(1)}\left(\pi_{(3,2)}\right) + \mathcal{L}^{\text{TD}}_{(1)}\left(\pi_{(1,2)}\right)\right] \\ &+ \frac{1}{m^2}\left[2\mathcal{L}^{\text{TD}}_{(2)}\left(\pi_{(3,2)}\right) + \mathcal{L}^{\text{TD}}_{(1,1)}\left(\pi_{(3,2)}, \pi_{(1,2)}\right)\right] \\ &+ \frac{1}{\Lambda_5^5}\left[\mathcal{L}^{\text{TD}}_{(3)}\left(\pi_{(3,2)}\right) + \mathcal{L}^{\text{TD}}_{(1,2)}\left(\pi_{(3,2)}, \pi_{(1,2)}\right)\right] \\ &+ \frac{1}{\Lambda_4^8}\left[\mathcal{L}^{\text{TD}}_{(1,3)}\left(\pi_{(3,2)}, \pi_{(1,2)}\right) + \frac{1}{4}\pi_{(1,2),\lambda[\mu}\pi_{(3,2),\nu]}^{,\lambda}\pi_{(1,2)}^{,\rho[\mu}\pi_{(3,2),\rho}^{,\nu]}\right] + \cdots \end{aligned} \tag{5.16}$$

where $\mathcal{L}^{\text{TD}}_{(n,l)}(\pi, \phi)$ is the total derivative combination of n copies of $\partial^2\pi$ and l of $\partial^2\phi$, defined in (4.19). One sees that this takes the expected, safe form, *viz.* a total derivative, at quadratic and cubic order, but at quartic order a new type of term appears which is precisely the same[1] as that in (5.14) suppressed by Λ_4^8. Similarly the vector-scalar-scalar terms will consist of total derivatives along with the term from (5.14) suppressed by Λ_4^4.

Of course this is to be expected, as the different ways of introducing the Stückelberg fields are all equivalent; in fact, performing a gauge transformation on the final term in (5.15) (each term is gauge invariant, so they can be treated individually) with parameter $Y^{-1}_{(3,2)}$, and noting that $Y_{(1,2)} = Y^{-1}_{(2,1)}$ one has

$$\mathcal{S}\left[g_{(3)}, g_{(1)} \circ \left(Y^{-1}_{(2,1)} \circ Y^{-1}_{(3,2)}\right)\right], \tag{5.17}$$

[1] Recall that $Y_{(1,2)} = Y^{-1}_{(2,1)}$, and so $\pi_{(1,2)} = -\pi_{(2,1)} + \ldots$

which is identical to using a plaquette.

5.3 Absence of ghost in multi-vielbein theories

Below I show how the dangerous terms which arise in multi-metric theories do not do so in the case of multi-vielbein theories, which is to be expected since such theories have been shown to be ghost-free even when cycles are present in the theory graph [19] (though see [94, 97] for some questions about a hole in that proof). It is useful first to recapitulate how the analysis of the decoupling limit proceeds for a bi-vierbein theory.

5.3.1 Bi-vierbein decoupling Limit

Just as in the metric version one can then perturb about a flat background for the vierbeine,

$$E^a_\mu = \delta^a_\mu + \frac{1}{2M} h^a_\mu, \qquad F^a_\mu = \delta^a_\mu + \frac{1}{2M} l^a_\mu, \tag{5.18}$$

and about the identity for the Stückelberg fields,

$$\partial_\mu Y^\nu = \delta^\nu_\mu + \frac{1}{mM} \partial_\mu A^\nu + \Pi^\nu_\mu, \qquad \Lambda^a{}_b = e^{\frac{1}{mM} \omega^a{}_b}, \tag{5.19}$$

where the fields have already been canonically normalised, and $\Pi^\mu_\nu = \frac{1}{m^2 M} \pi_{,\nu}^{,\mu}$; for simplicity I have already included the necessary dimensional factors. The decoupling limit is then taken in the usual way:

$$M \to \infty, \qquad m \to 0, \qquad \Lambda_3 = \left(m^2 M\right)^{1/3} \quad \text{fixed.} \tag{5.20}$$

The normalisation of ω may seem arbitrary, since it has no kinetic term, however due to its antisymmetry, ω will only couple to $\partial_\mu A^\nu$ at leading order, and hence it must have the same scaling as that in order to survive the decoupling limit without generating any divergent terms.

Since it scales in the same way as $\partial_\mu A^\nu$, we know that no terms involving both ω and a helicity-2 field will survive the decoupling limit. Therefore the helicity-2/0 part of the action will be of exactly the same form as in the metric version. For the helicity-1/0 part one finds

$$\mathcal{S}_{\rm int} = \frac{m^2 M_{\rm Pl}^2}{2}\sum_{n=0}^{4}\frac{\beta_n}{n!(4-n)!}\int \epsilon_{a_1\cdots a_{4-n}b_1\cdots b_n}E^{a_1}\wedge\cdots\wedge E^{a_{4-n}}\wedge F^{b-1}\wedge\cdots\wedge F^{b_n} \tag{5.21}$$

$$\begin{aligned}\to \mathcal{S}_{1/0} = &-\frac{1}{4}\sum_{n=0}^{4}\beta_n\int \mathrm{d}^4x\,\Big\{\left(\omega(1+\Pi)\right)_{\mu\nu}X^{\nu\mu}_{(1,n-2)}\left(G,1+\Pi\right)+(\omega G)_{\mu\nu}X^{\nu\mu}_{(n-1)}(1+\Pi)\\ &-\left(\omega^2(1+\Pi)\right)_{\mu\nu}X^{\nu\mu}_{(n-1)}\left(1+\Pi\right)-\left(\omega(1+\Pi)\right)_{\mu\nu}X^{\nu\mu}_{(1,n-2)}\left(\omega(1+\Pi),1+\Pi\right)\Big\}\\ =&-\frac{1}{4}\sum_{n=0}^{4}\beta_n\int \mathrm{d}^4x\,\omega_{\mu\nu}X^{\mu\nu}_{(1,n-1)}\left(G-\omega(1+\Pi),1+\Pi\right)\end{aligned} \tag{5.22}$$

where $G_{\mu\nu}=2\partial_{[\mu}A_{\nu]}$. The Lorentz Stückelberg field is an auxiliary field, and its equation of motion is

$$\sum_{n=1}^{4}\beta_n\,X^{[\mu\nu]}_{(1,n-1)}\left(G-2\omega(1+\Pi),1+\Pi\right)=0. \tag{5.23}$$

Since $1+\Pi$ is symmetric, this reduces to[2]

$$G_{\mu\nu}=2(\omega_{\mu\nu}+\omega_{[\mu|\lambda}\Pi^{\lambda}_{\ |\nu]}). \tag{5.24}$$

As a matrix equation this is the Lyapunov equation, which has solution

$$\omega_{\mu\nu}=\int_0^\infty \mathrm{d}u\,e^{-2u}e^{-u\Pi^\rho_\mu}G_{\rho\lambda}e^{-u\Pi^\lambda_\nu}=\sum_{i=0}^{\infty}\sum_{j=0}^{i}\frac{(-1)^i}{2^{1+i}}{}^iC_j\left(\Pi^{i-j}G\Pi^j\right)_{\mu\nu}, \tag{5.25}$$

and upon substitution of this into (5.22), and a little algebra, one finds

$$\begin{aligned}\mathcal{S}_{1/0}=&\frac{1}{16}\int \mathrm{d}^4x\,\sum_{n=0}^{3}(\alpha_{n+2}-\alpha_{n+1})\Big\{G_{\mu\nu}\left(X^{\mu\nu}_{(1,n)}(G,\Pi)-X^{\mu\nu}_{(1,n)}(G\Pi,\Pi)\right)\\ &-\sum_{i=2}^{\infty}\sum_{j=0}^{i}\sum_{k=0}^{i-j}\sum_{l=0}^{j}\frac{(-1)^i}{2^i}{}^{i-j}C_k\left({}^jC_l-2^{j-1}C_{l-1}\right)\left(\Pi^{i-j-k}G\Pi^k\right)_{\mu\nu}X^{\mu\nu}_{(1,n)}\left(\Pi^{j-l}G\Pi^l,\Pi\right)\Big\}.\end{aligned} \tag{5.26}$$

The equivalent calculation in the metric version is slightly more involved but one can confirm that it nonetheless yields the same result, as it should given the equivalence, for bi-gravity, demonstrated in section 2.4.2. In particular, one sees that the kinetic term for A_μ is $-\frac{1}{16}(\alpha_2-\alpha_1)G_{\mu\nu}G^{\mu\nu}$, in agreement with (3.14).

[2] This form could also be derived by considering (2.45) in the decoupling limit.

5.3.2 A simple cycle

Let us now see what the vierbein version of a tri-metric cycle looks like. For simplicity I will specialise to interaction terms which only have β_1 non-zero, however the result is completely general. The fact that not all three Lorentz Stückelberg fields are independent means that from their *e.o.m.* one now has the two equations

$$|E_{(1)}|E^a_{(1),[\mu|}\eta_{ab}(\Lambda_{(2,1)}E_{(2)})^b_{|\nu]} - |E_{(2)}|E^a_{(2),[\mu|}\eta_{ab}(\Lambda_{(3,2)}E_{(3)})^b_{|\nu]} = 0, \tag{5.27}$$

$$|E_{(2)}|E^a_{(2),[\mu|}\eta_{ab}(\Lambda_{(3,2)}E_{(3)})^b_{|\nu]} - |E_{(3)}|E^a_{(3),[\mu|}\eta_{ab}(\Lambda_{(1,3)}E_{(1)})^b_{|\nu]} = 0, \tag{5.28}$$

which are clearly not equivalent to the DvN symmetric vierbein condition (2.45) being enforced for each pair of vierbeine, and thus one no longer has direct equivalence with the metric version. The dangerous terms found in section 5.2 did not involve the helicity-2 mode, so let us focus on the helicity-1/0 part:

$$\begin{aligned}-4\mathcal{L}_{1/0} =& \sum_{i=1}^{2}\left[\omega_{(i+1,i)\mu\nu}G^{\mu\nu}_{(i+1,i)} + \left(1+\Pi_{(i+1,i)}\right)^{\mu}_{\nu}\omega^{\nu}_{(i+1,i)\rho}\omega^{\rho}_{(i+1,i)\mu}\right] \\ &+ m^2M^2\left(\omega_{(1,3)\mu\nu}G^{\mu\nu}_{(1,3)} + (1+\Pi_{(1,3)})^{\mu}_{\nu}\omega^{\nu}_{(1,3)\rho}\omega^{\rho}_{(1,3)\mu}\right) + \mathcal{O}(\omega^3).\end{aligned} \tag{5.29}$$

As in the metric case, the plaquette (5.4) can be used to replace the diff Stückelberg fields, which I will write as

$$A^{\mu}_{(1,3)} = \sum_{n=2}^{\infty}\frac{1}{\Lambda_3^{3n}}a^{\mu}_{n} + \frac{1}{\Lambda_2^2}\sum_{n=0}^{\infty}\frac{1}{\Lambda_3^{3n}}b^{\mu}_{n} + \mathcal{O}\left(\frac{1}{\Lambda_2^4}\right), \tag{5.30}$$

$$\pi_{(1,3)} = \sum_{n=1}^{\infty}\frac{1}{\Lambda_3^{3n}}\sigma_n + \mathcal{O}\left(\frac{1}{\Lambda_2^2}\right). \tag{5.31}$$

Similarly the Lorentz Stückelberg field $\omega_{(1,3)}$ can be related to the others via

$$e^{\omega_{(1,3)}} = \Lambda_{(1,3)} = \Lambda^{-1}_{(2,1)}\Lambda^{-1}_{(3,2)} = e^{-\omega_{(2,1)}}e^{-\omega_{(3,2)}} = e^{-\frac{1}{\Lambda_2^2}(\omega_{(2,1)}+\omega_{(3,2)}) + \frac{1}{2\Lambda_2^4}[\omega_{(2,1)},\omega_{(3,2)}] - \cdots}, \tag{5.32}$$

where, as in the acyclic case, $\omega_{(2,1)}$ and $\omega_{(3,2)}$ have been normalised by Λ_2^2. (5.29) then becomes

$$\begin{aligned}-4\mathcal{L}_{1/0} =& \frac{1}{2}\omega_{+\mu\nu}\left\{G^{\mu\nu}_{+} - 4\partial^{[\mu}(b+\Lambda_2^2 a)^{\nu]}\right\} + \frac{1}{2}\omega_{-\mu\nu}G^{\mu\nu}_{-} \\ &+ \left(1+\partial^{\mu}_{\nu}\sigma + \partial^{\mu}a_{\nu}\right)\omega^{\nu}_{+\rho}\omega^{\rho}_{+\mu} + \frac{1}{4}\Big\{(2+\Pi_{+})^{\mu}_{\nu}\left(\omega^{\nu}_{+\rho}\omega^{\rho}_{+\mu} + \omega^{\nu}_{-\rho}\omega^{\rho}_{-\mu}\right) \\ &+ (\Pi^{\mu\nu}_{-} + \partial^{[\mu}a^{\nu]})\omega_{+\nu\rho}\omega^{\rho}_{-\mu} + (\Pi^{\mu\nu}_{-} - \partial^{[\mu}a^{\nu]})\omega_{-\nu\rho}\omega^{\rho}_{+\mu}\Big\} + \mathcal{O}\left(\frac{1}{\Lambda_2^2}\right),\end{aligned} \tag{5.33}$$

where $\omega_{\pm} = \omega_{(3,2)} \pm \omega_{(2,1)}$, *etc.* The terms shown are those which naïvely would survive the decoupling limit holding Λ_3 constant (*which I do not yet take*). One derives the following equations of motion for the Lorentz Stückelberg fields:

$$\left(G_+ - 4\partial(b + \Lambda_2^2 a) + \omega_+(6 + \Pi_+ + 4\partial^2\sigma + 2(\partial a + (\partial a)^{\mathrm{T}})) - \omega_-(\Pi_- + (\partial a - (\partial a)^{\mathrm{T}}))\right)^{[\mu\nu]} + \mathcal{O}\left(\frac{1}{\Lambda_2^2}\right) = 0, \tag{5.34}$$

$$\left(G_- + \omega_-(2 + \Pi_+) - \omega_+(\Pi_- - (\partial a - (\partial a)^{\mathrm{T}}))\right)^{[\mu\nu]} + \mathcal{O}\left(\frac{1}{\Lambda_2^2}\right) = 0, \tag{5.35}$$

and can attempt to solve them via an expansion in powers of Λ_2 and Λ_3. Doing so one finds that the leading terms are

$$\omega_+^{\mu\nu} = \frac{4}{3}\frac{\Lambda_2^2}{\Lambda_3^6}\partial^{[\mu} a_2^{\nu]} + \ldots, \qquad \text{and} \qquad \omega_-^{\mu\nu} = \frac{4}{3}\frac{\Lambda_2^2}{\Lambda_3^6}\Pi^{[\mu}_{-\lambda}\partial^{\lambda} a_2^{\nu]} + \ldots. \tag{5.36}$$

But immediately a problem is apparent: with these solutions, terms in (5.33) which have been ignored in fact will contribute at a level equivalent to those that have been kept. Or in other words, ω should not be normalised by Λ_2, but by Λ_3, and so if one wants to take the decoupling limit keeping Λ_3 fixed, one must include terms with arbitrary powers of ω. Whilst this does not mean that taking such a decoupling limit is impossible, it certainly complicates matters, to the extent that unfortunately I am unable to explicitly show the absence of the ghost in this way.

