Constraints and symmetries in theories of interacting spin-2 fields

Anders Lundkvist
Constraints and symmetries in theories of interacting spin-2 fields

Anders Lundkvist

Academic dissertation for the Degree of Doctor of Philosophy in Theoretical Physics at Stockholm University to be publicly defended on Friday 21 February 2020 at 13.00 in FA32, AlbaNova universitetscentrum, Roslagstullsbacken 21.

Abstract

The Hassan-Rosen bimetric theory describes two interacting spin-2 fields, one massless and one massive. In this thesis, a complete canonical analysis of this theory is performed in the metric formulation and all constraints are computed. In particular, a secondary constraint, whose existence was in doubt, is shown to exist and evaluated explicitly, bringing the total number of constraints up to six. This, together with general covariance, is enough to eliminate the Boulware-Deser ghost and ensure that the theory propagates the appropriate seven degrees of freedom. The requirement that the constraints are preserved in time leads to a linear relation between the lapse functions of the two metrics. Knowing the explicit form of the ratio of the lapses is necessary for solving initial value problems. The ratio is computed for the special case where the metrics share the same spherical symmetry.

Since the bimetric theory is diffeomorphism invariant, it must contain four first class constraints whose Poisson brackets form a certain algebra. In general, it is possible to use this algebra to identify a metric. In this thesis, the four first class constraints of bimetric theory are identified and it is shown that their Poisson brackets indeed forms the algebra required by diffeomorphism invariance. However, the metric identified from the algebra turns out not to be unique, but to depend on a choice of variables. Additionally, it need not coincide with the gravitational metric.

The candidate nonlinear partially massless bimetric theory is also investigated in this thesis. It is shown that for this theory, the partially massless symmetry cannot be extended beyond cubic order in the action. This result is generalized to the most general two derivative theory of only two interacting spin-2 fields, showing that such a theory cannot possess the partially massless symmetry beyond cubic order.

Keywords: modified gravity, bimetric theory, spin-2 fields, classical field theory, partially massless symmetry.
CONSTRAINTS AND SYMMETRIES IN THEORIES OF INTERACTING SPIN-2 FIELDS

Anders Lundkvist
Constraints and symmetries in theories of interacting spin-2 fields

Anders Lundkvist
Abstract

The Hassan-Rosen bimetric theory describes two interacting spin-2 fields, one massless and one massive. In this thesis, a complete canonical analysis of this theory is performed in the metric formulation and all constraints are computed. In particular, a secondary constraint, whose existence was in doubt, is shown to exist and evaluated explicitly, bringing the total number of constraints up to six. This, together with general covariance, is enough to eliminate the Boulware-Deser ghost and ensure that the theory propagates the appropriate seven degrees of freedom. The requirement that the constraints are preserved in time leads to a linear relation between the lapse functions of the two metrics. Knowing the explicit form of the ratio of the lapses is necessary for solving initial value problems. The ratio is computed for the special case where the metrics share the same spherical symmetry.

Since the bimetric theory is diffeomorphism invariant, it must contain four first class constraints whose Poisson brackets form a certain algebra. In general, it is possible to use this algebra to identify a metric. In this thesis, the four first class constraints of bimetric theory are identified and it is shown that their Poisson brackets indeed forms the algebra required by diffeomorphism invariance. However, the metric identified from the algebra turns out not to be unique, but to depend on a choice of variables. Additionally, it need not coincide with the gravitational metric.

The candidate nonlinear partially massless bimetric theory is also investigated in this thesis. It is shown that for this theory, the partially massless symmetry cannot be extended beyond cubic order in the action. This result is generalized to the most general two derivative theory of only two interacting spin-2 fields, showing that such a theory cannot possess the partially massless symmetry beyond cubic order.
Sammanfattning


I den här avhandlingen undersöks också en kandidat till en icke-linjär delvis masslös bimetrisk teori. Det visas att denna teori inte är invariant under delvis masslösa gaugetransformationer bortom kubisk ordning i verkan. Detta resultat generaliseras till den mest generella teorin för endast två spinn-2 fält och högst andraderivator i verkan, vilket visar att en sådan teori inte kan ha en delvis masslös symmetri bortom kubisk ordning.

\(^1\)Författaren har inte lyckats hitta en svensk översättning.
List of papers


Acknowledgments

First of all, I would like to express my gratitude to my supervisor Fawad Hassan for your guidance and encouragement, as well as for always providing clear and pedagogical answers to my questions. I would also like to thank my other collaborators: Luis Apolo, Mikica Kocic and Francesco Torsello, without whom this work would not have been possible. A special thanks goes to Luis for explaining xAct. I have greatly enjoyed working with all of you. In addition, I am thankful to Ingemar Bengtsson, Markus Högås, Edvard Mörtsell and Angnis Schmidt-May for helpful discussions and comments.

I am grateful to my mentor Sören Holst for always listening to my problems and giving good advice on how to deal with the difficulties of doing a PhD.

My thanks goes to my past and present office mates: Andrea Chiappo, Jessica Elefant, Julius Engelsøy, Irina Galstyan, Laura Hangard, Markus Högås, Mikica Kocic, Daniel Palm, Tanja Petrushevska and Francesco Torsello. Your company has made every day at work so much better. I would also like to extend my gratitude to all members of the CoPS group and the Oskar Klein Centre for providing such an engaging and inspiring atmosphere.

Thanks to David Andersson for being such a great friend throughout our years as undergrads and PhD students. I am truly grateful for having shared this journey with you. Additionally, I would like to thank all of my other friends for being there for me during these years. Finally, thanks to Mom, Dad and Christer for all of your love and support.
Preface

This doctoral thesis is a thesis by publication consisting of two parts. Part I consists of a comprehensive summary which provides the necessary theoretical background, as well as describes the results of the scientific papers on which the thesis is based. These papers are included in part II. An outline of part I is provided below.

Outline

Chapter 1 provides a brief historical account of the development of bimetric theory, as well as motivation for the work done in the papers. Chapter 2 deals with the analysis of constrained systems, in the Lagrangian as well as the Hamiltonian formalism. In chapter 3, the $3+1$ formulation of general relativity is explained. Chapter 4 describes the consistent linear and nonlinear theories of massive gravity. An overview of the Hassan-Rosen bimetric theory, with emphasis on the parts of the constraint analysis that had been performed prior to Paper I, is provided in chapter 5. In chapter 6 the analysis of the constraints performed in Paper I and Paper II is described. Chapter 7 deals with partial masslessness, describing the candidate nonlinear PM bimetric theory as well as the results of Paper III. The conclusions are summarized and discussed in chapter 8.
Contribution to papers

Paper I
I performed the calculations and wrote the majority of the text. The results were regularly discussed between the authors.

Paper II
I proved the linear relation between the lapses using the canonical formalism and wrote the corresponding sections of the paper. All of the authors discussed the results and contributed to polishing the writing.

Paper III
I performed the calculations showing that the PM gauge symmetry cannot be extended beyond cubic order for a generic quartic action and wrote the corresponding section of the paper. The calculations showing that the gauge symmetry cannot be extended for the candidate PM bimetric theory were originally performed by Luis Apolo and double checked by me.

Material from licentiate thesis
Some of the material in part I comes from the author’s licentiate thesis, Computation of constraints in bimetric theory, 2019 (unpublished) and is reused here with some modifications and updates. Specifically, this applies to large parts of chapter 1, chapters 2 and 3, chapter 5 (except section 5.1), section 6.1 and section 8.1.
Contents

I Comprehensive summary 1

1 Introduction 3
  1.1 Conventions 5

2 Constrained systems 7
  2.1 First class constraints 10

3 General relativity 13
  3.1 The 3+1 formalism 14
  3.2 The diffeomorphism algebra 17

4 Massive gravity 19
  4.1 The Fierz-Pauli theory 19
  4.2 Nonlinear massive gravity 21

5 Bimetric theory 25
  5.1 Proportional background solutions 27
  5.2 Bimetric theory in the 3+1 formalism 29
  5.3 The extra secondary constraint 32
  5.4 The geometrical metric in bimetric theory 34

6 Analysis of constraints 35
  6.1 Results of Paper I 35
     6.1.1 The extra secondary constraint 35
     6.1.2 Degrees of freedom 39
     6.1.3 First class constraints and the geometrical metric 39
  6.2 Results of Paper II 42

7 Partial masslessness 45
  7.1 The linear PM theory 45
Part I

Comprehensive summary
Chapter 1

Introduction

Our present day understanding of physics rests upon two theories. Gravitational phenomena are described by Einstein’s theory of general relativity, and everything else by the standard model of particle physics.

These are both examples of field theories. Fields are characterized by spin and mass according to Wigner’s classification [1]. When the fields are quantized, the field quanta are particles of the corresponding spin and mass. Consistent theories exist for massless and massive fields of spin-0, 1/2 and 1 at both the classical and quantum level. These fields are the building blocks of the standard model, which consists of multiplets of such quantum fields, both massless and massive, interacting with each other. In contrast, general relativity is a theory of a single classical massless spin-2 field, which interacts with fields of lower spin.

It is therefore natural to ask if it is possible to construct consistent theories of massive and interacting spin-2 fields. Such theories can be thought of as modifications of general relativity. Aside from purely theoretical interest, modifying general relativity in this way can be motivated by the observations of dark matter and dark energy, neither of which is understood. Since these phenomena have only been observed through gravitational effects, it may be that they can be explained by modifying our theory of gravity.

A theory of a single massive spin-2 field is referred to as a theory of massive gravity, while the term bimetric theory is used for a theory of two interacting spin-2 fields. Attempts to construct these types of theories go back all the way to the 1930s, but have generally been difficult due to the presence of ghost instabilities, i.e. degrees of freedom with negative kinetic energy.

The first theory of massive gravity was constructed in 1939 by Fierz and Pauli at the linear level, by adding a mass term to the linearized action of
general relativity [2]. By giving this mass term a specific form they ensured
that this theory is ghost free at the linear level and propagates the appropriate five degrees of freedom. However, in the limit where the graviton mass approaches zero, this theory does not reproduce general relativity. This is known as the vDVZ (van Dam, Veltman and Zakharov) discontinuity [3, 4]. Observations would therefore rule out even a very small graviton mass, had it not been for the Vainshtein mechanism, which allows a nonlinear completion of the Fierz-Pauli theory to avoid the vDVZ discontinuity [5]. Vainshtein’s discovery was that the Fierz-Pauli linear approximation becomes invalid within a certain radius of a massive source. As the graviton mass approaches zero, this radius becomes infinite. This means that one must consider a nonlinear theory.

However, in 1972 it was argued by Boulware and Deser that any nonlinear completion of the Fierz-Pauli theory necessarily contains a ghost instability [6, 7]. It was not until 2010 that a loophole in this argument was found by de Rham, Gabadadze and Tolley (dRGT), who were able to construct a nonlinear theory of massive gravity and show that it is free of the Boulware-Deser ghost in a particular decoupling limit [8, 9]. In order to construct a nonlinear mass term for the metric $g_{\mu\nu}$, its indices must be contracted with some reference metric in order to create scalar terms in the Lagrangian. dRGT take this reference metric to be the Minkowski metric.

Hassan and Rosen (HR) reformulated the dRGT theory and generalized it to an arbitrary (but still nondynamical) reference metric $f_{\mu\nu}$ in [10]. Using a Hamiltonian analysis, they were able to prove that the theory is ghost free away from the decoupling limit [11–13]. In this analysis, a pair of constraints among the equations of motion is responsible for eliminating the Boulware-Deser ghost mode. HR also provided the reference metric $f_{\mu\nu}$ with dynamics, thereby generalizing the massive gravity theory to a bimetric theory, describing interactions between a massless and a massive spin-2 field [14]. This bimetric theory was defined more rigorously in [15]. Unlike massive gravity, HR bimetric theory is covariant. It was argued in [13] that the bimetric theory also has a sufficient number of constraints to be free of the Boulware-Deser ghost, but this was disputed in [16,17], who argued that one of the constraints did not exist. This provides motivation for one of the aims of this thesis, which is to explicitly compute the constraints of bimetric theory and show that they are sufficient to render the theory ghost free. The explicit form of all constraints is also necessary if one wishes to solve initial value problems in bimetric theory, providing further motivation for their computation.

An additional aim for this thesis is inspired by an observation made by
Bengtsson and Peldán in [18]. Specifically, they note that in a covariant field theory a metric can always be identified from the Poisson algebra of the constraints. In this thesis, the Poisson algebra of the constraints is therefore computed for the HR bimetric theory, and the possibility of identifying a metric is investigated.

The final subject of this thesis concerns theories of a partially massless spin-2 field. It was shown in [19] that the linear Fierz-Pauli theory of massive gravity on an Einstein-de Sitter background has an extra gauge symmetry if the mass of the spin-2 field satisfies $m_{FP}^2 = \frac{2}{3} \Lambda$, where $\Lambda$ is the cosmological constant. The theory then describes a so called partially massless spin-2 field, with four degrees of freedom rather than five. This naturally raises the question if this symmetry can be extended to nonlinear theories of spin-2 fields. A unique candidate for a nonlinear PM theory among the class of HR bimetric theories was found in [20]. Investigating if this theory possesses a nonlinear PM symmetry is the final aim of this thesis.

1.1 Conventions

Spacetime indices are denoted by Greek letters and spatial indices by Latin letters (except in chapter 2). Since we consider three dimensions of space and one of time, $\mu, \nu \in \{0, 1, 2, 3\}$ and $i, j \in \{1, 2, 3\}$. We make use of the Einstein summation convention, where repeated indices are summed over, e.g.

$$v^\mu w_\mu = \sum_{\mu=0}^{3} v^\mu w_\mu. \quad (1.1)$$

The metric signature is “mostly plus”, i.e. $n_{\mu\nu} = \text{diag}(-1, +1, +1, +1)$. The symbol $\approx$ is used to denote weak equality, that is an equality which only holds when certain constraints are imposed.
Chapter 2

Constrained systems

Conservative physical systems can be described by an action functional

$$S = \int dt L(q_i, \dot{q}_i),$$

(2.1)

where the Lagrangian $L$ depends on some generalized coordinates $q_i(t)$, with $i \in \{1, \ldots, N\}$, and their time derivatives$^1$. By demanding that the variation of the action vanishes for arbitrary variations of $q_i$ that keeps the endpoints fixed, we obtain equations of motion for $q_i$, which are known as the Euler-Lagrange equations

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) = \frac{\partial L}{\partial q_i}. \quad (2.2)$$

Suppose that only some of the $q_i$ are dynamical, in the sense that the Lagrangian depends on their time derivatives. Denote the dynamical variables by $q_m$, where $m \in \{1, \ldots, M\}$, and the nondynamical ones by $Q_r$, where $r \in \{M + 1, \ldots, N\}$, so that we have a Lagrangian $L(q_m, \dot{q}_m, Q_r)$. Then the equations of motion are

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_m} \right) = \frac{\partial L}{\partial q_m}, \quad \frac{\partial L}{\partial Q_r} = 0. \quad (2.3)$$

Some of the nondynamical equations (possibly all or none) may be solved for some of the $Q_r$, while the rest impose constraints on the dynamical variables. The equations of motion (2.3) contain all information about the dynamics and constraints of the system.

$^1$This can be generalized to field theories, where $q_i$ are also functions of space.
The equations of motion may also be derived in the Hamiltonian formulation. In this setup, one first defines the canonical momenta $p^i$ of all generalized coordinates as $p^i = \partial L / \partial \dot{q}_i$. Note that

$$P^r = \frac{\partial L}{\partial \dot{Q}_r} = 0$$

(2.4)
since the Lagrangian is independent of $\dot{Q}_r$. The equations (2.4) are called primary constraints. We can now define the Hamiltonian of the system as

$$H(q_m, p^m, Q_r) = p^i \dot{q}_i - L(q_m, \dot{q}_m, Q_r).$$

(2.5)

However, in order to ensure that the Hamiltonian equations of motion obey the primary constraints, the full Hamiltonian is defined as

$$H' = H + \lambda_r P^r,$$

(2.6)

where $\lambda_r$ are Lagrange multipliers whose equations of motion enforce (2.4). The time derivative of a quantity $A$ is given by its Poisson bracket with the Hamiltonian, $\dot{A} = \{A, H\}$, where the Poisson bracket is defined as

$$\{A, B\} = \frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p^j} - \frac{\partial A}{\partial p^j} \frac{\partial B}{\partial q_i}.$$

(2.7)

Hamilton’s equations of motion are therefore

$$\dot{q}_m = \{q_m, H'\} = \frac{\partial H'}{\partial p^m},$$

(2.8)

$$\dot{p}^m = \{p^m, H'\} = -\frac{\partial H'}{\partial q_m},$$

(2.9)

$$\dot{Q}_r = \{Q_r, H'\} = \lambda_r,$$

(2.10)

$$\dot{P}^r = \{P^r, H'\} = -\frac{\partial H'}{\partial Q_r}.$$  

(2.11)

Since the primary constraints (2.4) should hold at all times, we must impose $\dot{P}^r = 0$. This leads to a set of new constraint equations,

$$C^r = \frac{\partial H'}{\partial Q_r} = 0,$$

(2.12)

2In general, primary constraints can be more complicated relations among the phase space variables that arise from the definition of momenta. We make use of the $(q_m, Q_r)$ basis to ensure that the primary constraints are of the simple form (2.4).
which are called secondary constraints. They are equivalent to the second set of equations in (2.3).