5.3.2.1 Without a plaquette

One can of course analyse the tri-metric cycle in the same manner as section 5.2.1—mapping everything to one site. The parts of the interaction Lagrangian involving just one Stückelberg field, *i.e.* $\mathcal{L}[E_{(1)} \circ Y_{(1,2)}, E_{(2)}] + \mathcal{L}[E_{(2)}, E_{(3)} \circ Y_{(3,2)}]$, will have standard forms and so one just needs to consider the part involving two Stückelberg fields:

$$\mathcal{L}[E_{(3)} \circ Y_{(3,2)}, E_{(1)} \circ Y_{(1,2)}] = -\frac{m^2 M^2}{2}\frac{1}{3!}\delta^{\mu\nu\rho\sigma}_{abcd}(\Lambda_{(3,2)}E_{(3)}\partial Y_{(3,2)})^a_\mu(\Lambda_{(3,2)}E_{(3)}\partial Y_{(3,2)})^b_\nu(\Lambda_{(3,2)}E_{(3)}\partial Y_{(3,2)})^c_\rho(\Lambda_{(1,2)}E_{(1)}\partial Y_{(1,2)})^d_\sigma. \tag{5.37}$$

Expanding around a flat background and normalising the fields in the usual way this becomes

$$
\begin{aligned}
-4\mathcal{L} = &\Lambda_3^3\left(\frac{1}{3}h_{(1)\mu\nu}\tilde{X}^{\mu\nu}_{(0,3)} + h_{(3)\mu\nu}\tilde{X}^{\mu\nu}_{(1,2)}\right)\\
&+\frac{1}{3}\left(\omega_{(1,2)\mu\lambda}\partial_\nu A^\lambda_{(1,2)} + \frac{1}{2}\left(1+\Pi_{(1,2)}\right)^\lambda_\mu \omega_{(1,2)\lambda\rho}\omega^\rho_{(1,2)\nu}\right)\tilde{X}^{\mu\nu}_{(0,3)}\\
&+\left(\omega_{(3,2)\mu\lambda}\partial_\nu A^\lambda_{(3,2)} + \frac{1}{2}\left(1+\Pi_{(3,2)}\right)^\lambda_\mu \omega_{(3,2)\lambda\rho}\omega^\rho_{(3,2)\nu}\right)\tilde{X}^{\mu\nu}_{(1,2)}\\
&+(\omega_{(1,2)\mu\nu}+\partial_\nu A_{(1,2)\mu})(\omega_{(1,2)\rho\sigma}+\partial_\sigma A_{(1,2)\rho})\tilde{X}^{\mu\nu\rho\sigma}_{(0,2)}\\
&+2(\omega_{(3,2)\mu\nu}+\partial_\nu A_{(3,2)\mu})(\omega_{(3,2)\rho\sigma}+\partial_\sigma A_{(3,2)\rho})\tilde{X}^{\mu\nu\rho\sigma}_{(1,1)}\\
&+\Lambda_2^2\left(\frac{1}{3}(\omega_{(1,2)\mu\nu}+\partial_\nu A_{(1,2)\mu})\tilde{X}^{\mu\nu}_{(0,3)} + (\omega_{(3,2)\mu\nu}+\partial_\nu A_{(3,2)\mu})\tilde{X}^{\mu\nu}_{(1,2)}\right)\\
&+\Lambda_2^2\left(\frac{1}{3}\omega_{(1,2)\mu\nu}\Pi^\nu_{(1,2)\lambda}\tilde{X}^{\mu\lambda}_{(0,3)} + \omega_{(3,2)\mu\nu}\Pi^\nu_{(3,2)\lambda}\tilde{X}^{\mu\lambda}_{(1,2)}\right) + \mathcal{O}\left(\frac{1}{\Lambda_2^2}\right),
\end{aligned}
\tag{5.38}
$$

where $\tilde{X}^{\mu\nu\cdots\rho\sigma}_{(n,m)} = \eta^{\nu\tilde{\nu}}\cdots\eta^{\sigma\tilde{\sigma}}\delta^{\mu\cdots\rho\alpha_1\cdots\alpha_n\gamma_1\cdots\gamma_m}_{\tilde{\nu}\cdots\tilde{\sigma}\beta_1\cdots\beta_n\delta_1\cdots\delta_m}(1+\Pi_{(1,2)})^{\beta_1}_{\alpha_1}\cdots(1+\Pi_{(3,2)})^{\delta_1}_{\gamma_1}\cdots$. The terms on the final two lines would lower the cutoff since they are suppressed by a scale below Λ_3; one sees that those in the penultimate line do not contribute because $\omega_{\mu\nu}\tilde{X}^{\mu\nu}=0$ since $\tilde{X}$ is symmetric whereas ω is antisymmetric, and $\partial_\nu A_\mu \tilde{X}^{\mu\nu} = \partial_\nu(A_\mu\tilde{X}^{\mu\nu})$ since $\partial_\mu\tilde{X}^{\mu\nu}=0$; the terms in the final line however do not disappear and, as will be shown below, one arrives at a similar conclusion to the previous section.

Including the contributions from the other links, leads to

$$
\begin{aligned}
\mathcal{L}\supset &-\omega_{(1,2)\mu\nu}\left(\left(3\tilde{X}^{\mu\lambda}_{(2,0)}+\frac{1}{3}\tilde{X}^{\mu\lambda}_{(0,3)}\right)G^\nu_{(1,2)\lambda} + \left(8\tilde{X}^{\mu\nu\lambda\rho}_{(1,0)}+2\tilde{X}^{\mu\nu\lambda\rho}_{(0,2)}\right)G_{(1,2)\lambda\rho} + \frac{\Lambda_2^2}{3}\Pi^\nu_{(1,2)\lambda}\tilde{X}^{\mu\lambda}_{(0,3)}\right)\\
&-\omega_{(3,2)\mu\nu}\left(\left(\tilde{X}^{\mu\lambda}_{(1,2)}-\eta^{\mu\lambda}\right)G^\nu_{(3,2)\lambda} + 4\tilde{X}^{\mu\nu\lambda\rho}_{(1,1)}G_{(3,2)\lambda\rho} + \Lambda_2^2\Pi^\nu_{(3,2)\lambda}\tilde{X}^{\mu\lambda}_{(1,2)}\right)\\
&+\omega_{(1,2)\mu\nu}\omega_{(1,2)\lambda\rho}\left(\left(\frac{3}{2}\tilde{X}^{\mu\sigma}_{(2,0)}+\frac{1}{6}\tilde{X}^{\mu\sigma}_{(0,3)}\right)(1+\Pi_{(1,2)})^\rho_\sigma\eta^{\nu\lambda} + 4\tilde{X}^{\mu\nu\lambda\rho}_{(1,0)} + \tilde{X}^{\mu\nu\lambda\rho}_{(0,2)}\right)\\
&+\omega_{(3,2)\mu\nu}\omega_{(3,2)\lambda\rho}\left(\left(\frac{1}{2}\tilde{X}^{\mu\sigma}_{(1,2)}(1+\Pi_{(3,2)})^\rho_\sigma + \Pi^{\mu\rho}_{(3,2)} - \eta^{\mu\rho}\right)\eta^{\nu\lambda} + 2\tilde{X}^{\mu\nu\lambda\rho}_{(1,1)}\right) + \mathcal{O}\left(\frac{1}{\Lambda_2^2}\right),
\end{aligned}
\tag{5.39}
$$

from which are derived the following equations of motion for the Lorentz Stückelberg fields:

$$
T^{\mu\nu}_i = \omega^{\mu\nu}_{(i,2)} - 2\left(\omega^{[\mu}_{(i,2)\lambda}C^{\nu]\lambda}_i + A^{\lambda[\mu}_i\omega_{(i,2)\lambda\rho}B^{\nu]\rho}_i\right) + \mathcal{O}\left(\frac{1}{\Lambda_2^2}\right), \tag{5.40}
$$

where $i = 1, 3$ and T_i, A_i, B_i, and C_i are given in appendix C. Compared to (5.24), the equivalent for a single link, equation (5.40) is more complicated and for completeness its solution neglecting the terms which are naïvely suppressed by Λ_2^2 is given in appendix C. For now I only need look at the terms suppressed by the lowest scale, for which are found

$$\omega^{\mu\nu}_{(1,2)} = 2\frac{\Lambda_2^2}{\Lambda_3^6}\Pi^{[\mu}_{(1,2)\lambda}\Pi^{\nu]\lambda}_{(3,2)} + \dots, \qquad \omega^{\mu\nu}_{(3,2)} = -6\frac{\Lambda_2^2}{\Lambda_3^6}\Pi^{[\mu}_{(1,2)\lambda}\Pi^{\nu]\lambda}_{(3,2)} + \dots, \tag{5.41}$$

where Λ_3^{-3} has been explicitly extracted from each Π. These exhibit the same scaling as (5.36) in the previous section. The conclusion is thus the same: in order to consistently take the decoupling limit holding Λ_3 fixed one must consider terms with an arbitrary number of ω's.

In the absence of an explicit re-summation of the ω-dependent contributions I cannot prove ghost-freedom in this way, but the fact that this is different from the metric version, and the results of [19] (though see [94, 97] for some questions about a hole in that proof) inspire confidence in the ghost-freedom of the vielbein version.

5.4 Cycles and deconstructing dimensions

Dimensional deconstruction [46–50] is the idea that a theory placed on a discrete, periodic extra dimension is in some sense equivalent to the truncation of the infinite tower of modes which arises from a standard KK reduction on S^1. In this way it allows one to consider whether a low energy effective theory can be derived from the compactification of a higher dimensional theory.

Such a discrete, periodic extra dimension can clearly be represented as a cycle theory graph [48, 49], as in figure 2.2, and thus analysis of the dimensional deconstruction paradigm requires analysis of theory graphs containing cycles. In particular, as we will want to take the $N \to \infty$ naïve continuum limit,[3] it is necessary to consider larger cycles than in the previous sections.

5.4.1 Larger plaquettes

Whilst in the tri-metric case it is possible to avoid the use of plaquettes, simplifying matters slightly, for larger cycles the use of plaquettes (or plaquette-like constructions) is unavoidable, since now not every site is one link removed from every other site. Thus

[3] Precisely what is meant by this will be discussed later.

I now look at plaquettes of larger size.

In the case of a cycle of N metrics the plaquette expression (5.2) is extended in the obvious way and the equivalent of the final term in (5.10) is

$$A^{\mu}_{(1,N)} \supset -\sum_{i=1}^{N-2}\sum_{j=i+1}^{N-1} \pi^{,\mu}_{(i+1,i),\lambda}\pi^{,\lambda}_{(j+1,j)}. \tag{5.42}$$

The analysis proceeds in the same way as in the tri-metric case, leading to ghost-inducing terms of the same form as (5.14), except that now there are $\frac{1}{2}(N-1)(N-2)$ vector-scalar-scalar terms[4] and $\frac{1}{8}(N-1)(N-2)(N^2-3N+4)$ tetra-scalar terms, all suppressed by Λ_4.

The scalar and vector modes are each mixed at the quadratic level and the kinetic (and mass, in the case of the scalar) terms must be diagonalised in order to find the propagating modes; doing so will then introduce an N dependence to the previously $\mathcal{O}(1)$ coefficients in front of the Λ_4-suppressed terms, which then means that the actual cutoff can in fact be much lower.

As described in section 3.4.2, to find the propagating modes one must diagonalise the 'kinetic matrix,' K, and then canonically normalise by dividing each mode by the square root of the appropriate eigenvalue of K; an analogous procedure applies for the vectors.[5]

Before introducing the plaquette, the parts of the kinetic terms involving $\pi_{(1,N)}$ are

$$\pi^{,\mu}_{(1,N)}\pi_{(1,N),\mu} - \pi^{,\mu}_{(1,N)}\left(\pi_{(2,1),\mu} + \pi_{(N,N-1),\mu}\right), \tag{5.43}$$

which, upon the plaquette substitution (just taken to lowest order, $\pi_{(1,N)} = -\sum_i \pi_{(i+1,i)}$, since here we are only interested in overall quadratic terms), becomes

$$\sum_{i,j} \pi^{,\mu}_{(i+1,i)}\pi_{(j+1,j),\mu} + \sum_i \pi^{,\mu}_{(i+1,i)}\left(\pi_{(2,1),\mu} + \pi_{(N,N-1),\mu}\right). \tag{5.44}$$

Thus the kinetic matrix takes the form

$$K = \begin{pmatrix} 2 & -1 & \\ -1 & 2 & \ddots \\ & \ddots & \ddots \end{pmatrix} + \begin{pmatrix} 2 & 2 & \cdots \\ 2 & 2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} + \begin{pmatrix} 1 & \cdots & 1 \\ & & \\ 1 & \cdots & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \end{pmatrix}, \tag{5.45}$$

[4]One might expect an additional factor of $N-1$ from $\sum_i \partial A_i$, however as noted in section 5.1.1, the vector modes must be de-mixed and $\sum_i A_i$ will be precisely one of the propagating modes.

[5]For the scalars one must also diagonalise their mass matrix, however as explained in section 6.1.4 this turns out not to affect the scaling with N.

where blank entries are zero and the ellipsis denotes repetition, so the first matrix is tri-diagonal, the last only has non-zero entries in the first and last columns *etc.* The first term is just the kinetic matrix for a path graph of length N (see figure 2.2, and chapter 6), whilst the second term is due to the first term in (5.44), and the last two due to the last two in (5.44). Upon diagonalisation and normalisation (5.42) becomes

$$\sum_{i=1}^{N-2}\sum_{j=i+1}^{N-1}\pi^{;\mu}_{(i+1,i),\lambda}\pi^{;\lambda}_{(j+1,j)} \propto \sum_{n,m=1}^{N-1}\left(\frac{1}{\sqrt{\lambda_n\lambda_m}}\sum_{i=1}^{N-2}\sum_{j=i+1}^{N-1}(v_n)_i(v_m)_j\right)\chi^{;\mu}_{(n),\lambda}\chi^{;\lambda}_{(m)}, \tag{5.46}$$

where v_n is the n-th normalised eigenvector of K, λ_n the corresponding eigenvalue, and χ denotes the propagating modes. Taking this sum to be dominated by terms involving the smallest eigenvalue of K, which decreases to zero as N^{-2},[6] and taking $(v_n)_i \sim N^{-\frac{1}{2}}$, one finds that the largest coefficient in (5.46) scales like

$$N^2\sum_{i=1}^{N-2}\sum_{j=i+1}^{N-1}\left(\frac{1}{\sqrt{N}}\right)^2 \sim N^3. \tag{5.47}$$

Remarkably the validity of these simple arguments is borne out by full numerical analysis of (5.46); one should also diagonalise the mass matrix for the scalar fields as well, however this turns out not to affect the scaling with N.

In the absence of a cycle the vectors are not mixed at quadratic level; this however is changed by the presence of a cycle and introduction of a plaquette:

$$\partial_{[\mu}A_{(1,N)\nu]}\partial^{[\mu}A^{\nu]}_{(1,N)} \to \sum_{i,j}\partial_{[\mu}A_{(i+1,i)\nu]}\partial^{[\mu}A^{\nu]}_{(j+1,j)}. \tag{5.48}$$

Thus the kinetic matrix for the vectors takes the form

$$K = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & \end{pmatrix} + \begin{pmatrix} 1 & 1 & \cdots \\ 1 & 1 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \tag{5.49}$$

and one sees that $\sum_{i=1}^{N-1} A_{(i+1,i)}$, which is precisely the combination appearing in the $\frac{1}{\Lambda_4^4}\partial A(\partial^2\pi)^2$ terms, is an eigenvector, with eigenvalue N.

Therefore schematically one finds, for the terms with the largest coefficients,

$$\frac{1}{\Lambda_4^4}\partial A(\partial^2\pi)^2 \sim \frac{N^{\frac{5}{2}}}{\Lambda_4^4}\partial\tilde{A}(\partial^2\chi)^2 \qquad \text{and} \qquad \frac{1}{\Lambda_4^8}(\partial^2\pi)^4 \sim \frac{N^6}{\Lambda_4^8}(\partial^2\chi)^4, \tag{5.50}$$

[6]This can be shown numerically, or using the techniques of chapter 6.

where $\tilde{A}$ are the propagating vector modes which diagonalise (5.49). Thus the cutoff behaves as $\Lambda \sim \Lambda_4 N^{-5/8}$ for the first terms and $\Lambda \sim \Lambda_4 N^{-3/4}$ for the latter, which is interesting as the explicit N dependence is marginally stronger than if one looks just at the Λ_3 suppressed terms for which one finds $\Lambda \sim \Lambda_3 N^{-1/2}$ [50, 77, 80], and see chapter 6.

Care must be taken however to precisely define what is meant by the continuum limit, as, if one wants the theory to continue to resemble a KK theory, then the size of the extra dimension must be kept fixed in terms of the four-dimensional Planck mass. In a KK theory the size of the extra dimension is given (roughly) by the inverse of the mass of the lightest state, and so one has

$$R \sim m_1^{-1} \sim N m^{-1}, \tag{5.51}$$

where the mass of the lightest massive graviton for a theory graph C_N can be determined using, for example, the techniques of section 6.2. Meanwhile the four- and five-dimensional Planck masses, along with the Planck mass parameter entering the multi-gravity action are related by [50]

$$M_{\rm Pl}^2 \sim M_5^3 R \sim M^2 m R. \tag{5.52}$$

Requiring that $RM_{\rm Pl}$ remain fixed then leads to $m/M \sim N^{3/2}$, which means that Λ_λ will itself scale with N (relative to $M_{\rm Pl}$), and in particular one has $\Lambda_4/M_{\rm Pl} \sim N^{5/8}$. This finally means that of the two terms in (5.50), the physical cutoff scale due to the first is independent of N, whilst that due to the second *decreases* as $\Lambda \sim N^{-1/8}$.

The fact that the cutoff decreases as the number of sites is increased prevents one from taking the continuum limit of the cycle theory and arriving at Einstein gravity compactified on a circle. In fact it is worth mentioning that this is still not possible if the theory is formulated in terms of vierbeine [50]. In that case the cutoff no longer scales with N (as the increase of $\Lambda_3/M_{\rm Pl} \sim N^{1/2}$ compensates for the explicit scaling $\Lambda_3 N^{-1/2}$), however the requirement that $m/M \sim N^{3/2}$ means that no matter how one chooses their overall scaling, the relative scaling of m and M is counter to that which is required by the decoupling limit. This means that terms suppressed by $\Lambda_{\lambda<3}$ may become relevant again; similarly at some point one reaches a situation in which the heaviest massive graviton ($m_N \sim m$ for C_{N+1}) has a mass which exceeds the Planck mass parameter in the multi-gravity action (and the five-dimensional Planck mass), thereby invalidating the analysis. This means that one can only sensibly increase N to $(M_{\rm Pl}R)^{2/3} = 10^{10}\left(\frac{1\text{TeV}}{1/R}\right)^{2/3}$.

It is worth now making contact with other work that has been done linking dimensional deconstruction and multi-gravity. In [50] it is noted that taking higher dimensional GR and naïvely discretising the metric in the dimension to be compactified will involve interaction terms polynomial in $g^{(i+1)}_{\mu\nu} - g^{(i)}_{\mu\nu}$ (where $g^{(i)}_{\mu\nu}$ is the effective lower dimensional metric at position i in the discretised dimension), which will necessarily introduce a Boulware-Deser ghost [6]. I have now shown that even when one uses interactions which are individually ghost-free, constructing an extra gravitational dimension using metrics will introduce a ghost (in essence this is a 'bottom up' approach).

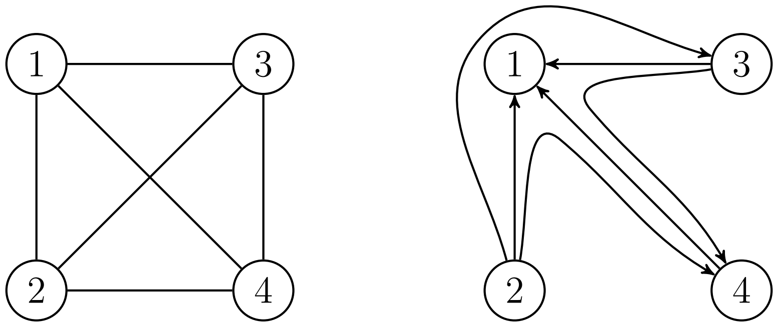
Secondly, in [46–49] it is argued that the truncated KK theory corresponds not to a single cyclic theory graph, but to a *complete* graph,[7] in which the interaction strength for a given link decays as a power law in the distance in the extra dimension between the two sites (and thus the theory is non-local in the extra dimension). Unfortunately this does not seem to help in the multi-metric situation, as there will still always be dangerous terms which do not cancel, and which are suppressed by a finite scale (and thus cannot be ignored); I will briefly show this below.

5.4.2 Complete graphs and plaquettes

There are multiple ways one could introduce plaquettes to a theory of a complete graph, depending on precisely which Stückelberg fields are taken to be independent, and in terms of which the other Stückelberg fields will be represented. Represented in figure 5.2 (for a tetra-metric example) is probably the simplest approach, which is to pick one vertex, and retain as independent all of the Stückelberg fields corresponding to the edges connected to that vertex (thus turning it into a star graph, see figure 2.2)—in this way all vertices are just a distance two from one another, and so each plaquette is of the same simple form as in the tri-gravity example discussed at the start of this chapter.

Consider the terms equivalent to (5.14) arising from a plaquette linking two vertices, i and j, which are separated by a distance $d_{(i,j)}$ in the extra dimension and from the 'host' vertex, H, (the special vertex described in the previous paragraph) by distances

[7] *i.e.* one in which every site is linked to every other



$$Y_{(2,3)} = Y_{(2,1)} \circ Y^{-1}_{(3,1)}$$
$$Y_{(2,4)} = Y_{(2,1)} \circ Y^{-1}_{(4,1)}$$
$$Y_{(3,4)} = Y_{(3,1)} \circ Y^{-1}_{(4,1)}$$

Figure 5.2: Introducing plaquettes for the complete graph of four vertices.

$d_{(i,H)}$ and $d_{(j,H)}$ respectively. Assuming that all Planck masses are equal,[8] schematically one has

$$m^2_{(i,j)}M^2(\partial A_{(i,j)})^2 \supset \frac{m^2_{(i,j)}}{m^2_{(i,H)}m^2_{(j,H)}M}\partial^2\pi_{(i,H)}\partial^2\pi_{(j,H)}\left(\frac{\partial A_{(i,H)}}{m_{(i,H)}} + \frac{\partial A_{(j,H)}}{m_{(j,H)}}\right) + \frac{m^2_{(i,j)}}{m^4_{(i,H)}m^4_{(j,H)}M^2}\left(\partial^2\pi_{(i,H)}\partial^2\pi_{(j,H)}\right)^2 . \tag{5.53}$$

Clearly any other pair of vertices will lead to terms involving a different pair of Stückelberg fields connected to the host vertex, and thus there is no possibility of cancellations. One might hope that by making $m_{(i,j)}$ small when $d_{(i,j)}$ is large (*i.e.* decreasing the interaction strength for large separations in the extra dimension) one can render such terms irrelevant, however it is easy to see that vertices which are simultaneously close together and very distant from the host vertex (*i.e.* $d_{(i,j)} \ll d_{(i,H)}, d_{(j,H)}$) will then lead to dangerous contributions which have a large pre-factor.

5.5 Conclusions

In this chapter a key question concerning the consistency of theories of multiple, interacting spin-2 fields has been answered: does a cycle of interactions, when formulated using metrics, lead to the presence of a ghost (which is not present in the absence of the cycle)? It has been shown that, even when the individual interaction terms are ghost-free, with a strong coupling scale of $\Lambda_3 = (m^2M)^{1/3}$, the cycle introduces higher-derivative terms, suppressed by the lower scale $\Lambda_4 = (m^3M)^{1/4}$, which will inevitably lead to the appearance of a ghost associated with that scale.

[8] This is valid when the extra dimension is not warped.

This was demonstrated in two ways: i) by using a plaquette construction to eliminate the 'extra' Stückelberg field which is introduced due to the number of symmetry breaking interactions being larger than the number of broken symmetries, and ii) by introducing a reduced number of Stückelberg fields at the start. Both methods give the same form for the dangerous terms which appear, confirming the validity of this result. The structure of interactions in the vielbein version of the theory has also been investigated and it was argued why the same ghost does not appear, which it should not, since this version is known to be ghost free [19].

The cycles considered in this chapter were built only from bi-gravity type interaction terms which directly involve two fields, however the results which were found rely only upon the plaquette construction, and not on the form of the interaction terms, and therefore I suspect that they can be extended to cover the multi-gravity type interaction terms mentioned in section 3.4, which directly involve more than two fields. In such an interaction term, every vertex can be considered to be adjacent to every other vertex, which of course suggests a complete graph, however plaquettes are not present since, as this is a single interaction term, Stückelberg fields do not need to be introduced which 'close the cycle.'[9] Therefore simply further connecting two of the vertices in such a term with another interaction term (of either bi- or multi-gravity type) would not require a plaquette, and potentially leaves the theory healthy, whereas connecting two of the vertices with a longer string of interactions, *e.g.* a path graph $P_{N>2}$, would necessitate a plaquette and presumably leads to a ghost at Λ_4 in the metric version.

The results presented in this chapter are interesting not just intrinsically, but also for their relation to dimensional deconstruction; the consequences on the latter of a ghost in the metric version were examined by considering cycles of general size N and it was found that the previously noted problem of a low strong coupling scale which decreases like $N^{1/2}$ [50] is even more pronounced when the ghost is taken into account. More specifically the cutoff of the theory will in this case decrease like $N^{3/4}$, a further impediment to taking the continuum limit (and recovering the full KK theory), at least in the metric version.

[9]For example, in the term depicted on the right in figure 2.1, if Stückelberg fields $Y_{(3,1)}$ and $Y_{(2,1)}$ are introduced, then the action will be gauge invariant, and so unlike the case of the cycle graph C_3, one does not need to introduce a field $Y_{(2,3)}$ which would then be re-expressed using a plaquette.

Further work remains to be done exploring the link between cyclic theories and dimensional deconstruction, and it would be especially interesting to see if and how the ghosts which we have found here are present when one directly truncates the full KK theory. Similarly it would be worthwhile and useful to investigate whether it is possible to find a dimensional deconstruction procedure involving gravity for which a well-defined continuum limit does exist, and if so, also how such a theory would behave from the 'bottom up' perspective of cyclic theory graphs as analysed in this chapter.

CHAPTER

6

The Graph Structure and The Strong-Coupling Scale

In this chapter, which is closely based on [80], I consider to what extent the energy scale at which a multi-gravity theory becomes strongly coupled is influenced and controlled by the structure of its theory graph. In section 6.1 bounds on the strong coupling scale are derived which depend only on properties of the graph. In section 6.2 this is related to the graviton mass spectrum and it is examined to what extent one can arrive at the effective strong-coupling scale by just considering the lightest massive graviton; in order to do this the possible interpretation of general multi-gravity theories in terms of discretisation of higher-dimensional theories is discussed, and I briefly touch upon the coupling of these theories to matter. Throughout this chapter I will work in four spacetime dimensions, though an extension to arbitrary dimensions does not essentially change the conclusions.

6.1 Theory graphs and the strong-coupling scale

Already some structural properties of the theory graph are known to have a bearing on the theory. In particular, as explained in section 2.5.3, if the theory graph contains cycles, then the equivalence between the metric and vielbein versions of the theory breaks down [19], and in the previous chapter it was shown that a cycle causes a ghost to appear at a scale lower than Λ_3 in the metric version of the theory, which lowers the strong coupling scale and the cutoff. Furthermore, the vielbein version of the theory, when examined in the decoupling limit, which makes explicit the strong coupling scale, becomes exceedingly complicated when there is a cycle in the graph [79].

For these reasons, in what follows I will focus on theory graphs which are acyclic (*i.e.* they are tree graphs), and thus it does not matter which version (metric or vielbein) of the theory one uses; for concreteness the metric version will be used.

Also note that, as mentioned in section 2.5.3, to encode the full information about the theory this would have to be a weighted graph, with each edge perhaps carrying a vector which describes the coefficients of the corresponding interaction terms, however in order to focus on the effect of the structure of the graph I will choose these coefficients so that the we can consider simply an unweighted graph.

6.1.1 Relation between the theory graph and the strong coupling scale

As explained in section 3.4.2, the strong coupling scale of a multi-gravity theory (in which all the interaction mass scales are identical) is not simply Λ_3 because one must de-mix and canonically normalise the Stückelberg scalar fields, which leads to a hierarchy in the values of the coefficients in front of the non-linear interaction terms (*i.e.* the Wilson coefficients), and the strong-coupling scale will generically be controlled by the largest of these coefficients. From (3.72) one sees that the eigenvalues of the Stückelberg scalar kinetic matrix for the de-mixed modes in question are a key ingredient in the size of the coefficient. Thus a first step in finding the strong coupling scale is to find the smallest eigenvalue of the kinetic matrix.

In order to see how this is related to the theory graph let us return to the tri-metric

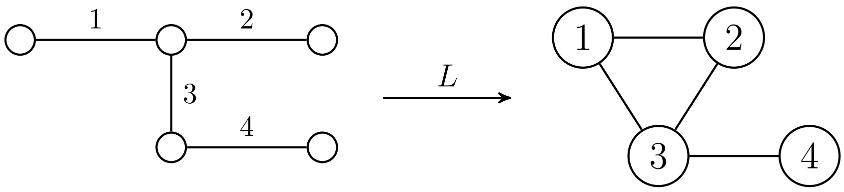


Figure 6.1: The operation of the line graph operator, L.

theory considered in section 3.4.2, whose theory graph is simply the path graph on three vertices,[1] P_3, as shown in figure 3.3. If one chooses the direction of the Stückelberg mapping so that $\sigma = 1$, then one sees from (3.67) that (in matrix notation) $K = 2I_2 + A(P_2)$, where $A(G)$ is the *adjacency matrix* of a graph G, which is defined as

$$A(G)_{ij} = \begin{cases} 1 & \text{if vertices } i \text{ and } j \text{ are joined by an edge,} \\ 0 & \text{else;} \end{cases} \tag{6.1}$$

also note that P_2 is the *line graph* of P_3, where the line graph, L_G, is the graph whose vertices are the edges of G, and two vertices in L_G are joined by an edge if the corresponding edges in G are connected to the same vertex. This is illustrated by an example in figure 6.1.

The choice $\sigma = 1$ corresponds to choosing the direction of the Stückelberg mapping such that either the two outer vertices, $g^{(2)}$ and $g^{(3)}$, are both mapped to the transformation properties of the inner vertex, $g^{(1)}$, in their respective interaction terms, *or* the inner vertex is mapped to the transformation properties of the two outer vertices in their respective interaction terms. Or in other words, that each vertex in the theory graph either only has fields being mapped to it, or away from it, as shown in figure 3.4.

For a general multi-gravity theory one can consider all Stückelberg scalars in this way, and furthermore for a tree graph one can always orient the edges so that each vertex has only inward or only outward edges (corresponding to $\sigma = 1$; on the other hand, choosing $\sigma = -1$ cannot be consistently extended to graphs with vertices of degree higher than two). Thus for a general theory graph T with N vertices[2] one has

$$K(T) \propto 2I_{N-1} + A(L_T). \tag{6.2}$$

[1] Note that as only acyclic graphs are being considered here, this is the only possibility with three vertices (up to vertex relabelling).

[2] And since this is a tree graph it will have $N - 1$ edges.

Recall that the expression (3.67) for the kinetic matrix derived in section 3.4.2 required the vanishing of the interaction term coefficient α_1, and so this will be a requirement of the analysis in this chapter.[3]

Thus the problem of finding the smallest eigenvalue of the kinetic matrix is the same as finding the smallest eigenvalue of the line graph. As we shall later see, $\lambda_{\min}(K(T))$ is equal to the *algebraic connectivity* of the graph T. The following subsection is devoted to placing bounds on the smallest eigenvalue of the line graph of a tree graph.

6.1.2 Placing bounds on the smallest eigenvalue of $A(L_T)$

Some bounds are already known about the smallest eigenvalue of the line graph of a tree; in particular from [154] we have

$$-2\cos\left(\frac{\pi}{N}\right) \leq \lambda_{\min}(L_T) \leq -1, \tag{6.3}$$

where the lower bound is saturated iff $T = P_N$, and the upper bound iff $T = S_N$—see figure 2.2. There is quite a gulf between these two graphs and their corresponding bounds, and in fact the former results in a strong coupling scale which goes $\sim 1/\sqrt{N}$ whereas the latter in one which does not scale with N. This makes it difficult to draw conclusions about the strong coupling scale for a general tree. Thus I will attempt to find bounds which are tighter, though note that they must necessarily depend on further properties of the graph, as the bounds (6.3) are saturated in certain circumstances.

Incidentally, note that from (6.3) one has $\lambda_{\min} > -2$ (for any finite N) and hence K is positive definite (*cf.* equation (6.2)), which agrees with expectations since we know that in the absence of cycles all the Stückelberg scalar modes have a kinetic term, all of which are of the right sign.[4]

[3] Recall that this is equivalent to the absence at the perturbative level of tadpole terms in the Lagrangian, whose presence at any rate would be equivalent to considering expanding around a different background.

[4] Extending this to graphs which contain cycles one has that, unless G is a tree or contains at least one odd cycle, $\lambda_{\min}(L_G) = -2$ (and $\lambda_{\min}(L_G) \geq -2$ in general) [155], which means that not all of the Stückelberg scalar modes have a kinetic term—leading to all sorts of problems (for details see the previous chapter and appendix B). One might be tempted to conclude that this means that theory graphs containing at least one odd cycle can avoid the problems, but note that the prescription concerning the directions of Stückelberg mapping fails in the presence of an odd cycle and one no longer has $K \propto 2I + A(L_G)$.

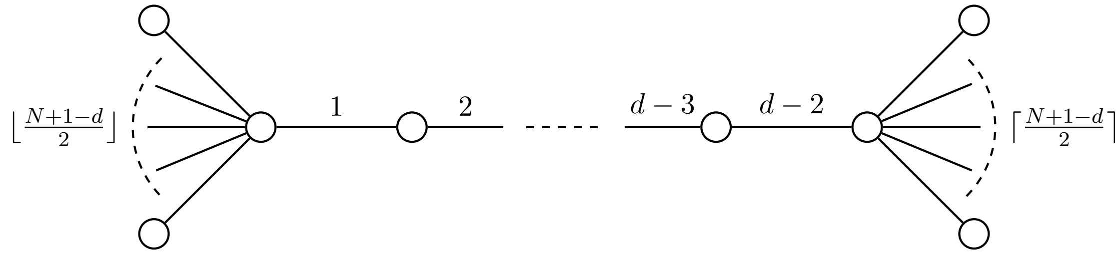


Figure 6.2: The tree which minimises $\lambda_{\min}(L_T)$ for a given number of edges $N-1$ and diameter d. $\lfloor X \rfloor$ denotes the largest integer less than or equal to X and vice versa for $\lceil X \rceil$.

6.1.2.1 Upper bound

Consider adding a vertex (and edge) to T to form a new tree T'. By suitably relabelling vertices one has

$$A(L_{T'}) = \begin{pmatrix} A(L_T) & a \\ a^{\mathrm{T}} & 0 \end{pmatrix}, \tag{6.4}$$

where $a^{\mathrm{T}} = (0, \ldots, 0, \overbrace{1, \ldots, 1}^{m})$, and m is the degree of the vertex in T to which the new vertex is attached. Let v be the eigenvector of $A(L_T)$ with the smallest eigenvalue $\lambda_{\min}(L_T)$, and let $v'^{\mathrm{T}} = (v^{\mathrm{T}}, 0)$, then by Rayleigh's theorem one has

$$\lambda_{\min}(L_{T'}) \leq \frac{v'^{\mathrm{T}} A(L_{T'}) v'}{v'^{\mathrm{T}} v'} = \lambda_{\min}(L_T). \tag{6.5}$$

Thus one sees that as a tree grows, the minimum eigenvalue of its line graph can never increase. As shall be seen later, this means the strong coupling scale of a theory can never be raised by adding new vertices and/or edges in this way.

The *diameter* of a graph, $d(G)$, is defined as the maximum of the set of minimum distances between all pairs of vertices; thus for example $d(P_N) = N - 1$ and $d(S_N) = 2$. Any tree can then be considered to be $P_{d(T)+1}$ with certain 'extrusions' attached to it, and thus one immediately has

$$\lambda_{\min}(L_T) \leq \lambda_{\min}(L_{P_{d(T)+1}}) = -2\cos\left(\frac{\pi}{d(T)+1}\right). \tag{6.6}$$

6.1.2.2 Lower bound

The lower bound of (6.3) can also be improved. Given $N - 1$ edges and a diameter d, the tree whose line graph has the smallest eigenvalue is shown in figure 6.2.