The secondary constraints must also hold at all times, hence we impose
\[
\dot{C}_r = \{C^r, H\} = \{C^r, H\} + \lambda_s\{C^r, P^s\} \approx 0, \tag{2.13}
\]
where the symbol \(\approx\) is used to denote an equality that holds on the constraint surface specified by (2.4) and (2.12). This is referred to as a weak equality. Note that the constraints are imposed after the Poisson brackets have been evaluated off the constraint surface. Hence, a bracket involving \(C^r\) does not necessarily vanish on the constraint surface just because \(C^r\) does. If the matrix \(\{C^r, P^s\}\) is invertible, these equations can be solved for the Lagrange multipliers \(\lambda_r\). If it’s singular, some of the Lagrange multipliers are undetermined and some linear combinations of \(\{C^r, H\}\) must vanish on the constraint surface. This condition may be satisfied identically or it may be satisfied provided that
\[
\Theta^A(q_m, p^m, Q_r) = 0, \tag{2.14}
\]
where \(\Theta^A\) are some functions of the phase space variables. These equations then provide additional secondary constraints on the dynamical variables. Since these constraints must also be preserved, we impose \(\dot{\Theta}^A = \{\Theta^A, H'\} \approx 0\). As before, some of these equations may be solvable for some of the Lagrange multipliers while others may give rise to even more secondary constraints. If there are more constraints the procedure is repeated. Note that this process cannot continue indefinitely, since the total number of constraints cannot exceed the number of phase space variables. In the end, some Lagrange multipliers (possibly all or none) will be determined and we will have some number of secondary constraints.

The Lagrangian and Hamiltonian equations of motion are equivalent to one another. While it’s easy to see that (2.12) is equivalent to the second set of equations in (2.3), it’s not as trivial to see where the constraints (2.14) appear in the Lagrangian formulation. In fact, they are hidden among the first set of equations in (2.3).

It is also possible to derive the Hamiltonian equations of motion by variation of the action. In order to do this, the action is written in terms of phase space variables
\[
S = \int dt \left[ p^m \dot{q}_m - H(q_m, p^m, Q_r) \right]. \tag{2.15}
\]
Varying this action with respect to \(q_m, p^m\) and \(Q_r\) yields the equations of motion
\[
\dot{q}_m = \frac{\partial H}{\partial p^m}, \quad \dot{p}^m = -\frac{\partial H}{\partial q_m}, \quad C^r = \frac{\partial H}{\partial Q_r} = 0, \tag{2.16}
\]
which are called secondary constraints. They are equivalent to the second set of equations in (2.3).
which are completely equivalent to (2.8), (2.9) and (2.12). They are also equivalent to the Lagrangian equations of motion in (2.3) and the definition of the momenta \( p^m \). Since the equations (2.16) are derived using the action principle, they are valid at all times. All information about dynamics and constraints is contained in them. However, in order to disentangle the constraints, that are of the type (2.14), it is convenient to write down the preservation of \( C^r \) in terms of Poisson brackets

\[
\dot{C}^r = \{C^r, H\} + \frac{\partial C^r}{\partial Q_s} \dot{Q}_s \approx 0, \tag{2.17}
\]

which is equivalent to (2.13). Note that here, the Poisson bracket is defined as

\[
\{A, B\} = \frac{\partial A}{\partial q_m} \frac{\partial B}{\partial p^m} - \frac{\partial A}{\partial p^m} \frac{\partial B}{\partial q_m}, \tag{2.18}
\]

which is why the second term in (2.17) is needed to take the time dependence of \( Q_r \) into account. If \( C^r = 0 \) can be solved for \( Q_r \), equation (2.17) can be solved for \( \dot{Q}_r \) (this is analogous to the situation where (2.13) is solved for a \( \lambda_r \)). However, for those equations \( C^r = 0 \) which cannot be solved for any \( Q_r \), the vanishing of \( \dot{C}^r \) enables us to isolate some equations (2.14) among the dynamical equations. These can either be solved for some of the \( Q_r \) or impose constraints\(^3\) on the dynamical variables \( q_m \) and \( p^m \). In the former case, the equations \( \dot{\Theta}^A = 0 \) can be solved for \( Q_r \), while in the latter they help isolate new constraints on the dynamical variables or algebraic equations for \( Q_r \). If there are more constraints the process is repeated. This is the formalism we use when computing the constraints of bimetric theory.

### 2.1 First class constraints

A natural way of classifying the constraints arises from the notion of first class and second class functions. A function \( F \) of the canonical variables is referred to as a first class function if its Poisson bracket with every constraint vanishes weakly. That is, if \( \Phi_a = 0 \) are all the constraints, then

\[
\{F, \Phi_a\} \approx 0, \tag{2.19}
\]

\(^3\)In the standard terminology, all nondynamical equations are referred to as constraints. In the rest of the thesis, only those nondynamical equations that can eliminate some dynamical variables in terms of other dynamical variables are called constraints, since this is the type that is needed to eliminate the ghost in bimetric theory.
if $F$ is first class. Functions which are not first class are referred to as second class. This allows us to classify the constraints themselves as first class or second class. Note that the Poisson bracket of two first class functions is itself first class [21]. From this it follows that the Poisson brackets of the first class constraints with each other form a closed algebra, i.e. the Poisson bracket of two first class constraints is itself a first class constraint.

It was conjectured by Dirac that all first class constraints are generators of gauge transformations [22]. While this is not always true (counterexamples can be constructed) it can be proven to hold under certain conditions, which apply to the theories discussed here [21].
Chapter 3
General relativity

In general relativity (GR), spacetime is a four dimensional Lorentzian manifold with a metric tensor $g_{\mu\nu}$. The curvature of such a manifold is given by the Riemann curvature tensor, $R^{\mu}_{\nu\alpha\beta}$, which can be defined by

$$R^{\mu}_{\nu\alpha\beta}v^\nu = [\nabla_\alpha, \nabla_\beta]v^\mu,$$  \hspace{1cm} (3.1)

where $v^\mu$ is an arbitrary vector and $\nabla_\mu$ is the covariant derivative compatible with $g_{\mu\nu}$, meaning that $\nabla_\alpha g_{\mu\nu} = 0$. The covariant derivative of a vector is given by

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\nu_{\mu\lambda}v^\lambda,$$  \hspace{1cm} (3.2)

where $\Gamma^\nu_{\mu\lambda}$ are the Christoffel symbols, which can be expressed as

$$\Gamma^\alpha_{\mu\nu} = \frac{1}{2}g^{\alpha\beta}(\partial_\mu g_{\nu\beta} + \partial_\nu g_{\mu\beta} - \partial_\beta g_{\mu\nu}).$$  \hspace{1cm} (3.3)

Equation (3.3) follows from the compatibility of $\nabla_\mu$ with the metric and the condition that there is no torsion. The Riemann tensor can be written in terms of the Christoffel symbols as

$$R^{\mu}_{\nu\alpha\beta} = \partial_\alpha \Gamma^\mu_{\beta\nu} - \partial_\beta \Gamma^\mu_{\alpha\nu} + \Gamma^\mu_{\alpha\lambda} \Gamma^\lambda_{\beta\nu} - \Gamma^\mu_{\beta\lambda} \Gamma^\lambda_{\alpha\nu}. $$ \hspace{1cm} (3.4)

By contracting the indices of the Riemann tensor, one can define the Ricci tensor

$$R_{\mu\nu} = R^\alpha_{\mu\alpha\nu},$$ \hspace{1cm} (3.5)

and the Ricci scalar

$$R = g^{\mu\nu}R_{\mu\nu}. $$ \hspace{1cm} (3.6)

The latter can be used to construct the action

$$S = \int d^4x M_p^2 \sqrt{-g}(R - 2\Lambda) + S_M,$$ \hspace{1cm} (3.7)
where \( g = \det (g_{\mu\nu}) \), \( M_P \) is the reduced Planck mass\(^1\), \( \Lambda \) is the cosmological constant and \( S_M \) is the action for the matter fields. The gravitational part of \( S \) is known as the Einstein-Hilbert action. Varying \( S \) with respect to the metric, the action principle yields the Einstein field equations

\[
G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{1}{2M_P^2} T_{\mu\nu},
\]

(3.8)

where \( G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \) denotes the Einstein tensor and \( T_{\mu\nu} \) is the stress-energy tensor of matter, defined by

\[
T_{\mu\nu} = -\frac{2}{\sqrt{-g}} \delta S_M \delta g^{\mu\nu},
\]

(3.9)

where \( \delta \frac{\delta}{\delta g_{\mu\nu}} \) denotes the functional derivative with respect to \( g_{\mu\nu} \). The Einstein tensor satisfies the Bianchi identities

\[
\nabla^\mu G_{\mu\nu} = 0,
\]

(3.10)

which, combined with the Einstein field equations imply \( \nabla^\mu T_{\mu\nu} = 0 \), i.e. local energy-momentum conservation\(^2\).

If the cosmological constant is zero and no matter Lagrangian is included, the Einstein field equations reduce to

\[
R_{\mu\nu} = 0,
\]

(3.11)

which are known as the vacuum field equations.

### 3.1 The 3+1 formalism

In order to analyze the constraints of general relativity, it is convenient to use the 3+1 formalism. One way of doing this was formulated by Arnowitt, Deser and Misner (ADM) in [23]. In this formalism, spacetime is foliated into a family of spacelike hypersurfaces. These hypersurfaces are equipped with an induced metric \( \gamma_{ij} \), which is positive definite. The spacetime metric can then be expressed as

\[
g_{\mu\nu} = \left( -N^2 + N_i N^i N_j \gamma_{ij} \right),
\]

(3.12)

\(^1\)The reduced Planck mass is here defined as \( M_P^2 = 16\pi G_N \), where \( G_N \) is Newton’s constant, ensuring the correct factor in the Einstein field equations.

\(^2\)Local energy-momentum conservation also follows from the covariance of the matter action.
and its inverse as

\[ g^{\mu\nu} = \frac{1}{N^2} \begin{pmatrix} -1 & N^j \\ N^i & N^2 \gamma^{ij} - N^i N^j \end{pmatrix}. \]  

(3.13)

Here, \( \gamma^{ij} \) is the inverse of \( \gamma_{ij} \). \( N \) is called the lapse function and \( N^i \) the shift vector. The index of \( N^i \) is raised and lowered using the metric \( \gamma_{ij} \).

We can now express the Einstein-Hilbert action (3.7) in terms of \( \gamma^{ij} \), \( N^i \) and \( N \), as well as their derivatives. It turns out that the action is independent of time derivatives of the lapse and shift, meaning that they are nondynamical. The dynamical variables are the components of \( \gamma_{ij} \) and their canonical momenta \( \pi^{ij} \), which are given by

\[ \pi^{ij} \equiv \frac{\partial L}{\partial \dot{\gamma}^{ij}} = \frac{M_\text{P}^2 \sqrt{\gamma}}{2N} \left[ \left( \gamma^{ik} \gamma^{jl} - \gamma^{ij} \gamma^{kl} \right) \dot{\gamma}_{kl} \\
- \nabla^i N^j - \nabla^j N^i + 2\gamma^{ij} \nabla^k N_k \right], \]  

(3.14)

where \( \nabla \) is the covariant derivative compatible with \( \gamma_{ij} \). From this it follows that \( \pi^{ij} \) is not a tensor, but a tensor density of weight +1. This means that under a general coordinate transformation, it behaves like a tensor except that it is also multiplied by the absolute value of the Jacobian determinant of the coordinate transformation. Note that \( \pi^{ij} \) can be expressed as

\[ \pi^{ij} = M_\text{P}^2 \sqrt{\gamma} \left( K^{ij} - \gamma^{ij} K \right), \]  

(3.15)

where \( K^{ij} \) is the extrinsic curvature of the spacelike hypersurface\(^3\). The Einstein-Hilbert action (without matter coupling and with \( \Lambda = 0 \)) can then be expressed as

\[ S = \int d^4 x \left( \pi^{ij} \dot{\gamma}_{ij} + NR^{(g)} + N^i R_i^{(g)} \right), \]  

(3.16)

where \( R^{(g)} \) and \( R_i^{(g)} \) are given by\(^4\)

\[ R^{(g)} = M_\text{P}^2 \sqrt{\gamma} R(\gamma) + \frac{1}{M_\text{P}^2 \sqrt{\gamma}} \left( \frac{1}{2} \pi^2 - \pi^{ij} \pi_{ij} \right), \]  

(3.17)

\[ R_i^{(g)} = 2 \sqrt{\gamma} \nabla^j \pi_{ij}, \]  

(3.18)

\(^3\)Sometimes \( K^{ij} \) are used as variables in the 3 + 1 formulation rather than \( \pi^{ij} \). However, for our purposes \( \pi^{ij} \) is more convenient.

\(^4\)Since \( \pi^{ij} \) is a tensor density of weight +1, its covariant derivative is given by \( \nabla_k \pi^{ij} = \partial_k \pi^{ij} + \Gamma^i_{kl} \pi^{lj} + \Gamma^j_{kl} \pi^{il} - \Gamma^l_{kl} \pi^{ij} = \sqrt{\gamma} \nabla_k \left( \frac{\pi^{ij}}{\sqrt{\gamma}} \right) \).
where $R(\gamma)$ is the Ricci scalar of the spatial metric $\gamma_{ij}$. Variation of the action (3.16) with respect to $\pi^{ij}$, $\gamma_{ij}$, $N$ and $N^i$ yields the following equations of motion

$$\dot{\gamma}_{ij} = \frac{2N}{M_P^2 \sqrt{\gamma}} (\pi_{ij} - \frac{1}{2} \gamma_{ij} \pi) + \nabla_i N_j + \nabla_j N_i,$$

(3.19)

$$\dot{\pi}_{ij} = \frac{N \gamma^{ij}}{2M_P^2 \sqrt{\gamma}} \left( \pi^{mn} \pi_{mn} - \frac{1}{2} \pi^2 \right) - \frac{2N}{M_P^2 \sqrt{\gamma}} \left( \pi^{ik} \pi_k^j - \frac{1}{2} \pi \pi^{ij} \right)
- M_P^2 N \sqrt{\gamma} \left( R^{ij}(\gamma) - \frac{1}{2} \gamma^{ij} R(\gamma) \right) + \nabla_k \left( N^k \pi^{ij} \right)
- 2\pi^{k(i} \nabla_k N^{j)} + M_P^2 \sqrt{\gamma} \left( \nabla^i \nabla^j N - \gamma^{ij} \nabla^k \nabla_k N \right),$$

(3.20)

$$R(\gamma) = 0, \quad R_i^{(g)} = 0.$$  

(3.21)

The equation of motion for $\gamma_{ij}$, (3.19) is equivalent to the definition of $\pi^{ij}$ in (3.14), whereas (3.20) and (3.21) are equivalent to the the vacuum field equations (3.11). Note that since $N$ and $N^i$ appear in the action as Lagrange multipliers, their equations of motion become constraints on $\gamma_{ij}$ and $\pi^{ij}$.

As explained in chapter 2, the time derivatives of the constraints (3.21) vanish on the constraint surface, since the constraints are valid at all times. Thus, it is necessary to compute these derivatives to see if their vanishing gives rise to any additional constraints, or to equations that can be solved for the nondynamical variables. Since $R(\gamma)$ and $R_i^{(g)}$ are independent of lapse and shift, their time derivatives are simply given by their Poisson brackets with the Hamiltonian,

$$\dot{R}(\gamma) = \{R(\gamma), H\}, \quad \dot{R}_i^{(g)} = \{R_i^{(g)}, H\},$$

(3.22)

where the Poisson bracket is defined by

$$\{A, B\} = \int d^3 z \left( \frac{\delta A}{\delta \gamma_{mn}(z)} \frac{\delta B}{\delta \pi^{mn}(z)} - \frac{\delta A}{\delta \pi^{mn}(z)} \frac{\delta B}{\delta \gamma_{mn}(z)} \right),$$

(3.23)

and the Hamiltonian is given by

$$H = -\int d^3 x \left( N R(\gamma) + N^i R_i^{(g)} \right),$$

(3.24)

as can be seen from (3.16). Since the Hamiltonian density is a linear combination of $R(\gamma)$ and $R_i^{(g)}$, we require explicit expressions for their Poisson
brackets with each other. These are given by [24],

\[
\{ R^{(g)}(x), R^{(g)}(y) \} = \left[ \gamma^{ij}(x) R^{(g)}_j(x) + \gamma^{ij}(y) R^{(g)}_j(y) \right] \frac{\partial}{\partial y^i} \delta^3(x, y),
\]

\[
\{ R^{(g)}(x), R^{(g)}_i(y) \} = R^{(g)}(y) \frac{\partial}{\partial y^i} \delta^3(x, y),
\]

\[
\{ R^{(g)}_i(x), R^{(g)}_j(y) \} = R^{(g)}_j(x) \frac{\partial}{\partial y^i} \delta^3(x, y) + R^{(g)}_i(y) \frac{\partial}{\partial y^j} \delta^3(x, y),
\]

where \( \delta^3(x, y) \) is the Dirac delta function in three dimensions. Since these brackets are proportional to the constraints themselves, it follows that \( \dot{R}^{(g)} \) and \( \dot{R}^{(g)}_i \) vanish provided that (3.21) holds. This implies that there are no additional constraints and that \( N \) and \( N^i \) are left undetermined (unless a gauge is chosen, as explained below).

We are now in a position to count the number of degrees of freedom in GR. \( \gamma^{ij} \) and \( \pi^{ij} \) each have six components, for a total of twelve dynamical variables. In addition, we have the four non-dynamical variables \( N \) and \( N^i \). The constraints (3.21) can be solved for four components of \( \gamma^{ij} \) and \( \pi^{ij} \). An additional four components can be eliminated using general coordinate transformations. Their equations of motion can then be solved for \( N \) and \( N^i \). We are left with four dynamical variables, i.e. two propagating degrees of freedom, corresponding to a massless spin-2 field. The number of degrees of freedom is not affected by adding a cosmological constant to the Einstein-Hilbert action or by minimal coupling matter to \( g_{\mu\nu} \). The reason is that both the cosmological constant term and the matter Lagrangian are linear in \( N \) and \( N^i \), which means that we still get four constraints. While they have extra terms compared to (3.21), the structure of the algebra (3.25) remains the same, and hence the counting above is still valid.