Thus the smallest eigenvalue of the associated line graph will give a lower bound for a given diameter and number of edges. The derivation of the characteristic polynomial of the adjacency matrix of the line graph of figure 6.2 can be found in appendix D, and the result is that the smallest eigenvalue is $-2\cos\theta^*_{N,d}$, where $\theta^*_{N,d}$ is the smallest non-zero root of

$$\begin{aligned} f_{N,d}(\theta) = &\sin((d+1)\theta) + (N-1-d)\sin(d\theta) + \left(\frac{1}{4}(N-3-d)^2 - 1\right)\sin((d-1)\theta) \\ &- \frac{1}{2}(N-1-d)^2\sin((d-2)\theta) + \frac{1}{4}(N-1-d)^2\sin((d-3)\theta). \end{aligned} \tag{6.7}$$

Expanding this about $\theta = 0$ ($\lambda = -2$) one can get a lower bound (valid for $d > 2$):[5]

$$\theta^*_{N,d} = \frac{2}{\sqrt{N(d-2)}}\left[1 + \mathcal{O}\left(\frac{1}{Nd}, \left(\frac{d}{N}\right)^2\right)\right]. \tag{6.8}$$

We now have an improved set of bounds for the smallest eigenvalue of the line graph of any tree of given N, d:[6]

$$-2\cos\theta^*_{N,d} \le \lambda_{\min}(L_T) \le -2\cos\left(\frac{\pi}{d+1}\right) \tag{6.9}$$

where the lower bound is saturated by the graph in figure 6.2. For large N, d this leads to the following bounds on the smallest eigenvalue of the kinetic matrix

$$\frac{4}{N(d-2)} < \lambda_{\min}(K(T)) < \frac{\pi^2}{(d+1)^2}. \tag{6.10}$$

6.1.3 Bounds on the strong coupling scale

To get to the strong coupling scale however requires more than just the eigenvalues of the kinetic matrix, one also needs to deal with the sum in (3.72) involving the initial matrices of coefficients, $C^{A,B}$, and the eigenvectors of K. From table 3.2 one sees that the form of $C^{A,B}$ depends not just on the relative orientation of the links (related to the choice of Stückelberg fields) connected to a given vertex (as for K), but also on the absolute orientation. Also relevant are the coefficients in front of the different interaction terms, and thus to avoid unnecessary complication, and to focus on the effect of the structure of the graph, in the sequel it is demanded that the interaction Lagrangian of each link is of identical form, *with the direction of the links in mind.*

[5] When $d = 2$ the tree must be a star, and one has $\lambda_{\min}(K(S_N)) = 1$.

[6] These bounds do require knowledge of the diameter of the tree, however in general this is much easier to compute than the spectrum ($\mathcal{O}(N)$ vs. $\mathcal{O}(N^3)$).

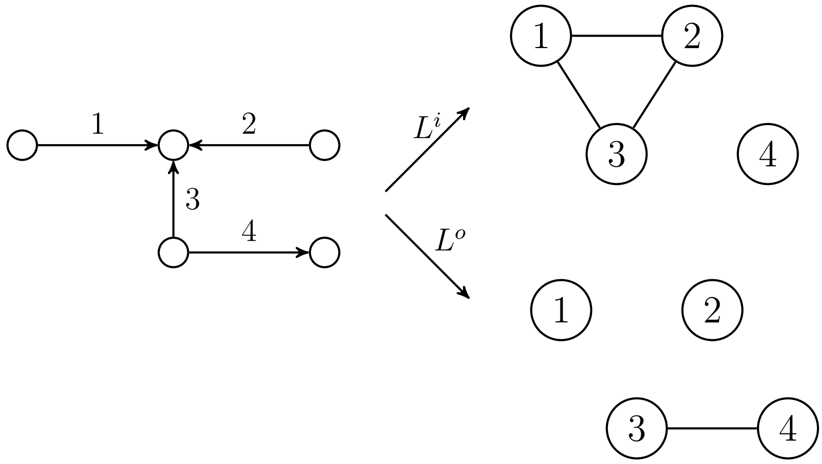


Figure 6.3: The operation of the additional line graph operators, L^i and L^o.

Let me illustrate what is meant by this last point with an example: in the tri-metric theory considered previously, in (3.62), since the directions of the links to site 3 are the same, *i.e.* they are both of the form $\sum_n \beta_n^{(i)} \sqrt{-\left|g_{(3)}\right|}\, e_n \left(\sqrt{g_{(3)}^{-1} g_{(i)}}\right)$, one requires $\beta_n^{(1)} = \beta_n^{(2)}$; if one were to add another site, connected to site 3, the interaction term would have to be $\sum_n \beta_n \sqrt{-\left|g_{(3)}\right|}\, e_n \left(\sqrt{g_{(3)}^{-1} g_{(4)}}\right)$, whereas if it was connected to site 1, the interaction term would have to be $\sum_n \beta_n \sqrt{-\left|g_{(4)}\right|}\, e_n \left(\sqrt{g_{(4)}^{-1} g_{(1)}}\right)$.

In order to write expressions for $C^{A,B}$ for a general tree, it is required to introduce two more line graph operators. For a directed graph, the standard definition of its line graph is that two edges are connected if they share a vertex and one edge is oriented towards the vertex, and the other away from it. I introduce L^i which connects edges if they share a vertex and are both directed *towards* it, and L^o which connects edges if they share a vertex and are both directed *away from* it. This is illustrated by an example in figure 6.3.

With these in hand one can write

$$C^B \propto C^o = I_{N-1} + A(L_G^o), \tag{6.11}$$

$$C^A \propto b_L C^i - b_R C^o, \tag{6.12}$$

and given a theory graph, one can now use these expressions and (3.72) to determine the largest Wilson coefficient, and hence the effective strong coupling scale.

From (3.72) one sees that due to the inverse powers of $\lambda(K(G))$, the largest coefficient will likely come from a term in which two of the fields are that with the minimal value of $\lambda(K(G))$, and so this dependence is isolated by writing

$$\tilde{C}_{\max} = \frac{f(G)}{\lambda_{\min}(K(G))}, \tag{6.13}$$

where

$$f(G) = \lambda_{\min}(K(G)) \max\left(\frac{1}{\sqrt{\lambda_i\lambda_j\lambda_k}}\left|\sum_{lm} U_{li}U_{lj}C_{lm}U_{mk}\right|\right). \tag{6.14}$$

One may wonder why a third factor of $\lambda_{\min}(K(G))^{-1/2}$ is not extracted, and the reason is twofold: whilst $C^{A,B}$ depends on $L_G^{i,o}$, and not just L_G, the similarities are still such that $\sum_m C_{lm}U_{mk} \sim \lambda_k$. This means that $\tilde{C}$ will not in general scale inversely with three powers of $\lambda(K(G))$ and so the maximum coefficient will not generally have *three* fields with minimal λ; secondly the form (6.13) will prove to have a nice interpretation in terms of graviton mass eigenstates, as explained in section 6.2.

Having placed bounds on $\lambda_{\min}(K(G))$ it remains to examine the behaviour of $f(G)$. Unfortunately, for given diameter, it is more difficult to determine which graphs extremise this, and in fact the dependence of those extreme values on the diameter is much weaker than for $\lambda_{\min}(K(G))$. (The same is true if one considers its dependence on the maximum degree of the graph, for example.)

Nonetheless one can say something about its average value (taking the prior distribution to be uniform on the set of unlabelled trees of a given number of vertices); I find

$$\langle f(T)\rangle \approx \frac{1}{\sqrt{d-1}}. \tag{6.15}$$

Thus, given (6.10) one has for $d > 2$ (note that $d = 2$ must be star graph, for which $\Lambda = \Lambda_3$)

$$1.5\left(\frac{\sqrt{d-1}}{N(d-2)}\right)^{1/3} \lesssim \frac{\Lambda}{\Lambda_3} \lesssim 2.1\left(\frac{\sqrt{d-1}}{(d+1)^2}\right)^{1/3}, \tag{6.16}$$

or for large d:

$$\frac{1.5}{\sqrt{d}}\left(\frac{d}{N}\right)^{1/3} \lesssim \frac{\Lambda}{\Lambda_3} \lesssim \frac{2.1}{\sqrt{d}}. \tag{6.17}$$

Although these bounds are not strict (since only the average value of $f(T)$ has been used) they are fairly inclusive. Taking the set of all (non-isomorphic) trees with 15 (resp. 20) vertices, only 2.5% (resp. 2.0%) exceed the upper bound, and 0.4% (resp. 0.5%) fall below the lower bound.

6.1.3.1 Higher order interaction terms

Having focussed only on the cubic scalar interactions I now briefly consider higher order (in the fields) interactions and explain why one would not expect them to lead to a lower effective cutoff than for cubic terms.

An extra scalar has to come in with two extra derivatives (because of the multi-gravity Λ_3 decoupling limit structure), and hence an $(n+2)$-th order term would look like

$$\frac{1}{\Lambda_3^{3n}}\left(\frac{1}{\sqrt{\lambda_{i_1}\cdots\lambda_{i_{n+2}}}}\sum_{lm}U_{li_{n+2}}\cdots U_{li_2}C_{lm}U_{mi_1}\right)\mathcal{L}(\chi_{(i_1)},\cdots,\chi_{(i_{n+2})}). \tag{6.18}$$

As before extract $n+1$ factors of $\sqrt{\lambda_{\rm min}}$, and writing the sum that is left as f_n one has that the effective cutoff is

$$\frac{\Lambda}{\Lambda_3}=\frac{\lambda_{\rm min}^{\frac{n+1}{6n}}}{f_n^{\frac{1}{3n}}}. \tag{6.19}$$

Since $\lambda_{\rm min}$ is raised to an increasingly small power, f_n would have to increase with n to compensate (*e.g.* for a path graph, $\Lambda\sim 1/\sqrt{N}$ requires $f_n\sim N^{\frac{n}{2}-1}$). This seems unlikely, especially as naïve power counting in N would suggest opposite behaviour, and so it is reasonable to conclude that the cubic order terms give the lowest effective cutoff.

6.1.4 Mass diagonalisation

Thus far I have ignored the fact that the Stückelberg scalars are mixed not only in their kinetic terms, but also in their mass terms, and that really, to find the propagating modes one must also find the mass eigenstates. In practice it is found that not including mass diagonalisation does not greatly affect the value of the strong coupling scale, and in this section I briefly investigate why this is.

Like their kinetic terms, the Stückelberg scalars do not initially have mass terms, but gain them from the spin-2 mass terms when the field redefinition $\sim h\to h-\pi\eta$ is applied. It turns out that their mass terms mix Stückelberg scalars which are *two* links apart, rather than just adjacent, and so the simplest example is a tetra-metric path theory, P_4, since this requires three Stückelberg scalars. Taking the interaction terms to be the same, as described in the previous section, the spin-2 mass terms can be written,

$$h^{\mathrm{T}}_{\mu\nu}\begin{pmatrix} m^2_{RR} & m^2_{RL} & 0 & 0 \\ m^2_{RL} & 2m^2_{LL} & m^2_{RL} & 0 \\ 0 & m^2_{RL} & 2m^2_{RR} & m^2_{RL} \\ 0 & 0 & m^2_{RL} & m^2_{LL} \end{pmatrix} h_{\alpha\beta}M^{\mu\nu\alpha\beta}, \tag{6.20}$$

where $M^{\mu\nu\alpha\beta}$ contains the information about the tensor structure of these terms and would be $\eta^{\mu\nu\alpha\beta}$ for Fierz-Pauli structure, whereas the other matrix of constant coefficients details the 'flavour' structure. Once the appropriate field redefinitions to remove the quadratic scalar-tensor mixing are applied, the scalar mass terms look like

$$\mathcal{L}_{\rho^2} \propto \rho^{\mathrm{T}} \begin{pmatrix} m^2_{RR} + 2m^2_{LL} - 2m^2_{RL} & 2(m^2_{LL} - m^2_{RL}) & -m^2_{RL} \\ 2(m^2_{LL} - m^2_{RL}) & 2(m^2_{RR} + m^2_{LL} - m^2_{RL}) & 2(m^2_{RR} - m^2_{RL}) \\ -m^2_{RL} & 2(m^2_{RR} - m^2_{RL}) & 2m^2_{RR} + m^2_{LL} - 2m^2_{RL} \end{pmatrix} \rho. \tag{6.21}$$

If one has $m^2_{RR} = m^2_{LL} = -m^2_{RL}$—which is the case if the coefficients of the interaction terms obey $\alpha_0 = \alpha_1 = 0$[7]—then the mass matrix in (6.21) becomes proportional to

$$\begin{pmatrix} 5 & 4 & 1 \\ 4 & 6 & 4 \\ 1 & 4 & 5 \end{pmatrix} = 4 + 4A(L_{P_4}) + A(L_{P_4})^2, \tag{6.22}$$

where one recalls that A is the adjacency matrix. It is easy to see that this property generalises to other trees, *i.e.*

$$\mathcal{L}_{\rho^2} \propto \rho^{\mathrm{T}} \left(4 + 4A(L_T) + A(L_T)^2\right) \rho. \tag{6.23}$$

Now since the eigenvectors of $K(T)$ are the eigenvectors of $A(L_T)$, and hence of $A(L_T)^2$, the mass diagonalisation becomes trivial in this case, being already performed by the kinetic diagonalisation.

[7] Recall that this is required for the expression for the kinetic matrix (6.2) to be valid, and is equivalent to the absence at the perturbative level of constant and tadpole terms in the Lagrangian, whose presence at any rate would be equivalent to considering expanding around a different background.

6.2 Relation to graviton masses

In the case of a path graph, one can write the strong coupling scale as [50]

$$\Lambda_{\text{eff}} = \left(m_1^2 M_{\text{Pl}}\right)^{1/3}, \tag{6.24}$$

where m_1 is the mass of the lightest massive graviton state, and M_{Pl} is the physical Planck mass, which are respectively related to the mass parameters m and M appearing in the action. Intuitively this makes sense, since if there is a gap between the lightest massive graviton and the next, then at energies below the mass of the next to lightest graviton the theory is effectively bi-gravity and so the strong coupling scale should be Λ_3 built out of mass parameters m_1 and M_{Pl}. In this section I investigate to what extent this is true in general.

First it is necessary to find the graviton mass spectrum for a generic theory; from (6.20) one can see that if $m_{LL}^2 = m_{RR}^2 = -m_{RL}^2 = m^2$, extending to a general graph the mass matrix satisfies

$$M(G) = m^2 L(G) = m^2 \left(D(G) - A(G)\right), \tag{6.25}$$

where $D(G)$ is the diagonal matrix whose ii entry is the degree of vertex i. The combination $L(G)$, of this and the adjacency matrix, is called the *Laplacian* (or *Kirchhoff*) matrix of the graph and its properties have been widely studied.

I note some basic facts about $L(G)$ [156] and their interpretation in terms of multi-gravity theories.

- $L(G)$ has exactly one zero eigenvalue, *i.e.* there is one massless graviton, since there is one unbroken copy of diff.[8]
- The eigenvector corresponding to the zero eigenvalue is $(1, \cdots, 1)^{\mathrm{T}}$, *i.e.* the sum mode $\sum_i h_{\mu\nu}^{(i)}$ is always massless.
- $L(G)$ is positive semi-definite, *i.e.* none of the gravitons are tachyonic.
- The smallest non-zero eigenvalue of $L(G)$ is called the algebraic connectivity, and is

[8]For graphs with multiple separate connected components, there is a corresponding number of zero eigenvalues, and thus a corresponding number of massless gravitons.

larger for graphs which are more "connected," *i.e.* more "connected" theories have a larger mass gap.

6.2.1 Relation to $K(G)$

Given a graph G, and an arbitrary orientation of its edges, the *incidence matrix* of a graph, is the $|E| \times |V|$ matrix with elements

$$B(G)_{ij} = \begin{cases} 1 & \text{if edge } i \textit{ leaves} \text{ vertex } j \\ -1 & \text{if edge } i \textit{ arrives} \text{ at vertex } j \\ 0 & \text{otherwise} \end{cases}, \tag{6.26}$$

and one then has

$$L(G) = B^{\mathrm{T}}B, \qquad \text{and} \qquad K(G) = B_+ B_+^{\mathrm{T}}, \tag{6.27}$$

where $B_{+,ij} = |B_{ij}|$. Now for a tree (or, indeed, any graph which is odd-cycle-free) one can consistently orient the edges such that each vertex has edges either all leaving, or all arriving, in which case one has $B_{ij} = (-1)^{\sigma_j} B_{+,ij}$, where $\sigma_j = 0, 1$ appropriately.

Thus for a tree one has $K(T) = BB^{\mathrm{T}}$, and hence the non-zero eigenvalues of $L(T)$ and $K(T)$ are identical, *i.e.* $\mathrm{spec}(L(T)) = \{0\} \cup \mathrm{spec}(K(T))$, where $\mathrm{spec}(X)$ is the spectrum of a matrix X. In particular this means that $\lambda_{\mathrm{min}}(K(T)) = m_1^2$, and hence from (6.13) one has

$$\Lambda_{\mathrm{eff}}^3 = m_1^2 \frac{M}{f(T)}, \tag{6.28}$$

and (6.24) is recovered if $f(T) = M/M_{\mathrm{Pl}}$. Thus in order to consider whether (6.24) gives an accurate expression for the strong coupling scale one needs an expression for the physical Planck mass to compare with $M/f(G)$; I now consider two approaches to this.

6.2.2 Dimensional deconstruction

A path or cycle graph has a natural interpretation as a discretised extra dimension [46, 48–51, 157], which leads to a relation between M and M_{Pl}. As mentioned previously, one has $M_{\mathrm{Pl}}^2 \sim M_5^3 R$, where M_5 is the five-dimensional Planck mass and $R \sim Nm^{-1}$ the size of the extra dimension; the largest graviton mass $m_N \sim m$ sets the lattice spacing and gives $M^2 = M_5^3 m^{-1}$, leading to $M_{\mathrm{Pl}}^2 \sim NM^2$ [50], which indeed agrees with $M_{\mathrm{Pl}} = M/f(P_N)$.

More complicated graphs do not have so obvious an interpretation in terms of higher dimensional theories, however if one still expects $M_{\mathrm{Pl}}^2 \sim M_5^3 R$ and $M^2 \sim M_5^3 m_N^{-1}$ to hold for general dimensionally reduced theories, one has

$$f(T) = \frac{M}{M_{\mathrm{Pl}}} = \frac{1}{\sqrt{m_N R}}, \tag{6.29}$$

and there are several ways to proceed: from the usual KK expression $R \sim m_1^{-1}$ one has a second relation between R and the graph structure, with which to compare, or alternatively one could take (6.29) as a definition of R, and investigate its behaviour.