### 3.2 The diffeomorphism algebra

From (3.25) it follows that all four constraints of GR are first class, since the Poisson bracket of any constraint with any other constraint vanishes weakly. As was remarked in the previous section, adding a cosmological constant or a minimal coupling to matter does not change the structure of the constraint algebra (3.25). In fact, the algebra is even more general, any covariant field theory contains four first class constraints whose Poisson brackets satisfy (3.25), once all second class constraints have been imposed [25, 26]. The first class constraints are the generators of diffeomorphisms, specifically, \( R^{(g)}_i \) generates diffeomorphisms on the spacelike
hypersurface and $\mathcal{R}(g)$ generates diffeomorphisms orthogonal to the hypersurface. Moreover, once any eventual second class constraints have been imposed, the Hamiltonian density of such a theory is of the form

$$\mathcal{H} = -M\hat{\mathcal{R}} - M^i\hat{\mathcal{R}}_i,$$  \hspace{1cm} (3.26)

where $\hat{\mathcal{R}}$ and $\hat{\mathcal{R}}_i$ are the first class constraints and $M$ and $M^i$ are some Lagrange multipliers.

The inverse of the spatial metric, $\gamma^{ij}$, appears in the first line of (3.25). This means that for any covariant field theory, some rank 2 tensor will appear in the corresponding place. From the analysis of [26] it follows that this tensor is to be interpreted as the spatial part of the geometrical metric (i.e. the metric which defines the norm of vectors). Furthermore, comparing (3.26) to the Hamiltonian of GR in (3.24), the Lagrange multipliers $M$ and $M^i$ is to be interpreted as the lapse and shift. Combined with the spatial metric appearing in the algebra, they yield the full geometrical metric. A consequence of this, pointed out in [18], is that the diffeomorphism algebra can be used to identify a metric in a covariant field theory where it is not evident from the start what that metric is. This was first done in [18, 27] for generalizations of Ashtekar’s [28] canonical formulation of gravity.
Chapter 4

Massive gravity

We now turn our attention to massive gravity, i.e. theories describing a single massive spin-2 field. Such theories are constructed by adding a mass term for the metric to the action of a massless field. We will start by describing the linear theory before turning our attention to the nonlinear case.

4.1 The Fierz-Pauli theory

General relativity can be linearized around a flat background by writing the metric as

$$g_{\mu \nu} = \eta_{\mu \nu} + h_{\mu \nu},$$

(4.1)

where $\eta_{\mu \nu}$ is the Minkowski metric and $h_{\mu \nu}$ is a small perturbation. To quadratic order in $h_{\mu \nu}$, the Einstein-Hilbert action without matter coupling is

$$S = \frac{M_P^2}{2} \int d^4x h_{\mu \nu} \mathcal{E}^{\mu \nu \rho \sigma} h_{\rho \sigma},$$

(4.2)

where the linearized Einstein operator $\mathcal{E}^{\mu \nu \rho \sigma}$ is given by

$$\mathcal{E}^{\mu \nu \rho \sigma} = -\frac{1}{2} \left( \eta^{\rho \mu} \partial^{\sigma} \partial^{\nu} + \eta^{\rho \nu} \partial^{\sigma} \partial^{\mu} - \eta^{\rho \mu} \eta^{\sigma \nu} \partial^2 - \eta^{\rho \sigma} \partial^{\mu} \partial^{\nu} - \eta^{\mu \nu} \partial^{\rho} + \eta^{\mu \nu} \eta^{\rho \sigma} \partial^2 \right).$$

(4.3)

The indices are here raised and lowered using the background metric $\eta_{\mu \nu}$. This action is invariant under linearized diffeomorphisms, given by

$$h_{\mu \nu} \rightarrow h_{\mu \nu} + 2 \partial_{(\mu} \xi_{\nu)},$$

(4.4)
where $\zeta_\mu$ is a local gauge parameter. In order to construct a linear theory of a massive spin-2 field, a mass term should be added to this action. The most general such mass term is given by

$$S_{\text{FP}} = \frac{M_{\text{P}}^2}{2} \int d^4x \left[ \varphi_{\mu\nu} \varepsilon^{\mu\nu\rho\sigma} \varphi_{\rho\sigma} - \frac{m_{\text{FP}}^2}{2} \left( \varphi^{\mu\nu} \varphi_{\mu\nu} - a \varphi^2 \right) \right],$$  \hspace{1cm} (4.5)

The mass of the spin-2 field is denoted by $m_{\text{FP}}$ while $a$ is a dimensionless constant. Here, we use $\varphi_{\mu\nu}$ to denote the linear massive field, since we wish to reserve the notation $h_{\mu\nu}$ for the massless fluctuation. The trace is taken with respect to the background metric, $\varphi = \eta^{\mu\nu} \varphi_{\mu\nu}$. Note that adding this mass term breaks the gauge invariance (4.4). It was discovered by Fierz and Pauli that the only consistent choice for the mass term is given by $a = 1$, as can be seen in the following way [2]. Variation of (4.5) with respect to $\varphi_{\mu\nu}$ yields the equations of motion

$$\varepsilon^{\mu\nu\rho\sigma} \varphi_{\rho\sigma} - \frac{m_{\text{FP}}^2}{2} \left( \varphi^{\mu\nu} - a \eta^{\mu\nu} \varphi \right) = 0.$$  \hspace{1cm} (4.6)

Taking the divergence of these equations and using the linearized version of the Bianchi identities (3.10), $\partial^\mu \varepsilon^{\mu\nu\rho\sigma} \varphi_{\rho\sigma} = 0$, results in four constraint equations

$$\partial^\mu \varphi_{\mu\nu} - a \partial^\nu \varphi = 0.$$  \hspace{1cm} (4.7)

Additionally, the trace of (4.6) is

$$\partial^\nu \left( \partial^\mu \varphi_{\mu\nu} - \partial^\nu \varphi \right) - \frac{m_{\text{FP}}^2}{2} \left( 1 - 4a \right) \varphi = 0.$$  \hspace{1cm} (4.8)

If $a = 1$, this reduces to a fifth constraint,

$$\varphi = 0.$$  \hspace{1cm} (4.9)

when (4.7) are imposed. The constraints eliminate five of the ten independent components of $\varphi_{\mu\nu}$, leaving five propagating degrees of freedom, which corresponds to a massive spin-2 field. For $a \neq 1$, (4.8) is a dynamical equation of motion and the theory has six degrees of freedom. Additionally, this extra degree of freedom is a so called ghost, i.e. a mode with negative kinetic energy. A theory which contains a ghost is unphysical, so in order for the linear theory of massive gravity to be consistent the mass term must be of the Fierz-Pauli form, with $a = 1$.

The Fierz-Pauli theory describes a massive spin-2 field propagating on a flat background. This can be generalized to a generic Einstein background\(^1\)

\(^1\)For an Einstein spacetime, $G_{\mu\nu} + \Lambda g_{\mu\nu} = 0$. 

20
\(\bar{g}_{\mu\nu}\), resulting in the action
\[
S_{\text{FP}} = \frac{M_P^2}{2} \int d^4x \sqrt{-\bar{g}} \left[ \varphi_{\mu\nu} \bar{E}^{\mu\nu\rho\sigma} \varphi_{\rho\sigma} + \Lambda \left( \varphi^{\mu\nu} \varphi_{\mu\nu} - \frac{1}{2} \varphi^2 \right) \right. \\
\left. - \frac{m_{\text{FP}}^2}{2} \left( \varphi^{\mu\nu} \varphi_{\mu\nu} - \varphi^2 \right) \right].
\] (4.10)

Here the indices are raised and lowered using the background metric \(\bar{g}_{\mu\nu}\) and the operator \(\bar{E}^{\mu\nu\rho\sigma}\) is given by
\[
\bar{E}^{\mu\nu\rho\sigma} = -\frac{1}{2} \left( \bar{g}^{\rho\sigma} \nabla_\mu \nabla_\nu + \bar{g}^{\rho\mu} \nabla_\sigma \nabla_\nu - \bar{g}^{\rho\mu} \bar{g}^{\sigma\nu} \nabla^2 - \bar{g}^{\rho\sigma} \nabla_\mu \nabla_\nu \\
- \bar{g}^{\mu\nu} \nabla_\rho \nabla_\sigma + \bar{g}^{\mu\nu} \bar{g}^{\rho\sigma} \nabla^2 \right),
\] (4.11)

where \(\nabla_\mu\) is the covariant derivative compatible with \(\bar{g}_{\mu\nu}\). This theory has some features not shared by the flat space theory. These will be discussed in chapter 7.

4.2 Nonlinear massive gravity

In order to generalize the Fierz-Pauli theory to the nonlinear level, a mass term for the metric \(g_{\mu\nu}\) should be added to the nonlinear Einstein-Hilbert action (3.7). Such a mass term should be a scalar function of the metric which does not contain any derivatives. However, if the indices of the metric are contracted using the metric itself, such a term reduces to a cosmological constant, since \(g_{\mu\nu} g^{\mu\nu} = 4\). This means that a mass term can not be constructed using only the metric. It is therefore necessary to introduce an additional symmetric rank-2 tensor field, \(f_{\mu\nu}\), and use that to contract the indices of \(g_{\mu\nu}\). This field is referred to as the reference metric and is taken to be a fixed background metric in massive gravity. A generic nonlinear theory of massive gravity is then described by the action
\[
S_{\text{MG}} = \int d^4x \sqrt{-g} \left[ \right. M_P^2 R(g) - m^4 V(g, f) \left. \right] + S_M,
\] (4.12)

where \(V(g, f)\) is some scalar function of the metric and the reference metric. A necessary condition for this to be ghost free is that it reduces to the Fierz-Pauli theory when linearized. It was however argued in [6, 7] that even in this case the theory will propagate six degrees of freedom at the nonlinear level, with the sixth being a ghost. This is referred to as the Boulware-Deser
ghost. The argument is based on an ADM analysis of massive gravity. In terms of the 3 + 1 variables, action (4.12) takes the form
\[
S_{\text{MG}} = \int d^4x \left( \pi^{ij} \dot{\gamma}^{ij} + N^i \mathcal{R}^{(g)} + N^i \mathcal{R}^{(g)}_i - m^4 V(g, f) \right). \tag{4.13}
\]
As in general relativity, the lapse and shift of the metric \( g_{\mu\nu} \) will be non-dynamical variables. However, they will in general not be Lagrange multipliers, since \( V(g, f) \) is not a linear function of them. Their equations of motion will therefore be algebraic equations for the lapse and shift themselves, rather than constraints on the dynamical variables. In addition, the mass term breaks general covariance due to the fixed reference metric \( f_{\mu\nu} \), which means that no degrees of freedom can be removed using general coordinate transformations. None of the twelve components of \( \gamma^{ij} \) and \( \pi^{ij} \) can therefore be removed, and we are left with six degrees of freedom.

It is however possible to construct a nonlinear massive gravity theory which is ghost free even at the nonlinear level by giving the mass term a specific form, resulting in the action
\[
S = \int d^4x \left( M_5^2 \sqrt{-g} R(g) - 2m^4 \sqrt{-g} \sum_{n=0}^3 \beta_n e_n(S) \right) + S_M. \tag{4.14}
\]
Here, the matrix \( S \) is defined as the square root of the matrix \( g^{-1} f \), i.e. by \( S^\mu_\lambda S^\lambda_\nu = g^{\mu\lambda} f_{\lambda\nu} \). The functions \( e_n(S) \) are elementary symmetric polynomials of the eigenvalues of \( S \). Since \( S \) is a \( 4 \times 4 \) matrix, these are defined as
\[
e_0(S) = 1, \quad e_1(S) = [S], \quad e_2(S) = \frac{1}{2} ([S]^2 - [S^2]), \quad e_3(S) = \frac{1}{6} ([S]^3 - 3[S][S^2] + 2[S^3]), \quad e_4(S) = \frac{1}{24} ([S]^4 - 6[S]^2[S^2] + 3[S^2]^2 + 8[S][S^3] - 6[S^4]) = \det S, \quad e_k(S) = 0 \quad \text{for} \quad k > 4,
\]
where \([S] = \text{Tr}(S)\). The constants \( \beta_n \) are free parameters of the theory. This theory, often referred to as dRGT (de Rham, Gabadadze and Tolley) massive gravity, was first derived in [8, 9], following the approach of [29], with \( f_{\mu\nu} = \eta_{\mu\nu} \) and \( \beta_n \) parameters consistent with a flat background, and shown to be ghost free in a certain decoupling limit. It was generalized to an arbitrary reference metric and arbitrary \( \beta_n \) in [10], where it was also reformulated in the form (4.14), and was shown to be ghost free away from
the decoupling limit in [11–13]. Specifically, the theory was analyzed in the ADM formalism and it was shows that there are two constraints which remove the Boulware-Deser ghost.
Chapter 5

Bimetric theory

In the ghost free massive gravity theory described by (4.14), $f_{\mu\nu}$ is a fixed background metric, which is imposed on the theory by hand. Promoting $f_{\mu\nu}$ to a dynamical variable would render the theory background independent, i.e. all fields would be determined by the theory itself. It would also restore general covariance. Hassan and Rosen (HR) promoted $f_{\mu\nu}$ to a dynamical variable in [14], resulting in the HR bimetric theory. This is a nonlinear theory of two interacting spin-2 fields, one massless and one massive. The action of this theory is given by

$$S = \int d^4x \left( M_g^2 \sqrt{-g} R(g) + M_f^2 \sqrt{-f} R(f) ight)$$

$$- 2m^4 \sqrt{-g} \sum_{n=0}^{4} \beta_n e_n(S) \right) + S_{M}^{(g)} + S_{M}^{(f)}, \quad (5.1)$$

where $R(g)$ and $R(f)$ are the Ricci scalars of $g_{\mu\nu}$ and $f_{\mu\nu}$, respectively, while $M_g$ and $M_f$ are the Planck masses for the $g$- and $f$-sectors. Comparing this to the massive gravity action (4.14), note that there is now an Einstein-Hilbert term for each metric. $S_{M}^{(g)}$ and $S_{M}^{(f)}$ are actions for matter fields that are minimally coupled to $g_{\mu\nu}$ and $f_{\mu\nu}$, respectively. In order for the theory to be ghost free, each matter field can only couple to one metric [14,30–32]. For most of this thesis, we will consider this theory without matter couplings.

As mentioned in section 4.2, $S$ is defined as the square root of the matrix $g^{-1}f$, meaning that $S^\mu_\lambda S^\lambda_\nu = g^{\mu\lambda} f_{\lambda\nu}$. This is not enough to specify $S$ uniquely, so in order for the action to be well defined, a specific square root must be chosen. This can be done by demanding that the action
is invariant under general coordinate transformations. This requires $S$ to transform as a $(1,1)$ tensor, a property held only by the principal square root. Hence, $S$ is defined as the principal square root of $g^{-1}f$, which determines it uniquely [15]. From the form of the elementary symmetric polynomials (4.15), it follows that

$$\sqrt{-g} e_n(S) = \sqrt{-f} e_{4-n}(S^{-1}), \quad (5.2)$$

implying that the form of the action (5.1) is preserved if $g_{\mu\nu}$ and $f_{\mu\nu}$ are interchanged [14]. In the absence of matter couplings, the action is completely symmetric under the interchange

$$g_{\mu\nu} \leftrightarrow f_{\mu\nu}, \quad M_g \leftrightarrow M_f, \quad \beta_n \rightarrow \beta_{4-n}. \quad (5.3)$$

By varying the action with respect to $g_{\mu\nu}$ and $f_{\mu\nu}$, we obtain the bimetric field equations [14],

$$G_{\mu\nu}(g) + \frac{m^4}{M_g^2} V_{\mu\nu}^{(g)} = \frac{1}{2M_g^2} T_{\mu\nu}^{(g)}, \quad (5.4)$$

$$G_{\mu\nu}(f) + \frac{m^4}{M_f^2} V_{\mu\nu}^{(f)} = \frac{1}{2M_f^2} T_{\mu\nu}^{(f)}. \quad (5.5)$$

Here, $G_{\mu\nu}(g)$ and $G_{\mu\nu}(f)$ are the Einstein tensors of $g_{\mu\nu}$ and $f_{\mu\nu}$, while $T_{\mu\nu}^{(g)}$ and $T_{\mu\nu}^{(f)}$ are stress-energy tensors of the matter fields, defined by

$$T_{\mu\nu}^{(g)} = -\frac{2}{\sqrt{-g}} \frac{\delta S_{M}^{(g)}}{\delta g^{\mu\nu}}, \quad T_{\mu\nu}^{(f)} = -\frac{2}{\sqrt{-f}} \frac{\delta S_{M}^{(f)}}{\delta f^{\mu\nu}}. \quad (5.6)$$

The tensors $V_{\mu\nu}^{(g)}$ and $V_{\mu\nu}^{(f)}$, sometimes referred to as the bimetric stress-energy tensors, are given by the variations of the bimetric interaction terms in the action,

$$V_{\mu\nu}^{(g)} = -\frac{2}{\sqrt{-g}} \frac{\partial}{\partial g^{\mu\nu}} \left[ \sqrt{-g} \sum_{n=0}^{4} \beta_n e_n(S) \right], \quad (5.7)$$

$$V_{\mu\nu}^{(f)} = -\frac{2}{\sqrt{-f}} \frac{\partial}{\partial f^{\mu\nu}} \left[ \sqrt{-f} \sum_{n=0}^{4} \beta_{4-n} e_n(S^{-1}) \right]. \quad (5.8)$$

Note that, because of the identity (5.2), the expressions inside the square brackets of (5.7) and (5.8) are in fact the same. Using the usual Bianchi
identities for the Einstein tensors, as well as local energy-momentum conservation, it follows that the bimetric stress-energy tensors satisfy the equations
\[ \nabla_\mu V^{(g)}_{\mu\nu} = 0, \quad \nabla_\mu V^{(f)}_{\mu\nu} = 0, \] (5.9)
which are known as the Bianchi constraints\(^1\). Here \( \nabla_\mu \) is the covariant derivative compatible with \( g_{\mu\nu} \) and \( \hat{\nabla}_\mu \) with \( f_{\mu\nu} \).

In order for the equations of motion to be real, the principal square root matrix \( S \) must be real, which is not true for generic metrics \( g_{\mu\nu} \) and \( f_{\mu\nu} \). The metrics must therefore be restricted in such a way that \( S \) is always real. This restriction is equivalent to demanding that the metrics \( g_{\mu\nu} \) and \( f_{\mu\nu} \) are causally coupled, i.e. that they share both a common timelike direction and a common spacelike hypersurface [15]. As long as initial conditions are chosen such that \( S \) is real, the equations of motion ensures that it will remain real [15].