As an example, consider the star graph: as already mentioned, $f(S_N) = 1$, and one also has $m_N^2 = Nm^2$, and hence (6.29) gives $R = m^{-1}/\sqrt{N}$. At face value this is a little peculiar: keeping m^{-1}, fixed, as one increases the number of sites, the size of the extra dimension being described actually decreases. It is also in contradiction with $R \sim m_1^{-1} = m^{-1}$. In reality though we probably should not expect (6.24) to give a good expression for the strong-coupling scale in this case, because the graviton mass spectrum is highly degenerate:

$$\mathrm{spec}(L(S_N)) = \{0, 1^{N-2}, N\}. \tag{6.30}$$

i.e. there are $N-2$ gravitons with mass $m_1 = m$. Thus it is no longer true that at low enough energies one can treat the theory as simply bi-gravity, and one has no reason to expect the cutoff to be (6.24), with the attendant interpretation as resulting from a higher dimensional theory. This is not to say that a star graph theory has no place within dimensional deconstruction, merely that its interpretation is a little more tricky.

Denoting by Δ the largest degree of any vertex in T, one has the following bounds [158, 159]

$$\Delta + 1 \leq m_N^2 < \Delta + 2\sqrt{\Delta - 1}, \tag{6.31}$$

and hence from (6.15), $R \sim m^{-1} d/\sqrt{\Delta}$. Intuitively this makes sense since as one increases Δ (resp. d) whilst keeping d (resp. Δ) fixed, it is increasingly similar to $S_{\Delta-1}$ (resp. P_{d+1}). Thus a theory will avoid the counter-intuitive behaviour described above as its theory graph 'grows' if d grows more quickly than does $\sqrt{\Delta}$.

As a further example, consider the 'mace graph,' $M_{d,\Delta}$, which is defined in figure 6.4; this indeed has $R \sim d/\sqrt{\Delta}$. Curiously however, the fact that R decreases when Δ grows

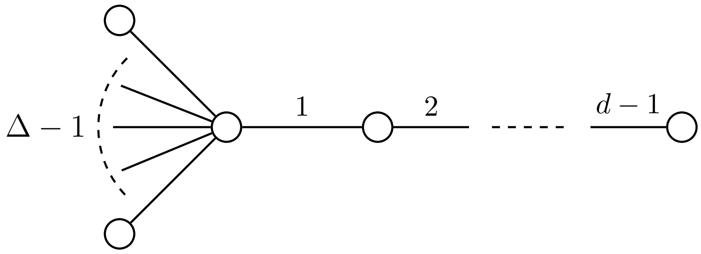


Figure 6.4: The 'mace graph,' $M_{d,\Delta}$.

quickly enough compared to d cannot in this case be blamed on there being more than one graviton at, or close to, mass m_1, as in the case of the star graph. In appendix D it is shown that for $d > 4$,

$$\frac{m_2}{m_1} > \frac{3}{2} \tag{6.32}$$

(and in fact numerically it is found that $m_2 > 2m_1$), *i.e.* the separation between the first two massive gravitons is always of order the mass gap of the theory, which is precisely the case for the path graph theory (which has $m_2 \approx 2m_1$.)

Finally, it is also interesting to briefly comment on a conjecture made in [52]. The authors of that paper were interested in multi-gravity theories in which the mass of the lightest massive graviton is much less than that of the second lightest massive graviton, and whilst a KK theory will usually yield a regular spacing of masses, they conjectured that if the manifold that is being compactified consists of two spaces connected by a thin tube, *i.e.* it is 'pinched,' then the first mass will be much smaller than the next.

It is possible to compare this with how the graviton mass spectrum is related to the properties of the theory graph. Bounds on the lightest mass were already computed in section 6.1.2, and now bounds on the second lightest mass are required, *i.e.* bounds on the second smallest non-zero eigenvalue of the Laplacian matrix of the theory graph. Whilst I am not aware of any general bounds which are helpful in this case, progress can be made by considering a specific class of graphs—those depicted in figure 6.2, which also have the advantage of somewhat resembling a discretisation of the setup proposed in [52].

By considering the large N limit of (6.7), one can derive, for $d > 5$, the following lower bound on it second smallest root: $\theta > \frac{\pi}{d-2}$. Meanwhile for $d \leq 5$ the second smallest, positive eigenvalue of the Laplacian comes not from solving (6.7), but from (D.8) it is simply 1, and thus one has

$$\frac{m_2}{m_1} > \frac{\sqrt{N(d-2)}}{2} \qquad \text{for } d \le 5 \tag{6.33}$$

$$\frac{m_2}{m_1} \gtrsim \frac{\pi}{2}\sqrt{\frac{N}{d-2}} \qquad \text{for } d > 5. \tag{6.34}$$

Therefore so long as N grows more quickly than d, the theory given by the graph in figure 6.2 will exhibit a splitting between the first and second massive gravitons (with $d = 5$ giving the largest splitting). In this circumstance, one would expect (6.24) to give a good expression for the strong coupling scale.

6.2.3 Coupling to matter

Outside of the framework of dimensional deconstruction an expression for the physical Planck mass can be derived by coupling the theory to matter, since this coupling strength is what determines the strength of gravity, and hence M_{Pl}. Considering this question in the detail it deserves is unfortunately beyond the scope of this work, except to make a few elementary observations.

Obviously in general this will depend on how one couples to matter (see section 2.6 for the possible consistent couplings)—for instance if one couples to just one site, then this will clearly depend on at least the graph structure in the vicinity of that vertex.

A more realistic scenario might be that one wants to couple to the (perturbative) massless mode. Regardless of graph structure, the massless mode is $h^{(m=0)}_{\mu\nu} = \sum_i h^{(i)}_{\mu\nu}$; this leads to a kinetic term $NM^2 h^{(m=0)}\mathcal{E}h^{(m=0)}$, and hence canonical normalisation involves $h^{(m=0)} \to \frac{1}{\sqrt{N}M}h^{(m=0)}$. At the non-linear level coupling to this could arise from coupling to an effective vielbein $E^{\text{eff}} = \sum_i E^{(i)}$; linearising, and then canonically normalising one has

$$\frac{1}{\sqrt{N}M}\int \mathrm{d}^4x\, h^{(m=0)}_{\mu\nu}T^{\mu\nu}, \tag{6.35}$$

and thus the coupling strength is $M_{\text{Pl}} = \sqrt{N}M$, *i.e.* the same as in dimensional deconstruction using a path graph. This gives $f = M/M_{\text{Pl}} \sim 1/\sqrt{N}$, which is satisfied for trees with $d \sim N$ (see section 6.1.3). However there are many trees which violate this, and hence violate (6.24); furthermore not all of these are ones which have a degenerate lightest graviton mass (for which one already does not expect (6.24) to hold). This would seem to advise against using (6.24) in general to estimate the strong coupling (at least when one is coupling matter to the massless mode).[9]

6.3 Conclusions

In this chapter I have considered the question of how the strong coupling scale of a multi-gravity theory is related to its topology, *i.e.* to the presence and structure of interactions between the different fields. I have shown that this can elegantly be rephrased in terms of various properties of the corresponding theory graph. In particular bounds, (6.17), have been placed on the strong coupling scale. The parameter to which these bounds are most sensitive is the diameter of the graph. One would expect that tighter bounds could in principle be derived by taking into account further graph properties (of course at the expense of complexity and/or generality). These bounds give us a sense of how the range of validity of the theory behaves as one changes the graph and hence the theory. One important finding is that the strong coupling scale of a multi-gravity theory can never be raised simply by adding extra spin-2 fields and/or interactions between them.

I have also considered when the expression (6.24) ($(m_1^2 M_{\mathrm{Pl}})^{1/3}$) for the strong coupling scale, which one gets by just considering the lightest massive graviton, is valid. The lightest graviton mass turns out to be the algebraic connectivity of the graph, and fits easily with the previously computed bounds. One also has to determine the physical Planck mass, and I have considered doing this via appealing to a higher-dimensional interpretation of the theory, and by considering the coupling of the theory to matter.

The latter approach implies that (6.24) is not universally valid, whilst the former yields curious behaviour in which, for a fixed discretisation scale, the size of the nominal extra dimension decreases as the size of the graph increases. Whilst one would not expect theories which have more than one graviton at mass scale m_1 to have a cutoff described by (6.24), interestingly it also does not hold in some theories which do have a separation between the first two massive gravitons.

It would be useful to further investigate the question of whether and how general

[9] Of course, this relies on starting off with order one parameters, and in particular coupling to $E^{\mathrm{eff}} = \sum_i E^{(i)}$, rather than, say $E^{\mathrm{eff}} = \frac{1}{\sqrt{N}} \sum_i E^{(i)}$—though if the latter were chosen one still would not have a result which would be valid for all graphs.

theory graphs might be related to higher dimensional theories, in the way that path and cycle graph theories can. The results here seem to indicate that the usual approach to interpretation is not appropriate. Finally, I have only considered bi-metric interaction terms, whilst, as mentioned in section 3.4, one may write interaction terms which involve more than two fields. It would be interesting to extend this work to those theories.

CHAPTER

7

Conclusions

This thesis has focussed on various theoretical aspects of theories in which there are multiple, interacting spin-2 fields, all but one of which are massive, and in particular I have considered issues which are specific to multi-gravity, and not present in theories of just two or one spin-2 fields. Given that a single, consistent, non-linear theory of a massive spin-2 field is known to exist, in the form of dRGT massive gravity, the multi-gravity theories studied in this thesis have been based on the interaction terms present in the dRGT theory.

A theory of a massless spin-2 field naturally possesses a gauge symmetry, in the form of diffeomorphism invariance, which is necessarily broken when one makes the field massive, however the analysis presented here has relied on (re)introducing this gauge symmetry to an action involving multiple, interacting spin-2 fields by introducing new fields into the action via the Stückelberg trick. This is a useful thing to do because at energies much higher than the mass of the gravitons (and below the strong coupling scale of the theory), the newly introduced Stückelberg fields behave as the helicity-1 and helicity-0 components of the massive gravitons, which makes the dynamics of those lower helicity components more transparent, and thus makes it easier to ascertain various properties of the theory, such as the scale at which it becomes strongly coupled.

I have considered in detail the structure of the interactions of the helicity-0 modes in the decoupling limit—given that the decoupling limit of dRGT massive gravity contains interactions which belong to a specific class of scalar field theories, the Galileons, and that these have a multi-field generalisation, one might expect that the multi-field generalisation of massive gravity would then contain multi-Galileons in its decoupling limit. Indeed, some of the interactions are manifestly of this form, however some are not, and would naïvely seem to give *e.o.m.* which are higher than second order, however the gauge invariance of the theory (after applying the Stückelberg trick), and duality between different Galileon theories that this implies, was crucial in showing that the resulting interactions do nonetheless only generate second order *e.o.m.*

Many questions one might ask about a given multi-gravity theory can be rephrased in terms of properties of the so-called theory graph, and when the theory contains more than two fields (*i.e.* it is truly multi- and not bi-gravity) the possibilities for this theory graph become numerous. In particular, whether the graph contains cycles turns out to be key. When there are cycles in the theory graph, if one applies the Stückelberg trick in the same way as for a tree graph, then it turns out that not all of the Stückelberg fields are independent, and one must re-express some of them in terms of the others using a plaquette. In a multi-metric theory these plaquettes unavoidably lead to the appearance of a ghost, which lowers the cutoff of the effective theory (compared to what it would be if the graph were acyclic). The case of a multi-vielbein theory is more complicated, as one has to first integrate out the Stückelberg fields which are used to introduce local Lorentz invariance, and whilst the cutoff of the effective theory is not reduced in the same way as in the multi-metric case, it is extremely difficult (and I have not been able) to determine the full structure of the terms present in the decoupling limit.

Theory graphs which contain cycles are especially interesting as they naturally arise in the context of dimensional deconstruction and can be used to describe extra dimensions (which have been discretised). The fact that a multi-metric cycle contains a ghost which lowers the cutoff is a further blow to the prospect of constructing an extra gravitational dimension from the bottom up, and it exacerbates previous problems which have been noted concerning being unable to take the continuum limit of the putative extra dimension.

Even in the absence of cycles, the structure of the theory graph has a large impact on the strong coupling scale of the theory. This is because, when there is more than one set of Stückelberg fields (again, this requires truly multi- and not bi-gravity), they will be mixed at the quadratic level, and one must diagonalise the action to discover the true propagating modes; this procedure then leads to coefficients in front of some of the suppressed terms in the action which are larger than $\mathcal{O}(1)$ and hence the actual suppression scale is lower than that given by just the dimensional factor present. I have shown that this means that in general one cannot raise the cutoff of the effective theory by adding more fields and interactions; on the other hand it is possible to relate the lowered cutoff to structural properties of the theory graph—in particular its diameter.

I have compared this effective strong coupling scale to that which one would calculate if one only considered the lightest massive graviton contained in the theory, along with some physical Planck mass. The difficulty with this is that one also needs somehow to relate the physical Planck mass to the Planck mass entering the gravitational action, and I have considered two possible ways to do this—by imagining the theory graph to describe some discretised, extra dimension, along the lines of dimensional deconstruction, and by appealing to the coupling of matter to the massless graviton mode—however each of these gives a result which cannot be consistent in all cases. And so it seems that multi-gravity theories cannot always be simply reduced to an effective bi-gravity theory.

Finally, given that the title of this thesis contains a question, it is perhaps appropriate for me to venture an answer. Multi-gravity is necessarily more complicated than bi-gravity and massive gravity, and this complexity generates many interesting structures, and affords qualitatively new possibilities. On the other hand, these novel things seem to often make the theory less appealing, be it through the appearance of ghosts in specific circumstances, or the reduction of the range of validity of the effective theory in rather more general circumstances. Therefore, in conclusion, one probably can have too much of a good thing.

APPENDIX

A

Alternative Derivation of (3.24)

In this appendix I will present an alternative derivation to that given in the main text of the mixing (to first order in the tensor field) between $l_{\mu\nu} = f_{\mu\nu} - \eta_{\mu\nu}$ and the scalar Stückelberg field from $F = f \circ Y$, that comes from the $\sqrt{-|g|}\, e_n\left(\sqrt{g^{-1}F}\right)$ interaction term.

Our starting point is the following expression for the l, π mixing, which can be derived by simply substituting the Stückelberg replacement (3.12) into the interaction term:

$$\frac{1}{4}m^2 M_{\rm eff}^2 l_{\mu\nu}(Y)\sum_{i=0}^{n}\left[{}^{D-i}C_{n-i}\eta^{\mu\nu}\mathcal{L}^{\rm TD}_{(i)}(\pi) - {}^{D-i-1}C_{n-i}X^{\mu\nu}_{(i)}(\pi)\right], \tag{A.1}$$

where $l_{\mu\nu}(Y) = l_{\mu\nu}(x + \partial\pi)$ has not been expanded.[1] If one was to expand this and proceed in the same way as for the h, π mixing, one would start encountering terms whose ghost-freedom is hard to establish, due to the issue of non-locality which appears due to the infinite number of derivatives acting on the tensor l. Naïvely one may think that this automatically leads to higher order equations of motion, and hence ghosts via

[1]Throughout this appendix I will drop all contributions from the Stückelberg vector.

Ostrogradsky's theorem. However, the non-local interaction Lagrangian generated in this fashion is degenerate, since there is no highest order derivative term, and one consequently cannot straightforwardly apply Ostrogradsky's theorem here. It is therefore necessary to re-sum terms in the expansion in order to produce a local theory to which Ostrogradsky's theorem can be applied (*cf.* [77, 83, 136, 160]).

The non-locality stems from the expansion of the original tensor field after the introduction of the Stückelberg fields. Recall that the introduction of Stückelberg fields essentially amounts to a co-ordinate transformation $x^\mu \to Y^\mu = x^\mu + \partial^\mu \pi$. Thus, if one re-expresses the rest of the scalar-tensor action in terms of coördinates Y^μ, instead of x^μ, one can remove the non-locality. Let us first analyse the way in which this works for a general N-metric coupling, in which Stückelberg fields are introduced to make everything transform under the symmetry of site 1:

$$\mathcal{S}_{\text{int}} = \int \mathrm{d}^D x\, f(g_{(1)}, g_{(2)}, \ldots, g_{(N)}) \to \int \mathrm{d}^D x\, f(g_{(1)}, g_{(2)} \circ Y_2, \ldots, g_{(N)} \circ Y_N). \tag{A.2}$$

To linear order in the tensor modes $h^{(i)}_{\mu\nu} = g^{(i)}_{\mu\nu} - \eta_{\mu\nu}$, the interactions with the Stückelberg fields take the form

$$\int \mathrm{d}^D x \left(h^{(1)}_{\mu\nu}(x) f_1^{\mu\nu}(\partial Y_2(x), \ldots, \partial Y_N(x)) + \sum_{i=2}^{N} h^{(i)}_{\mu\nu}(Y_i(x)) f_i^{\mu\nu}(\partial Y_2(x), \ldots, \partial Y_N(x)) \right). \tag{A.3}$$

The first term is straightforward to deal with since all the fields explicitly depend on x, so let us focus on one term of the remainder, which can be rewritten

$$\int \mathrm{d}^D x\, h^{(i)}_{\mu\nu}(Y_i(x)) f_i^{\mu\nu}(\partial Y_2(x), \ldots, \partial Y_N(x)) = \int \mathrm{d}^D Y_i\, h^{(i)}_{\mu\nu}(Y_i) F_i^{\mu\nu}(Y_i, x), \tag{A.4}$$

where in the new interaction term one has

$$F_i^{\mu\nu}(Y_i, x) = \left| \frac{\partial x}{\partial Y_i} \right| f_i^{\mu\nu} \left(\frac{\partial Y_2}{\partial Y_i} \frac{\partial Y_i}{\partial x}, \ldots, 1 \frac{\partial Y_i}{\partial x}, \ldots, \frac{\partial Y_N}{\partial Y_i} \frac{\partial Y_i}{\partial x} \right). \tag{A.5}$$

Now if the functional form of f_i is such that the determinental pre-factor on the RHS of (A.5) can eliminate the $\frac{\partial Y_i}{\partial x}$'s appearing within f_i, then the explicit dependence on x is eliminated and one has $F_i^{\mu\nu} = F_i^{\mu\nu}(Y_i)$. This means one can then rename the integration variable on the RHS of (A.4) $Y_i \to x$ and add it back to the rest of the action.