5.1 Proportional background solutions

As a theory of one massless and one massive spin-2 field, bimetric theory is expected to propagate 7 degrees of freedom. However, if the theory is expanded around general backgrounds it is not possible to diagonalize the fluctuations into a massless and a massive mode. The most general backgrounds around which such a diagonalization is possible are the proportional backgrounds described in [30], which we will review in this section.

As a starting point, consider the conformal ansatz for the two metrics,
\[ \bar{f}_{\mu\nu} = c^2(x)\bar{g}_{\mu\nu}, \] (5.10)
where \( c(x) \) is dependent on spacetime. If this ansatz is inserted into the Bianchi constraints (5.9), they reduce to the equation \( \partial_\mu c(x) = 0 \). This restricts the metrics to being proportional, i.e.
\[ \bar{f}_{\mu\nu} = c^2\bar{g}_{\mu\nu}, \] (5.11)
where \( c \) is now a constant. Imposing this on the bimetric field equations (5.4) and (5.5) reduces them to two copies of the Einstein field equations
\[ G_{\mu\nu}(\bar{g}) + \Lambda g_{\mu\nu} = \frac{1}{2\bar{M}_g^2} \bar{T}^{(g)}_{\mu\nu}, \quad G_{\mu\nu}(\bar{g}) + \Lambda f_{\mu\nu} = \frac{1}{2\bar{M}_f^2} \bar{T}^{(f)}_{\mu\nu}, \] (5.12)
\(^1\)Sometimes (5.9) are referred to as the bimetric conservation law.
where $G_{\mu\nu}(\bar{g})$ is the Einstein tensor of $\bar{g}_{\mu\nu}$. Note that $G_{\mu\nu}(\bar{g})$ appears in both equations, since the Einstein tensor is invariant under a constant rescaling of the metric. The cosmological constants come from the bimetric stress-energy tensors and are given by

$$\Lambda_g = \frac{m^4}{M_g^2} \left( \beta_0 + 3c\beta_1 + 3c^2\beta_2 + c^3\beta_3 \right), \quad (5.13)$$

$$\Lambda_f = \frac{m^4}{M_f^2 c^2} \left( c\beta_1 + 3c^2\beta_2 + 3c^3\beta_3 + c^4\beta_4 \right). \quad (5.14)$$

In order for the equations in (5.12) to be consistent, the cosmological constants and stress-energy tensors must satisfy

$$(\Lambda_g - \Lambda_f) \bar{g}_{\mu\nu} = \frac{1}{2} \left( M_g^{-2} \bar{T}^{(g)}_{\mu\nu} - M_f^{-2} \bar{T}^{(f)}_{\mu\nu} \right). \quad (5.15)$$

Any vacuum energy contributions to $\bar{T}^{(g)}_{\mu\nu}$ and $\bar{T}^{(f)}_{\mu\nu}$ can be absorbed into the cosmological constants by rescaling $\beta_0$ and $\beta_4$. After doing this, the two sides of equation (5.15) must vanish separately

$$\bar{T}^{(f)}_{\mu\nu} = \alpha^2 \bar{T}^{(g)}_{\mu\nu}, \quad \alpha \equiv \frac{M_f}{M_g}, \quad (5.16)$$

$$\Lambda_g = \Lambda_f. \quad (5.17)$$

Equation (5.17) is a polynomial equation in $c$ where the coefficients depend on $\beta_n$ and $\alpha$. In general, this equation can be solved for $c$, which further restricts the ansatz. The case where (5.17) cannot be solved for $c$ is the candidate partially massless bimetric theory, which is described in section 7.2.

We now consider perturbations around the proportional background solutions

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu}, \quad f_{\mu\nu} = c^2 \bar{g}_{\mu\nu} + \delta f_{\mu\nu}. \quad (5.18)$$

For simplicity, we here restrict ourselves to the case where the matter stress-energy tensors vanish, which means that we are expanding around an Einstein spacetime. The fluctuations $\delta g_{\mu\nu}$ and $\delta f_{\mu\nu}$ are neither purely massless nor massive. However, as mentioned previously, around proportional backgrounds the fluctuations can be diagonalized into a massless and a massive mode. This is done by defining the following linear combinations of the fluctuations

$$h_{\mu\nu} \equiv \frac{1}{1 + \alpha^2 c^2} \left( \delta g_{\mu\nu} + \alpha^2 \delta f_{\mu\nu} \right), \quad (5.19)$$

$$\phi_{\mu\nu} \equiv \frac{1}{1 + \alpha^2 c^2} \left( \delta f_{\mu\nu} - c^2 \delta g_{\mu\nu} \right), \quad (5.20)$$
in terms of which the original fluctuations can be expressed as

\[ \delta g_{\mu\nu} = h_{\mu\nu} - \alpha^2 \varphi_{\mu\nu}, \quad \delta f_{\mu\nu} = c^2 h_{\mu\nu} + \varphi_{\mu\nu}. \]  

(5.21)

In terms of these linear combinations, the linearized bimetric field equations can be written

\[ \bar{\mathcal{E}}_{\mu\nu}^{\rho\sigma} h_{\rho\sigma} + \Lambda_g \left( h_{\mu\nu} - \frac{1}{2} \bar{g}_{\mu\nu} h \right) = 0, \]  

(5.22)

\[ \bar{\mathcal{E}}_{\mu\nu}^{\rho\sigma} \varphi_{\rho\sigma} + \Lambda_g \left( \varphi_{\mu\nu} - \frac{1}{2} \bar{g}_{\mu\nu} \varphi \right) - \frac{m_{\text{FP}}^2}{2} (\varphi_{\mu\nu} - \bar{g}_{\mu\nu} \varphi) = 0. \]  

(5.23)

The traces of the fluctuations are here defined with respect to \( \bar{g}_{\mu\nu} \), i.e. \( h = \bar{g}^{\mu\nu} h_{\mu\nu} \) and \( \varphi = \bar{g}^{\mu\nu} \varphi_{\mu\nu} \). As can be seen from (5.22) and (5.23), \( h_{\mu\nu} \) and \( \varphi_{\mu\nu} \) are, respectively, the massless and massive mode of the spin-2 fluctuations. The Fierz-Pauli mass of the massive mode is given by

\[ m_{\text{FP}}^2 = \frac{m^4 (1 + \alpha^2 c^2)}{M_g^2 \alpha^2 c^2 \left( c \beta_1 + 2c^2 \beta_2 + c^3 \beta_3 \right)}. \]  

(5.24)

In terms of the massless and massive mode, the lowest order terms in the bimetric action can be written [33]

\[ S_2 = \frac{M_g^2}{2} (1 + \alpha^2 c^2) \int d^4x \sqrt{-\bar{g}} \left[ h_{\mu\nu} \bar{\mathcal{E}}^{\mu\nu\rho\sigma} h_{\rho\sigma} + \Lambda_g \left( h_{\mu\nu} h_{\mu\nu} - \frac{1}{2} h^2 \right) \right. \]

\[ \left. + \frac{\alpha^2}{c^2} \left\{ \varphi_{\mu\nu} \bar{\mathcal{E}}^{\mu\nu\rho\sigma} \varphi_{\rho\sigma} + \Lambda_g \left( \varphi_{\mu\nu} \varphi_{\mu\nu} - \frac{1}{2} \varphi^2 \right) \right\} \right] - \frac{m_{\text{FP}}^2}{2} (\varphi_{\mu\nu} \varphi_{\mu\nu} - \varphi^2). \]  

(5.25)

Note that the terms involving the massive mode have the structure of the Fierz-Pauli action in a curved background, given in (4.10). From this it can be seen that bimetric theory propagates 7 degrees of freedom at the linear level, 2 for the massless mode and 5 for the massive. Showing that the number of d.o.f is the same at the nonlinear level will require use of the 3 + 1 formalism, which is reviewed in the next section.

5.2 Bimetric theory in the 3+1 formalism

Similarly to GR, the dynamic content of bimetric theory is conveniently analyzed using 3 + 1 variables. In this formalism, both \( g_{\mu\nu} \) and \( f_{\mu\nu} \) are
decomposed into lapse, shift and spatial metric. The decomposition of $g_{\mu\nu}$ is given in equations (3.12) and (3.13), while the decomposition of $f_{\mu\nu}$ is given by

$$f_{\mu\nu} = \left( -L^2 + L_k L^k L_j \phi_{ij} \right), \quad (5.26)$$

and its inverse as

$$f^{\mu\nu} = \frac{1}{L^2} \left( -\frac{1}{L^i} L^j \phi_{ij} - L^i L_j \right), \quad (5.27)$$

where $\phi^{ij}$ is the inverse of $\phi_{ij}$ and the index of $L^i$ is raised and lowered with $\phi_{ij}$. Note that in order for this simultaneous 3 + 1 decomposition to be possible, $g_{\mu\nu}$ and $f_{\mu\nu}$ must share a common spacelike hypersurface. As noted above, this follows from the requirement that $S$ is real [15].