I do not here analyse what forms of f_i have the above property, but demonstrate the precise way in which this works in the context of (A.1). Use of the identities $X^{\mu\nu}_{(n)}(\pi) = \eta^{\mu\nu}\mathcal{L}^{\rm TD}_{(n)}(\pi) - \pi^{\mu}_{\lambda}X^{\lambda\nu}_{(n-1)}$, and ${}^{n}C_k - {}^{n-1}C_{k-1} = {}^{n-1}C_k$, and a little algebra, reveals that this can be re-written as

$$\frac{1}{4}m^2 M^2_{\rm eff} l_{\mu\nu}(Y)\partial_\lambda Y^\nu X^{\lambda\mu}_{(n-1)}(\partial Y), \tag{A.6}$$

where $X^{\mu\nu}_{(n)}(\partial Y)$ is similar to $X^{\mu\nu}_{(n)}(\pi)$ but just with the π^α_β's replaced replaced by $\partial_\beta Y^\alpha$. One can re-express it in the following way

$$\begin{aligned} X^{\mu}_{(n)\nu}(\partial Y) =& \frac{1}{n!(D-n-1)!}\epsilon^{\rho\mu_1\cdots\mu_D}\epsilon_{\nu\nu_1\cdots\nu_D}\frac{\partial Y^\lambda}{\partial x^\rho}\frac{\partial x^\mu}{\partial Y^\lambda}\frac{\partial Y^{\nu_1}}{\partial x^{\mu_1}}\cdots\frac{\partial Y^{\nu_n}}{\partial x^{\mu_n}} \\ &\times\frac{\partial Y^{\lambda_{n+1}}}{\partial x^{\mu_{n+1}}}\frac{\partial x^{\nu_{n+1}}}{\partial Y^{\lambda_{n+1}}}\cdots\frac{\partial Y^{\lambda_D}}{\partial x^{\mu_D}}\frac{\partial x^{\nu_D}}{\partial Y^{\lambda_D}} \\ =& \left|\frac{\partial Y}{\partial x}\right|\frac{1}{n!(D-n-1)!}\frac{\partial x^\mu}{\partial Y^\lambda}\epsilon^{\lambda\nu_1\cdots\nu_n\lambda_{n+1}\lambda_D}\epsilon_{\nu\nu_1\cdots\nu_D}\frac{\partial x^{\nu_{n+1}}}{\partial Y^{\lambda_{n+1}}}\cdots\frac{\partial x^{\nu_D}}{\partial Y^{\lambda_D}} \\ =& \left|\frac{\partial Y}{\partial x}\right|\frac{\partial x^\mu}{\partial Y^\lambda}X_{(D-n-1)\nu}{}^{\lambda}\left(\frac{\partial x}{\partial Y}\right), \end{aligned} \tag{A.7}$$

and thus one sees that (A.6) is a function purely of Y; upon renaming $Y \to x$, one has $\frac{\partial x}{\partial Y} \to \frac{\partial Y^{-1}}{\partial x}$, and so the l, π mixing can be written as

$$\frac{1}{4}m^2 M^2_{\rm eff} l_{\mu\nu}(x) X^{\mu\nu}_{(D-n)}\left(\partial Y^{-1}\right). \tag{A.8}$$

Finally, writing $\partial_\mu Y^{-1\nu} = \delta^\nu_\mu + \phi^\nu_\mu$, and expanding $X^{\mu\nu}_{(D-n)}(\partial Y^{-1})$, one arrives at (3.24):

$$\frac{1}{4}m^2 M^2_{\rm eff} l_{\mu\nu}\sum_{i=0}^{D-n}{}^{D-i-1}C_{D-n-i}X^{\mu\nu}_{(i)}(\phi). \tag{A.9}$$

APPENDIX

B

Necessity of Plaquettes

In this appendix the consequences of not eliminating one of the Stückelberg fields via construction of a plaquette will be investigated; for concreteness and simplicity I will work in $D = 4$, with a tri-metric theory in which all the interaction terms are identical, and all Planck masses and interaction mass scales are equal.

After introducing the Stückelberg fields and expanding about a flat background

$$g^{(i)}_{\mu\nu} = \eta_{\mu\nu} + h^{(i)}_{\mu\nu}, \qquad Y^{\mu}_{(i+1,i)} = x^{\mu} + A^{\mu}_{(i+1,i)} + \partial^{\mu}\pi_{(i+1,i)}, \tag{B.1}$$

where the indices in the labels are understood to be taken modulo 3, the scalar-tensor interaction terms which will survive in the decoupling limit are (*cf.* (3.22) and (3.24))

$$\mathcal{L}_{h\partial^2\pi} = \frac{m^2M^2}{8}\sum_{i=1}^{3} h^{(i)}_{\mu\nu}\sum_{n=1}^{4}\left((\alpha_n - \alpha_{n+1})X^{\mu\nu}_{(n)}(\pi_{(i+1,i)}) + \left(\sum_{i=0}^{n}(-1)^{in}C_i\alpha_{i+1}\right)X^{\mu\nu}_{(n)}(\phi_{(i,i-1)})\right), \tag{B.2}$$

where, as usual $\phi_{(i,j)}$ is the dual Galileon field associated with $\pi_{(i,j)}$, α_n is the coefficient of $\sqrt{-|g^{(i)}|}e_n\left(\sqrt{g^{-1}_{(i)}g^{(i+1)}} - I_4\right)$, and $\alpha_0 = \alpha_1 = 0$.

One then performs a linearised conformal transformation to remove the scalar-tensor mixing at quadratic order (*cf.* (3.27))

$$h^{(i)}_{\mu\nu} \to h^{(i)}_{\mu\nu} - \alpha_2 \frac{m^2 M^2}{8} \left(\pi_{(i+1,i)} + \phi_{(i,i-1)}\right) \eta_{\mu\nu}, \tag{B.3}$$

which will lead to a pure scalar part of the action which mixes all of the fields, and in particular to terms which look like $\pi_{(i+1,i)}\mathcal{L}^{\rm TD}_{(n)}(\pi_{(i+1,i)})$, $\phi_{(i+1,i)}\mathcal{L}^{\rm TD}_{(n)}(\phi_{(i+1,i)})$, $\pi_{(i+1,i)}\mathcal{L}^{\rm TD}_{(n)}(\phi_{(i,i-1)})$, $\phi_{(i,i-1)}\mathcal{L}^{\rm TD}_{(n)}(\pi_{(i+1,i)})$. All of which have been dealt with previously in chapters 3 and 4, and so let us re-express the ϕ fields in terms of the π fields, and arrive at the following action:

$$\mathcal{L}_{\pi\partial^2\pi} = \sum_{i=1}^{3}\sum_{n=1}^{4}\left[a_n \pi_{(i+1,i)}\mathcal{L}^{\rm TD}_{(n)}(\pi_{(i+1,i)}) + b_n \pi_{(i+1,i)}\mathcal{L}^{\rm TD}_{(n)}(\pi_{(i,i-1)}) + c_n \pi_{(i,i-1)}\mathcal{L}^{\rm TD}_{(n)}(\pi_{(i+1,i)})\right], \tag{B.4}$$

where a_n, b_n, c_n depend on $\{\alpha_i\}$. In particular, one has $a_1 = -2c_1 = -2b_1$, and so the kinetic matrix looks like

$$K \propto \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}. \tag{B.5}$$

Thus the Lagrangian is diagonalised at the quadratic level by the fields

$$\chi_{(1)} = \frac{1}{\sqrt{2}}\left(\pi_{(2,1)} - \pi_{(1,3)}\right) \tag{B.6}$$

$$\chi_{(2)} = \frac{1}{\sqrt{6}}\left(\pi_{(2,1)} - 2\pi_{(3,2)} + \pi_{(1,3)}\right) \tag{B.7}$$

$$\chi_{(3)} = \frac{1}{\sqrt{3}}\left(\pi_{(2,1)} + \pi_{(3,2)} + \pi_{(1,3)}\right), \tag{B.8}$$

of which the third is special since its eigenvalue of the kinetic and mass matrices is zero, and hence it drops out of the quadratic order action! This is an indication that it might be acceptable to not use a plaquette, since an appropriate combination corresponding to the constraint (5.1) might drop out of the action, leaving just two propagating modes. However this means that $\chi_{(3)}$ must not reappear other than linearly in higher order interaction terms. (It may appear linearly since in that case partial integration allows us to remove all the derivatives acting on $\chi_{(3)}$, reducing its role to a Lagrange multiplier enforcing a constraint on the dynamics of $\chi_{(1)}$ and $\chi_{(2)}$.) In particular, if it does appear as more than a Lagrange multiplier then, due to its lack of a kinetic term, it will be infinitely strongly coupled.

$\chi_{(3)}^3$	$a_2 + b_2 + c_2 = 0$
$\chi_{(3)}^4$	$a_3 + b_3 + c_3 = 0$
$\chi_{(3)}^2(\chi_{(1)}^2 + \chi_{(2)}^2)$	$4a_3 + b_3 + c_3 = 0$
$\chi_{(3)}^5$	$a_4 + b_4 + c_4 = 0$
$\chi_{(3)}^3(\chi_{(1)}^2 + \chi_{(2)}^2)$	$10a_4 + 4b_4 + 4c_4 = 0$
$\chi_{(3)}^2(\chi_{(1)}^3 - 3\chi_{(1)}\chi_{(2)}^2)$	$b_4 + c_4 = 0$
$\chi_{(3)}^2(3\chi_{(1)}^2\chi_{(2)} - \chi_{(2)}^3)$	$10a_4 + b_4 + c_4 = 0$

Table B.1: The row labelled $\chi_{(3)}^i\chi_{(1)}^j\chi_{(2)}^k$ indicates the condition derived from the vanishing of the coefficient of the (multi-Galileon) interaction term involving i, j, and k copies of the respective field—note that these interaction terms are symmetric in the fields; certain terms do not appear, *e.g.* $\chi_{(3)}^2\chi_{(1)}$, since they vanish automatically.

The constraints this requirement makes on the parameters $\{a_n, b_n, c_n\}$ arising from the cubic, quartic, and quintic terms are presented in table B.1. These can be reduced to five conditions

$$\begin{aligned} &a_2 + b_2 + c_2 = 0, \\ &a_3 = 0, \qquad b_3 + c_3 = 0, \\ &a_4 = 0, \qquad b_4 + c_4 = 0, \end{aligned} \tag{B.9}$$

on the three parameters α_2, α_3, and α_4, and so since this system is overdetermined, it would be surprising if they were satisfied for particular, let along arbitrary, values of α_n. And indeed, just the vanishing of a_3 and a_4 already fixes $\alpha_4 = \frac{3}{4}\alpha_3 = \frac{9}{16}\alpha_2$, which then does not allow the other conditions to be satisfied. Therefore $\chi_{(3)}$ will appear more than linearly in the action, without having a kinetic term, and thus will be infinitely strongly coupled, hindering any further analysis of the theory. This highlights the necessity of using plaquettes.

APPENDIX

C

Solution of Equation (5.40)

Written in matrix notation, and ignoring terms which naïvely are suppressed by Λ_2^2, equation (5.40) becomes (note that ω is antisymmetric, whilst A, B, and C are symmetric)

$$\omega - (\omega C + C\omega + A\omega B + B\omega A) = T, \tag{C.1}$$

which is a combination of the Sylvester and Stein equations. It can be solved by use of the vectorisation operation (which turns an $n \times n$ matrix into a vector of length n^2, made of the concatenated columns of the matrix) and the identity

$$\mathrm{vec}(XYZ) = (Z^{\mathrm{T}} \otimes X)\,\mathrm{vec}(Y), \tag{C.2}$$

where $\otimes$ represents the Kronecker product. Application of these to (C.1) leads to

$$(1 - (1 \otimes C + C \otimes 1 + A \otimes B + B \otimes A))\,\mathrm{vec}(\omega) = \mathrm{vec}(T), \tag{C.3}$$

where 1 represents an identity matrix of the appropriate size. This can then be solved as system of linear equations, and in particular if the matrix on the left hand side is not singular (which is assumed), one can multiply through by its inverse, which is then expanded in a power series

$$(1-(1\otimes C+C\otimes 1+A\otimes B+B\otimes A))^{-1}=\sum_{n=0}^{\infty}(1\otimes C+C\otimes 1+A\otimes B+B\otimes A)^n, \quad \text{(C.4)}$$

and re-written

$$1\otimes C+C\otimes 1+A\otimes B+B\otimes A=\frac{\partial}{\partial a}\frac{\partial}{\partial b}\frac{\partial}{\partial c}(a(A+cB)+b(1+cC))\otimes(a(A+cB)+b(1+cC))\Big|_{a=b=c=0}, \quad \text{(C.5)}$$

to get

$$\mathrm{vec}(\omega)=\sum_{n=0}^{\infty}\frac{\partial^3}{\partial a_1\partial b_1\partial c_1}\cdots\frac{\partial^3}{\partial a_n\partial b_n\partial c_n}(D_1\dots D_n)\otimes(D_1\dots D_n)\Big|_{a=b=c=0}\mathrm{vec}(T), \quad \text{(C.6)}$$

where $D_i=(a_i(A+c_iB)+b_i(1+c_iC))$. Finally turning each side back into a matrix gives

$$\omega=\sum_{n=0}^{\infty}\frac{\partial^3}{\partial a_1\partial b_1\partial c_1}\cdots\frac{\partial^3}{\partial a_n\partial b_n\partial c_n}D_1\dots D_nTD_n\dots D_1\Big|_{a=b=c=0}. \quad \text{(C.7)}$$

For completeness, the terms appearing in (5.40) are

$$T_1^{\mu\nu}=-\left(\frac{1}{2}\tilde{X}^{\mu\lambda}_{(2,0)}+\frac{1}{18}\tilde{X}^{\mu\lambda}_{(0,3)}\right)G^{\nu}_{(1,2)\lambda}+\left(\frac{4}{3}\tilde{X}^{\mu\nu\lambda\rho}_{(1,0)}+\frac{1}{3}\tilde{X}^{\mu\nu\lambda\rho}_{(0,2)}\right)G_{(1,2)\lambda\rho}+\frac{\Lambda_2^2}{18}\tilde{X}^{\mu\lambda}_{(0,3)}\Pi^{\nu}_{(1,2)\lambda} \quad \text{(C.8)}$$

$$A_1^{\mu\nu}=\frac{1}{6}\Pi^{\mu\nu}_{(2,3)} \quad \text{(C.9)}$$

$$B_1^{\mu\nu}=\frac{1}{6}\Pi^{\mu\nu}_{(2,3)} \quad \text{(C.10)}$$

$$\begin{aligned}C_1^{\mu\nu}=&-\Pi^{\mu\nu}_{(1,2)}-\frac{2}{3}\Pi^{\mu\nu}_{(2,3)}-\frac{1}{72}(1+\Pi_{(1,2)})^{\mu}_{\lambda}\left(2X^{\lambda\nu}_{(0,1)}+X^{\lambda\nu}_{(0,2)}+X^{\lambda\nu}_{(0,3)}+18X^{\lambda\nu}_{(1,0)}+9X^{\lambda\nu}_{(2,0)}\right)\\&-\frac{1}{3}(1+\Pi_{(2,3)})^{\mu}_{\lambda}X^{\lambda\nu}_{(0,1)}+\frac{1}{12}\eta^{\mu\nu}\left(\mathcal{L}^{\mathrm{TD}}_{(1,0)}+3\mathcal{L}^{\mathrm{TD}}_{(0,1)}+\mathcal{L}^{\mathrm{TD}}_{(0,2)}\right) \quad \text{(C.11)}\end{aligned}$$

$$T_3^{\mu\nu}=\frac{1}{2}\left(\tilde{X}^{\mu\lambda}_{(1,2)}+\eta^{\mu\lambda}\right)G^{\nu}_{(2,3)\lambda}-2\tilde{X}^{\mu\nu\lambda\rho}_{(1,1)}G_{(2,3)\lambda\rho}+\frac{\Lambda_2^2}{2}\tilde{X}^{\mu\lambda}_{(1,2)}\Pi^{\nu}_{(2,3)\lambda} \quad \text{(C.12)}$$

$$A_3^{\mu\nu}=-\Pi^{\mu\nu}_{(1,2)} \quad \text{(C.13)}$$

$$B_3^{\mu\nu}=\Pi^{\mu\nu}_{(2,3)} \quad \text{(C.14)}$$

$$\begin{aligned}C_3^{\mu\nu}=&-\Pi^{\mu\nu}_{(1,2)}-\frac{3}{4}\Pi^{\mu\nu}_{(2,3)}+\frac{1}{8}(1+\Pi_{(2,3)})^{\mu}_{\lambda}\left(2X^{\lambda\nu}_{(0,1)}+X^{\lambda\nu}_{(0,2)}+8X^{\lambda\nu}_{(1,0)}+X^{\lambda\nu}_{(1,1)}+X^{\lambda\nu}_{(1,2)}\right)\\&+(1+\Pi_{(1,2)})^{\mu}_{\lambda}X^{\lambda\nu}_{(0,1)}-\frac{1}{2}\eta^{\mu\nu}\left(3\mathcal{L}^{\mathrm{TD}}_{(0,1)}+3\mathcal{L}^{\mathrm{TD}}_{(1,0)}+\mathcal{L}^{\mathrm{TD}}_{(1,1)}\right). \quad \text{(C.15)}\end{aligned}$$

APPENDIX

D

The Spectrum of The Generalised Barbell Graph

The line graph of the graph depicted in figure 6.2 consists of two complete graphs of $\lfloor \frac{N+3-d}{2} \rfloor$ and $\lceil \frac{N+3-d}{2} \rceil$ vertices connected by a path graph with $d-3$ edges. This is an example of a 'generalised barbell graph', $B_{a,b,c}$, which is defined to consist of a path P_{b+2} which has each endpoint in common with each of two complete graphs, K_a and K_c.[1]

The adjacency matrix is

$$A(B_{a,b,c}) = \begin{pmatrix} J_{a,a} & \begin{matrix} 0_{a-1,1} & 0_{a-1,b-1} \\ 1 & 0_{1,b-1} \end{matrix} & 0_{a,c} \\ \begin{matrix} 0_{1,a-1} & 1 \\ 0_{b-1,a-1} & 0_{b-1,1} \end{matrix} & T_b & \begin{matrix} 0_{b-1,1} & 0_{b-1,c-1} \\ 1 & 0_{1,c-1} \end{matrix} \\ 0_{c,a} & \begin{matrix} 0_{1,b-1} & 1 \\ 0_{c-1,b-1} & 0_{c-1,1} \end{matrix} & J_{c,c} \end{pmatrix} - I_{a+b+c}, \quad \text{(D.1)}$$

[1]In this notation the ordinary barbell graph would be $B_{n,0,n}$.

where $J_{n,n}$ is the $n \times n$ matrix of 1's, and T_n is a tri-diagonal matrix of 1's. The characteristic polynomial $\Delta_{a,b,c}$ obeys the following recurrence relation in a:

$$\Delta_{a,b,c} = |\lambda I_{a+b+c} - A(B_{a,b,c})| = \lambda \Delta_{a-1,b,c} + (a-2)\Delta^{(1)}_{a-2,b,c} + \Delta^{(2)}_{a-2,b,c}, \tag{D.2}$$

where

$$\begin{aligned} \Delta^{(1)}_{a,b,c} &= \left| \lambda I_{a+b+c+1} - A(B_{a+1,b,c}) - \begin{pmatrix} \lambda+1 & 0_{1,a+b+c} \\ 0_{a+b+c,1} & 0_{a+b+c,a+b+c} \end{pmatrix} \right| \\ &= -\Delta_{a,b,c} + (a-1)\Delta^{(1)}_{a-1,b,c} + \Delta^{(2)}_{a-1,b,c}, \end{aligned} \tag{D.3}$$

and

$$\begin{aligned} \Delta^{(2)}_{a,b,c} &= \left| \lambda I_{a+b+c+1} - A(B_{a+1,b,c}) - \begin{pmatrix} \lambda+1 & 0_{1,a} & 1 & 0_{1,b-1+c} \\ 0_{a+b+c,1} & 0_{a+b+c,a} & 0_{a+b+c,1} & 0_{a+b+c,b-1+c} \end{pmatrix} \right| \\ &= -(\lambda+1)^a \Delta_{0,b,c}. \end{aligned} \tag{D.4}$$

These recurrence relations can be solved, noting that $\Delta_{1,b,c} = \Delta_{0,b+1,c}$, to get

$$\Delta_{a,b,c} = -(1+\lambda)^{a-2} \left((a-1)\Delta_{0,b,c} - (\lambda - (a-2))\Delta_{0,b+1,c} \right). \tag{D.5}$$

Now $\Delta_{0,b,c}$ obeys the following recurrence relation in b:

$$\Delta_{0,b,c} = \lambda \Delta_{0,b-1,c} - \Delta_{0,b-2,c}, \tag{D.6}$$

which, by using $\Delta_{0,0,c} = (\lambda+1)^{c-1}(\lambda-(c-1))$ and $\Delta_{0,1,c} = \lambda \Delta_{0,0,c} - (\lambda+1)^{c-2}(\lambda-(c-2))$, can be solved to give

$$\begin{aligned} \Delta_{0,b,c} =& \frac{1}{2^{b+3}}(\lambda+1)^{c-2} \Bigg[(c-1)\left(\left(\lambda - \sqrt{\lambda^2-4}\right)^{b+2} + \left(\lambda + \sqrt{\lambda^2-4}\right)^{b+2} \right) \\ &+ (\lambda(c-3) + 2(c-2)) \left(\frac{\left(\lambda - \sqrt{\lambda^2-4}\right)^{b+2} - \left(\lambda + \sqrt{\lambda^2-4}\right)^{b+2}}{\sqrt{\lambda^2-4}} \right) \Bigg]. \end{aligned} \tag{D.7}$$

Finally, performing a change of variables: $\lambda = -2\cos\theta$, substituting (D.7) in (D.5), and a little algebra yields

$$\Delta_{a,b,c} = (-1)^b (1 - 2\cos\theta)^{a+c-4} \operatorname{cosec}\theta \, f_{a,b,c}(\theta), \tag{D.8}$$

where

$$f_{a,b,c}(\theta) = \sin((b+5)\theta) + (a+c-4)\sin((b+4)\theta) + ((a-3)(c-3)-1)\sin((b+3)\theta) \\ - 2(a-2)(c-2)\sin((b+2)\theta) + (a-2)(c-2)\sin((b+1)\theta). \tag{D.9}$$

$f_{N,d}$ in (6.7) is then just (D.9) with $a = c = \frac{N+1-d}{2}$ and $b = d-4$.

D.1 Spectrum of $L(M_{d,\Delta})$

One has $L_{M_{d,\Delta}} = B_{\Delta,d-4,2}$,[2] and hence

$$\mathrm{spec}(L(M_{d,\Delta})) = \{0, 2(1-\cos\theta_i^*), 1^{d+\Delta-8}\}, \tag{D.10}$$

where θ_i^* is a root of

$$\begin{aligned} f_{\Delta,d-4,2}(\theta) &= \sin((d+1)\theta) + (\Delta-2)(\sin(d\theta) - \sin((d-1)\theta)) \\ &= (\Delta-1)\sin(d\theta)\sin(\theta)\left[\cot(d\theta) - \left(\frac{\Delta-3}{\Delta-1}\cot(\theta) - \frac{\Delta-2}{\Delta-1}\mathrm{cosec}(\theta)\right)\right]. \end{aligned} \tag{D.11}$$

From the square bracket one sees that $(n-\frac{1}{2})\frac{\pi}{d} < \theta_n^* < n\frac{\pi}{d}$. Hence one has for $d > 4$

$$\frac{m_2}{m_1} \approx \frac{\theta_2^*}{\theta_1^*} > \frac{3}{2}. \tag{D.12}$$

[2]Note that there is a degeneracy in the final two parameters: $B_{\Delta,d-4,2} = B_{\Delta,d-3,1} = B_{\Delta,d-2,0}$.

References

[1] S. Deser, *Self-Interaction and gauge invariance*, *Gen. Rel. Grav.* **1** (1970) 9–18, [gr-qc/0411023].

[2] L. M. Butcher, M. Hobson and A. Lasenby, *Bootstrapping gravity: A Consistent approach to energy-momentum self-coupling*, *Phys. Rev.* **D80** (2009) 084014, [0906.0926].

[3] T. Padmanabhan, *From gravitons to gravity: Myths and reality*, *Int. J. Mod. Phys.* **D17** (2008) 367–398, [gr-qc/0409089].

[4] N. Boulanger, T. Damour, L. Gualtieri and M. Henneaux, *Inconsistency of interacting, multigraviton theories*, *Nucl.Phys.* **B597** (2001) 127–171, [hep-th/0007220].

[5] M. Fierz and W. Pauli, *On relativistic wave equations for particles of arbitrary spin in an electromagnetic field*, *Proc.Roy.Soc.Lond.* **A173** (1939) 211–232.

[6] D. Boulware and S. Deser, *Can gravitation have a finite range?*, *Phys.Rev.* **D6** (1972) 3368–3382.

[7] J. B. Pitts, *Universally coupled massive gravity, III: dRGT–Maheshwari pure spin-2, Ogievetsky–Polubarinov and arbitrary mass terms*, *Annals Phys.* **365** (2016) 73–90, [1505.03492].

[8] A. Maheshwari, *Spin-2 field theories and the tensor-field identity*, *Nuovo Cim.* **A8** (1972) 319–330.

[9] C. de Rham and G. Gabadadze, *Generalization of the Fierz-Pauli Action*, *Phys.Rev.* **D82** (2010) 044020, [1007.0443].

[10] C. de Rham, G. Gabadadze and A. J. Tolley, *Resummation of Massive Gravity*, *Phys.Rev.Lett.* **106** (2011) 231101, [1011.1232].

[11] C. de Rham, G. Gabadadze and A. J. Tolley, *Helicity Decomposition of Ghost-free Massive Gravity*, *JHEP* **1111** (2011) 093, [1108.4521].

[12] C. de Rham, G. Gabadadze and A. J. Tolley, *Ghost free Massive Gravity in the Stückelberg language*, *Phys.Lett.* **B711** (2012) 190–195, [1107.3820].

[13] S. F. Hassan and R. A. Rosen, *On Non-Linear Actions for Massive Gravity*, *JHEP* **07** (2011) 009, [1103.6055].

[14] S. Hassan and R. A. Rosen, *Resolving the Ghost Problem in non-Linear Massive Gravity*, *Phys.Rev.Lett.* **108** (2012) 041101, [1106.3344].

[15] S. Hassan, R. A. Rosen and A. Schmidt-May, *Ghost-free Massive Gravity with a General Reference Metric*, *JHEP* **1202** (2012) 026, [1109.3230].

[16] S. Hassan, A. Schmidt-May and M. von Strauss, *Proof of Consistency of Nonlinear Massive Gravity in the Stückelberg Formulation*, *Phys.Lett.* **B715** (2012) 335–339, [1203.5283].

[17] S. Hassan and R. A. Rosen, *Bimetric Gravity from Ghost-free Massive Gravity*, *JHEP* **1202** (2012) 126, [1109.3515].

[18] S. Hassan and R. A. Rosen, *Confirmation of the Secondary Constraint and Absence of Ghost in Massive Gravity and Bimetric Gravity*, *JHEP* **1204** (2012) 123, [1111.2070].

[19] K. Hinterbichler and R. A. Rosen, *Interacting Spin-2 Fields*, *JHEP* **1207** (2012) 047, [1203.5783].

[20] K. Hinterbichler, *Theoretical Aspects of Massive Gravity*, *Rev.Mod.Phys.* **84** (2012) 671–710, [1105.3735].

[21] C. de Rham, *Massive Gravity*, *Living Rev.Rel.* **17** (2014) 7, [1401.4173].

[22] A. Schmidt-May and M. von Strauss, *Recent developments in bimetric theory*, *J. Phys.* **A49** (2016) 183001, [1512.00021].

[23] G. Dvali, G. Gabadadze and M. Shifman, *Diluting cosmological constant in infinite volume extra dimensions*, *Phys. Rev.* **D67** (2003) 044020, [hep-th/0202174].

[24] N. Arkani-Hamed, S. Dimopoulos, G. Dvali and G. Gabadadze, *Nonlocal modification of gravity and the cosmological constant problem*, hep-th/0209227.

[25] G. Dvali, S. Hofmann and J. Khoury, *Degravitation of the cosmological constant and graviton width*, *Phys. Rev.* **D76** (2007) 084006, [hep-th/0703027].

[26] M. S. Volkov, *Exact self-accelerating cosmologies in the ghost-free bigravity and massive gravity*, *Phys. Rev.* **D86** (2012) 061502, [1205.5713].

[27] M. S. Volkov, *Exact self-accelerating cosmologies in the ghost-free massive gravity – the detailed derivation*, *Phys. Rev.* **D86** (2012) 104022, [1207.3723].

[28] H. van Dam and M. J. G. Veltman, *Massive and massless Yang-Mills and gravitational fields*, *Nucl. Phys.* **B22** (1970) 397–411.

[29] V. I. Zakharov, *Linearized gravitation theory and the graviton mass*, *JETP Lett.* **12** (1970) 312.

[30] Y. Iwasaki, *Consistency condition for propagators*, *Phys. Rev.* **D2** (1970) 2255–2256.

[31] A. I. Vainshtein, *To the problem of nonvanishing gravitation mass*, *Phys. Lett.* **B39** (1972) 393–394.

[32] C. Deffayet, G. R. Dvali, G. Gabadadze and A. I. Vainshtein, *Nonperturbative continuity in graviton mass versus perturbative discontinuity*, *Phys. Rev.* **D65** (2002) 044026, [hep-th/0106001].

[33] E. Babichev, C. Deffayet and R. Ziour, *The Recovery of General Relativity in massive gravity via the Vainshtein mechanism*, *Phys. Rev.* **D82** (2010) 104008, [1007.4506].

[34] E. Babichev and C. Deffayet, *An introduction to the Vainshtein mechanism*, *Class. Quant. Grav.* **30** (2013) 184001, [1304.7240].

[35] A. Nicolis and R. Rattazzi, *Classical and quantum consistency of the DGP model*, *JHEP* **06** (2004) 059, [hep-th/0404159].

[36] C. de Rham, G. Gabadadze, L. Heisenberg and D. Pirtskhalava, *Cosmic Acceleration and the Helicity-0 Graviton*, *Phys. Rev.* **D83** (2011) 103516, [1010.1780].

[37] C. Burrage, N. Kaloper and A. Padilla, *Strong Coupling and Bounds on the Spin-2 Mass in Massive Gravity*, *Phys. Rev. Lett.* **111** (2013) 021802, [1211.6001].

[38] N. Kaloper, A. Padilla, P. Saffin and D. Stefanyszyn, *Unitarity and the Vainshtein Mechanism*, *Phys. Rev.* **D91** (2015) 045017, [1409.3243].

[39] U. Aydemir, M. M. Anber and J. F. Donoghue, *Self-healing of unitarity in effective field theories and the onset of new physics*, *Phys. Rev.* **D86** (2012) 014025, [1203.5153].

[40] C. de Rham, A. J. Tolley and S.-Y. Zhou, *Non-compact nonlinear sigma models*, 1512.06838.

[41] C. de Rham, A. J. Tolley and S.-Y. Zhou, *The Λ_2 limit of massive gravity*, 1602.03721.

[42] K. Aoki and S. Mukohyama, *Bigravitons as dark matter and gravitational waves*, 1604.06704.

[43] E. Babichev, L. Marzola, M. Raidal, A. Schmidt-May, F. Urban, H. Veermäe et al., *Gravitational Origin of Dark Matter*, 1604.08564.

[44] C. J. Isham, A. Salam and J. A. Strathdee, *F-dominance of gravity*, *Phys. Rev.* **D3** (1971) 867–873.

[45] C. Sivaram and K. P. Sinha, *STRONG SPIN-2 INTERACTION AND GENERAL RELATIVITY*, *Phys. Rept.* **51** (1979) 111–187.

[46] N. Arkani-Hamed, A. G. Cohen and H. Georgi, *(De)constructing dimensions*, *Phys.Rev.Lett.* **86** (2001) 4757–4761, [hep-th/0104005].

[47] N. Arkani-Hamed, H. Georgi and M. D. Schwartz, *Effective field theory for massive gravitons and gravity in theory space*, *Annals Phys.* **305** (2003) 96–118, [hep-th/0210184].

[48] N. Arkani-Hamed and M. D. Schwartz, *Discrete gravitational dimensions*, *Phys.Rev.* **D69** (2004) 104001, [hep-th/0302110].

[49] M. D. Schwartz, *Constructing gravitational dimensions*, *Phys.Rev.* **D68** (2003) 024029, [hep-th/0303114].

[50] C. de Rham, A. Matas and A. J. Tolley, *Deconstructing Dimensions and Massive Gravity*, *Class. Quant. Grav.* **31** (2014) 025004, [1308.4136].

[51] C. Deffayet and J. Mourad, *Multigravity from a discrete extra dimension*, *Phys.Lett.* **B589** (2004) 48–58, [hep-th/0311124].

[52] T. Damour and I. I. Kogan, *Effective Lagrangians and universality classes of nonlinear bigravity*, *Phys. Rev.* **D66** (2002) 104024, [hep-th/0206042].

[53] I. I. Kogan, S. Mouslopoulos, A. Papazoglou, G. G. Ross and J. Santiago, *A Three three-brane universe: New phenomenology for the new millennium?*, *Nucl. Phys.* **B584** (2000) 313–328, [hep-ph/9912552].

[54] I. I. Kogan, *Multi(scale)gravity: A Telescope for the microworld*, 2001. astro-ph/0108220.

[55] G. R. Dvali, G. Gabadadze and M. Porrati, *Metastable gravitons and infinite volume extra dimensions*, *Phys. Lett.* **B484** (2000) 112–118, [hep-th/0002190].

[56] G. R. Dvali, G. Gabadadze and M. Porrati, *4-D gravity on a brane in 5-D Minkowski space*, *Phys. Lett.* **B485** (2000) 208–214, [hep-th/0005016].

[57] G. R. Dvali and G. Gabadadze, *Gravity on a brane in infinite volume extra space*, *Phys. Rev.* **D63** (2001) 065007, [hep-th/0008054].

[58] C. Deffayet, *Cosmology on a brane in Minkowski bulk*, *Phys. Lett.* **B502** (2001) 199–208, [hep-th/0010186].

[59] C. Deffayet, G. R. Dvali and G. Gabadadze, *Accelerated universe from gravity leaking to extra dimensions*, *Phys. Rev.* **D65** (2002) 044023, [astro-ph/0105068].

[60] M. A. Luty, M. Porrati and R. Rattazzi, *Strong interactions and stability in the DGP model*, *JHEP* **09** (2003) 029, [hep-th/0303116].

[61] A. Nicolis, R. Rattazzi and E. Trincherini, *The Galileon as a local modification of gravity*, *Phys.Rev.* **D79** (2009) 064036, [0811.2197].

[62] C. de Rham and A. J. Tolley, *DBI and the Galileon reunited*, *JCAP* **1005** (2010) 015, [1003.5917].

[63] A. Gruzinov, *All Fierz-Paulian massive gravity theories have ghosts or superluminal modes*, 1106.3972.

[64] C. Burrage, C. de Rham, L. Heisenberg and A. J. Tolley, *Chronology Protection in Galileon Models and Massive Gravity*, *JCAP* **1207** (2012) 004, [1111.5549].

[65] S. Deser and A. Waldron, *Acausality of Massive Gravity*, *Phys. Rev. Lett.* **110** (2013) 111101, [1212.5835].

[66] S. Deser, K. Izumi, Y. C. Ong and A. Waldron, *Massive Gravity Acausality Redux*, *Phys. Lett.* **B726** (2013) 544–548, [1306.5457].

[67] S. Deser, K. Izumi, Y. C. Ong and A. Waldron, *Problems of massive gravities*, *Mod. Phys. Lett.* **A30** (2015) 1540006, [1410.2289].

[68] I. T. Drummond and S. J. Hathrell, *QED Vacuum Polarization in a Background Gravitational Field and Its Effect on the Velocity of Photons*, *Phys. Rev.* **D22** (1980) 343.

[69] G. Goon, *Black Holes, Superluminality and EFT.* Oxford Cosmology Seminar, 19$^{\text{th}}$ Nov. 2015.

[70] G. M. Shore, *Superluminality and UV completion*, *Nucl. Phys.* **B778** (2007) 219–258, [hep-th/0701185].

[71] T. J. Hollowood and G. M. Shore, *Causality and Micro-Causality in Curved Spacetime*, *Phys. Lett.* **B655** (2007) 67–74, [0707.2302].

[72] T. J. Hollowood and G. M. Shore, *The Refractive index of curved spacetime: The Fate of causality in QED*, *Nucl. Phys.* **B795** (2008) 138–171, [0707.2303].

[73] C. de Rham, M. Fasiello and A. J. Tolley, *Galileon Duality*, *Phys. Lett.* **B733** (2014) 46–51, [1308.2702].

[74] C. De Rham, L. Keltner and A. J. Tolley, *Generalized galileon duality*, *Phys. Rev.* **D90** (2014) 024050, [1403.3690].

[75] P. Creminelli, M. Serone, G. Trevisan and E. Trincherini, *Inequivalence of Coset Constructions for Spacetime Symmetries*, *JHEP* **02** (2015) 037, [1403.3095].

[76] P. Baratella, P. Creminelli, M. Serone and G. Trevisan, *Inequivalence of Coset Constructions for Spacetime Symmetries: Coupling with Gravity*, *Phys. Rev.* **D93** (2016) 045029, [1510.01969].

[77] J. Noller, J. H. C. Scargill and P. G. Ferreira, *Interacting spin-2 fields in the Stückelberg picture*, *JCAP* **1402** (2014) 007, [1311.7009].

[78] J. Noller and J. H. C. Scargill, *The decoupling limit of Multi-Gravity: Multi-Galileons, Dualities and More*, *JHEP* **05** (2015) 034, [1503.02700].

[79] J. H. C. Scargill, J. Noller and P. G. Ferreira, *Cycles of interactions in multi-gravity theories*, *JHEP* **12** (2014) 160, [1410.7774].

[80] J. H. C. Scargill and J. Noller, *Strong-coupling scales and the graph structure of multi-gravity theories*, *JHEP* **01** (2016) 029, [1511.02877].

[81] S. Weinberg, *The Quantum theory of fields. Vol. 1: Foundations.* Cambridge University Press, 2005.

[82] E. A. Abbott, *Flatland: A Romance of Many Dimensions.* London: Seeley & Co., 1884.

[83] R. P. Woodard, *Avoiding dark energy with 1/r modifications of gravity*, *Lect. Notes Phys.* **720** (2007) 403–433, [astro-ph/0601672].

[84] F. Sbisà, *Classical and quantum ghosts*, *Eur. J. Phys.* **36** (2015) 015009, [1406.4550].

[85] J. Khoury, *Theories of Dark Energy with Screening Mechanisms*, 1011.5909.

[86] M. Banados, A. Gomberoff, D. C. Rodrigues and C. Skordis, *A Note on bigravity and dark matter*, *Phys.Rev.* **D79** (2009) 063515, [0811.1270].

[87] M. Jaccard, M. Maggiore and E. Mitsou, *Nonlocal theory of massive gravity*, *Phys. Rev.* **D88** (2013) 044033, [1305.3034].

[88] G. Cusin, J. Fumagalli and M. Maggiore, *Non-local formulation of ghost-free bigravity theory*, *JHEP* **09** (2014) 181, [1407.5580].

[89] D. Comelli, M. Crisostomi, K. Koyama, L. Pilo and G. Tasinato, *New Branches of Massive Gravity*, *Phys. Rev.* **D91** (2015) 121502, [1505.00632].

[90] M. Henneaux and C. Teitelboim, *Quantization of gauge systems.* Princeton University Press, 1992.

[91] R. L. Arnowitt, S. Deser and C. W. Misner, *The Dynamics of general relativity*, *Gen. Rel. Grav.* **40** (2008) 1997–2027, [gr-qc/0405109].

[92] A. Schmidt-May, *Classically Consistent Theories of Interacting Spin-2 Fields.* PhD thesis, Stockholm University, 2013.

[93] S. Hassan, A. Schmidt-May and M. von Strauss, *Metric Formulation of Ghost-Free Multivielbein Theory*, 1204.5202.

[94] C. Deffayet, J. Mourad and G. Zahariade, *Covariant constraints in ghost free massive gravity*, *JCAP* **1301** (2013) 032, [1207.6338].

[95] N. A. Ondo and A. J. Tolley, *Complete Decoupling Limit of Ghost-free Massive Gravity*, *JHEP* **11** (2013) 059, [1307.4769].

[96] G. Gabadadze, K. Hinterbichler, D. Pirtskhalava and Y. Shang, *Potential for general relativity and its geometry*, *Phys.Rev.* **D88** (2013) 084003, [1307.2245].

[97] C. Deffayet, J. Mourad and G. Zahariade, *A note on 'symmetric' vielbeins in bimetric, massive, perturbative and non perturbative gravities*, *JHEP* **1303** (2013) 086, [1208.4493].

[98] S. Alexandrov, K. Krasnov and S. Speziale, *Chiral description of ghost-free massive gravity*, *JHEP* **06** (2013) 068, [1212.3614].

[99] K. Nomura and J. Soda, *When is Multimetric Gravity Ghost-free?*, *Phys.Rev.* **D86** (2012) 084052, [1207.3637].

[100] J. Noller and S. Melville, *The coupling to matter in Massive, Bi- and Multi-Gravity*, *JCAP* **1501** (2015) 003, [1408.5131].

[101] C. de Rham, L. Heisenberg and R. H. Ribeiro, *On couplings to matter in massive (bi-)gravity*, *Class. Quant. Grav.* **32** (2015) 035022, [1408.1678].

[102] S. Melville and J. Noller, *Generalised matter couplings in massive bigravity*, *JHEP* **01** (2016) 094, [1511.01485].

[103] Y. Yamashita, A. De Felice and T. Tanaka, *Appearance of Boulware–Deser ghost in bigravity with doubly coupled matter*, *Int. J. Mod. Phys.* **D23** (2014) 1443003, [1408.0487].

[104] S. Folkerts, A. Pritzel and N. Wintergerst, *On ghosts in theories of self-interacting massive spin-2 particles*, 1107.3157.

[105] K. Hinterbichler, *Ghost-Free Derivative Interactions for a Massive Graviton*, *JHEP* **10** (2013) 102, [1305.7227].

[106] R. Kimura and D. Yamauchi, *Derivative interactions in de Rham-Gabadadze-Tolley massive gravity*, *Phys. Rev.* **D88** (2013) 084025, [1308.0523].

[107] C. de Rham, A. Matas and A. J. Tolley, *New Kinetic Interactions for Massive Gravity?*, *Class. Quant. Grav.* **31** (2014) 165004, [1311.6485].

[108] J. Noller, *On Consistent Kinetic and Derivative Interactions for Gravitons*, *JCAP* **1504** (2015) 025, [1409.7692].

[109] C. de Rham, A. Matas and A. J. Tolley, *New Kinetic Terms for Massive Gravity and Multi-gravity: A No-Go in Vielbein Form*, *Class. Quant. Grav.* **32** (2015) 215027, [1505.00831].

[110] C. de Rham, G. Gabadadze, L. Heisenberg and D. Pirtskhalava, *Nonrenormalization and naturalness in a class of scalar-tensor theories*, *Phys. Rev.* **D87** (2013) 085017, [1212.4128].

[111] M. Park, *Quantum Aspects of Massive Gravity*, *Class. Quant. Grav.* **28** (2011) 105012, [1009.4369].

[112] I. L. Buchbinder, D. D. Pereira and I. L. Shapiro, *One-loop divergences in massive gravity theory*, *Phys. Lett.* **B712** (2012) 104–108, [1201.3145].

[113] C. de Rham, L. Heisenberg and R. H. Ribeiro, *Quantum Corrections in Massive Gravity*, *Phys. Rev.* **D88** (2013) 084058, [1307.7169].

[114] L. Heisenberg, *Quantum corrections in massive bigravity and new effective composite metrics*, *Class. Quant. Grav.* **32** (2015) 105011, [1410.4239].

[115] G. 't Hooft and M. J. G. Veltman, *One loop divergencies in the theory of gravitation*, *Annales Poincare Phys. Theor.* **A20** (1974) 69–94.

[116] G. 't Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking*, *NATO Sci. Ser. B* **59** (1980) 135.

[117] S. Dimopoulos and L. Susskind, *Mass Without Scalars*, *Nucl. Phys.* **B155** (1979) 237–252.

[118] S. Deser and A. Waldron, *Partial masslessness of higher spins in (A)dS*, *Nucl. Phys.* **B607** (2001) 577–604, [hep-th/0103198].

[119] S. Deser and A. Waldron, *Stability of massive cosmological gravitons*, *Phys. Lett.* **B508** (2001) 347–353, [hep-th/0103255].

[120] S. Deser and A. Waldron, *Null propagation of partially massless higher spins in (A)dS and cosmological constant speculations*, *Phys. Lett.* **B513** (2001) 137–141, [hep-th/0105181].

[121] G. Gabadadze and A. Iglesias, *Special Massive Spin-2 on de Sitter Space*, *JCAP* **0802** (2008) 014, [0801.2165].

[122] C. de Rham, K. Hinterbichler, R. A. Rosen and A. J. Tolley, *Evidence for and obstructions to nonlinear partially massless gravity*, *Phys. Rev.* **D88** (2013) 024003, [1302.0025].

[123] S. Deser, M. Sandora and A. Waldron, *Nonlinear Partially Massless from Massive Gravity?*, *Phys. Rev.* **D87** (2013) 101501, [1301.5621].

[124] S. Garcia-Saenz and R. A. Rosen, *A non-linear extension of the spin-2 partially massless symmetry*, *JHEP* **05** (2015) 042, [1410.8734].

[125] S. F. Hassan, A. Schmidt-May and M. von Strauss, *On Partially Massless Bimetric Gravity*, *Phys. Lett.* **B726** (2013) 834–838, [1208.1797].

[126] S. F. Hassan, A. Schmidt-May and M. von Strauss, *Bimetric theory and partial masslessness with Lanczos–Lovelock terms in arbitrary dimensions*, *Class. Quant. Grav.* **30** (2013) 184010, [1212.4525].

[127] S. F. Hassan, A. Schmidt-May and M. von Strauss, *Higher Derivative Gravity and Conformal Gravity From Bimetric and Partially Massless Bimetric Theory*, *Universe* **1** (2015) 92–122, [1303.6940].

[128] S. F. Hassan, A. Schmidt-May and M. von Strauss, *Particular Solutions in Bimetric Theory and Their Implications*, *Int. J. Mod. Phys.* **D23** (2014) 1443002, [1407.2772].

[129] S. F. Hassan, A. Schmidt-May and M. von Strauss, *Extended Weyl Invariance in a Bimetric Model and Partial Masslessness*, *Class. Quant. Grav.* **33** (2016) 015011, [1507.06540].

[130] S. Garcia-Saenz, K. Hinterbichler, A. Joyce, E. Mitsou and R. A. Rosen, *No-go for Partially Massless Spin-2 Yang-Mills*, *JHEP* **02** (2016) 043, [1511.03270].

[131] E. C. G. Stueckelberg, *Théorie de la radiation de photons de masse arbitrairement petite*, *Helv. Phys. Acta* **30** (1957) 209–215.

[132] J. M. Cornwall, D. N. Levin and G. Tiktopoulos, *Derivation of Gauge Invariance from High-Energy Unitarity Bounds on the s Matrix*, *Phys. Rev.* **D10** (1974) 1145.

[133] C. E. Vayonakis, *Born Helicity Amplitudes and Cross-Sections in Nonabelian Gauge Theories*, *Lett. Nuovo Cim.* **17** (1976) 383.

[134] B. W. Lee, C. Quigg and H. B. Thacker, *Weak Interactions at Very High-Energies: The Role of the Higgs Boson Mass*, *Phys. Rev.* **D16** (1977) 1519.

[135] M. S. Chanowitz and M. K. Gaillard, *The TeV Physics of Strongly Interacting W's and Z's*, *Nucl. Phys.* **B261** (1985) 379.

[136] M. Fasiello and A. J. Tolley, *Cosmological Stability Bound in Massive Gravity and Bigravity*, *JCAP* **1312** (2013) 002, [1308.1647].

[137] P. Creminelli, A. Nicolis, M. Papucci and E. Trincherini, *Ghosts in massive gravity*, *JHEP* **09** (2005) 003, [hep-th/0505147].

[138] M. V. Ostrogradsky, *Mémoires sur les équations différentielles relatives au problème des isopérimètres*, *Mem. Acad. St. Petersbourg* **VI 4** (1850) 385–517.

[139] A. Salvio and A. Strumia, *Quantum mechanics of 4-derivative theories*, 1512.01237.

[140] Y. Akrami, S. F. Hassan, F. Könnig, A. Schmidt-May and A. R. Solomon, *Bimetric gravity is cosmologically viable*, *Phys. Lett.* **B748** (2015) 37–44, [1503.07521].

[141] D. Fairlie, J. Govaerts and A. Morozov, *Universal field equations with covariant solutions*, *Nucl.Phys.* **B373** (1992) 214–232, [hep-th/9110022].

[142] G. W. Horndeski, *Second-order scalar-tensor field equations in a four-dimensional space*, *Int.J.Theor.Phys.* **10** (1974) 363–384.

[143] M. Trodden and K. Hinterbichler, *Generalizing Galileons*, *Class. Quant. Grav.* **28** (2011) 204003, [1104.2088].

[144] G. Goon, K. Hinterbichler, A. Joyce and M. Trodden, *Galileons as Wess-Zumino Terms*, *JHEP* **06** (2012) 004, [1203.3191].

[145] A. Padilla, P. M. Saffin and S.-Y. Zhou, *Multi-galileons, solitons and Derrick's theorem*, *Phys. Rev.* **D83** (2011) 045009, [1008.0745].

[146] C. Deffayet, S. Deser and G. Esposito-Farese, *Arbitrary p-form Galileons*, *Phys. Rev.* **D82** (2010) 061501, [1007.5278].

[147] K. Hinterbichler, M. Trodden and D. Wesley, *Multi-field galileons and higher co-dimension branes*, *Phys. Rev.* **D82** (2010) 124018, [1008.1305].

[148] K. Kampf and J. Novotny, *Unification of Galileon Dualities*, *JHEP* **10** (2014) 006, [1403.6813].

[149] G. Gabadadze, K. Hinterbichler, J. Khoury, D. Pirtskhalava and M. Trodden, *A Covariant Master Theory for Novel Galilean Invariant Models and Massive Gravity*, *Phys.Rev.* **D86** (2012) 124004, [1208.5773].

[150] M. Zumalacárregui and J. García-Bellido, *Transforming gravity: from derivative couplings to matter to second-order scalar-tensor theories beyond the Horndeski Lagrangian*, *Phys.Rev.* **D89** (2014) 064046, [1308.4685].

[151] J. Gleyzes, D. Langlois, F. Piazza and F. Vernizzi, *Healthy theories beyond Horndeski*, *Phys. Rev. Lett.* **114** (2015) 211101, [1404.6495].

[152] H. R. Afshar, E. A. Bergshoeff and W. Merbis, *Interacting spin-2 fields in three dimensions*, *JHEP* **01** (2015) 040, [1410.6164].

[153] N. Khosravi, N. Rahmanpour, H. R. Sepangi and S. Shahidi, *Multi-Metric Gravity via Massive Gravity*, *Phys.Rev.* **D85** (2012) 024049, [1111.5346].

[154] C. Yan, *Properties of spectra of graphs and line graphs*, *Applied Mathematics-A Journal of Chinese Universities* **17** (2002) 371–376.

[155] M. Doob, *An interrelation between line graphs, eigenvalues, and matroids*, *Journal of Combinatorial Theory, Series B* **15** (1973) 40 – 50.

[156] R. Brualdi, *The Mutually Beneficial Relationship of Graphs and Matrices.* Regional conference series in mathematics. American Mathematical Soc.

[157] N. Kan and K. Shiraishi, *Multigraviton theory: A Latticized dimension, and the cosmological constant*, *Class. Quant. Grav.* **20** (2003) 4965–4972, [gr-qc/0212113].

[158] R. Grone and R. Merris, *The laplacian spectrum of a graph ii*, *SIAM Journal on Discrete Mathematics* **7** (1994) 221–229, [http://dx.doi.org/10.1137/S0895480191222653].

[159] D. Stevanović, *Bounding the largest eigenvalue of trees in terms of the largest vertex degree*, *Linear Algebra and its Applications* **360** (2003) 35 – 42.

[160] T.-j. Chen, M. Fasiello, E. A. Lim and A. J. Tolley, *Higher derivative theories with constraints: Exorcising Ostrogradski's Ghost*, *JCAP* **1302** (2013) 042, [1209.0583].