Since the kinetic terms in (5.1) are of the Einstein-Hilbert form, the action contains no time derivatives of the lapses and shifts, meaning that their canonical momenta are zero. Denoting the canonical momenta of $\gamma_{ij}$ and $\phi_{ij}$ by $\pi_{ij}$ and $p_{ij}$, the Lagrangian density of (5.1) (without matter couplings) takes the form

$$\mathcal{L} = \left( \pi^{ij} \dot{\gamma}_{ij} + N \mathcal{R}^{(g)} + N^i \mathcal{R}^{(g)}_i \right) + \left( p^{ij} \dot{\phi}_{ij} + L \mathcal{R}^{(f)} + L^i \mathcal{R}^{(f)}_i \right)
- 2m^4 N \sqrt{\det \gamma} \sum_{n=0}^{4} \beta_n \epsilon_n(S), \quad (5.28)$$

where $\mathcal{R}^{(g)}$ and $\mathcal{R}^{(g)}_i$ are defined as in (3.17) and (3.18), while $\mathcal{R}^{(f)}$ and $\mathcal{R}^{(f)}_i$ are analogously defined as

$$\mathcal{R}^{(f)} = M_f^2 \sqrt{\phi} R(\phi) + \frac{1}{M_f^2 \sqrt{\phi}} \left( \frac{1}{2} p^2 - p^{ij} p_{ij} \right), \quad (5.29)$$

$$\mathcal{R}^{(f)}_i = 2 \sqrt{\phi} \nabla^j p_{ij}, \quad (5.30)$$

where $\nabla_i$ is the covariant derivative compatible with $\phi_{ij}$. While the lapses and shifts appear linearly in the Einstein-Hilbert terms of the bimetric action, as can be seen from (5.28), they appear nonlinearly in the interaction term. Naively, one might therefore expect their equations of motion to be eight algebraic equations which could be solved for the eight components of the lapses and shifts. However, as was mentioned in section 3.1, any covariant field theory contains four first class constraints. After imposing other nondynamical equations of motion, the Hamiltonian of such a theory
is of the form (3.24). It should therefore be possible to rewrite a generic
covariant bimetric action in this form, where four combinations of the lapses
and shifts appear as Lagrange multipliers. Four constraints are however not
enough to eliminate the Boulware-Deser ghost and reduce the theory to the
correct number of d.o.f.

For the particular form of the interaction term in (5.28), it is possible to
make a field redefinition so that five combinations of the lapses and shifts
appear as Lagrange multipliers. In order to do this, it is useful to note
that the interaction term of (5.28) only depends on the shifts through the
combination $N^i - L^i$ [12]. Hence, introducing the new shift like variables
$n^i$, defined by [9,11,12],

$$N^i - L^i = Ln^i + ND^i_j n^j,$$  \hspace{1cm} (5.31)

and expressing one shift in terms of them ensures linearity of the action in
the other shift. The specific form of (5.31), with the matrix $D$ defined as
the solution to the equation

$$\sqrt{x}D = \sqrt{(\gamma^{-1} - Dnn^T D^T)} \phi,$$  \hspace{1cm} (5.32)

has been chosen to ensure that the action is also linear in both lapses
[11,12,14]. Here, $x$ is defined as

$$x = 1 - n^i \phi_{ij} n^j.$$  \hspace{1cm} (5.33)

Expressing $N^i$ in terms of $n^i$ using (5.31), the Lagrangian (5.28) takes the
form [14],

$$\mathcal{L} = \pi^{ij} \dot{\gamma}_{ij} + p^{ij} \dot{\phi}_{ij} + NC + L\tilde{C} + L^i \mathcal{R}_i,$$  \hspace{1cm} (5.34)

with $\mathcal{C}$, $\tilde{\mathcal{C}}$ and $\mathcal{R}_i$ given by

$$\mathcal{C} = \mathcal{R}^{(g)} + \mathcal{R}_i^{(g)} D^i_j n^j - 2m^4 \sqrt{\gamma} V,$$  \hspace{1cm} (5.35)

$$\tilde{\mathcal{C}} = \mathcal{R}^{(f)} + n^i \mathcal{R}_i^{(g)} - 2m^4 \sqrt{\gamma} U,$$  \hspace{1cm} (5.36)

$$\mathcal{R}_i = \mathcal{R}_i^{(g)} + \mathcal{R}_i^{(f)}.$$  \hspace{1cm} (5.37)

The functions $U$ and $V$ are defined by

$$U = \beta_1 \sqrt{x} + \beta_2 \left[ xe_1(D) + n^i \phi_{ij} D^j_k n^k \right]$$

$$+ \beta_3 \left[ \sqrt{x} \left( e_1(D)n^i \phi_{ij} D^j_k n^k - D^i_k n^k \phi_{ij} D^j_l n^l \right) + x^{3/2} e_2(D) \right]$$

$$+ \beta_4 \frac{\sqrt{\phi}}{\sqrt{\gamma}},$$  \hspace{1cm} (5.38)

$$V = \beta_0 + \beta_1 \sqrt{x} e_1(D) + \beta_2 xe_2(D) + \beta_3 x^{3/2} e_3(D),$$  \hspace{1cm} (5.39)
where the elementary symmetric polynomials are again defined by (4.15).
Treating \( n^i \) as independent variables, \( N, L \) and \( L^i \) appear as Lagrange multipliers in (5.34), hence their equations of motion are

\[
\mathcal{C} = 0, \quad \tilde{\mathcal{C}} = 0, \quad \mathcal{R}_i = 0. \tag{5.40}
\]

In addition, using (5.32) it can be shown that [12],

\[
\frac{\partial \mathcal{C}}{\partial n^k} = \mathcal{C}_i \frac{\partial (D^i_j n^j)}{\partial n^k}, \quad \frac{\partial \tilde{\mathcal{C}}}{\partial n^k} = \mathcal{C}_k, \tag{5.41}
\]

where

\[
\mathcal{C}_i = \mathcal{R}_i^{(g)} + 2m^4 \sqrt{\gamma} \frac{n^l \phi_{lj}}{\sqrt{x}} \left[ \beta_1 \delta^j_i + \beta_2 \sqrt{x} \left( \delta^j_i D_k^k - D^j_i \right) \ight.
\]

\[
\left. + \beta_3 x \left( \delta^j_i e_2(D) + D^j_k D^k_i - D^j_i D^k_k \right) \right). \tag{5.42}
\]

From this it follows that varying the Lagrangian with respect to \( n^i \) yields the equations of motion

\[
\mathcal{C}_i \left[ L \delta^i_k + N \frac{\partial (D^i_j n^j)}{\partial n^k} \right] = 0. \tag{5.43}
\]

Since the expression inside the square brackets is an invertible matrix, it follows that the \( n^i \) equation of motion reduces to

\[
\mathcal{C}_i = 0, \tag{5.44}
\]

which are three algebraic equations for \( n^i \) that are independent of \( N, L \) and \( L^i \). Solving these gives \( n^i \) as a function of \( \gamma_{ij}, \phi_{ij} \) and \( \pi_{ij} \). Exact solutions to this equation are only known for the case where \( \beta_2 = \beta_3 = 0 \) and for spherically symmetric configurations with general \( \beta_n \), but it can be solved perturbatively for the general case [12,34]. These solutions are then used to eliminate \( n^i \) in the expressions for \( \mathcal{C} \) and \( \tilde{\mathcal{C}} \), ensuring that these only depend on the spatial metrics and their canonical momenta. It follows that, after imposing (5.44), (5.40) are five constraints on the dynamical variables [14].

### 5.3 The extra secondary constraint

The dynamical variables of bimetric theory are the spatial metrics, \( \gamma_{ij} \) and \( \phi_{ij} \), and their canonical momenta. Together, these have a total of 24 independent components. In order for the theory to describe a massless and
a massive spin-2 field, it should contain 7 dynamical d.o.f, i.e. 14 phase space variables, which means that 10 components of the spatial metrics and their momenta must be eliminated. The five constraints in (5.40) can be solved for five of these components, and general coordinate transformations can be used to eliminate another four. Unless there is some additional constraint, we are therefore left with an odd dimensional phase space, which is a highly unphysical situation.

As was described in chapter 2, the dynamical equations of motion may contain additional constraints, which become manifest by computing the time derivatives of the known constraints. This situation occurs in dRGT massive gravity, where the action is given by (4.14). Here, there is no Einstein-Hilbert term for \( f_{\mu\nu} \), since it's a nondynamical metric. Performing the field redefinition (5.31), the massive gravity Hamiltonian can be expressed as

\[
H = \int d^3 x \left( H_0 - NC \right),
\]  

(5.45)

where \( C \) is defined by (5.35) and \( H_0 \) is independent of \( N \). The \( N \) equation of motion is therefore the constraint \( C = 0 \).

The time derivative of \( C \) is given by

\[
\dot{C}(x) = \{ C(x), H \} + \frac{\partial C}{\partial f_{\mu\nu}} \dot{f}_{\mu\nu} 
= \int d^3 y \left( \{ C(x), H_0(y) \} - N(y) \{ C(x), C(y) \} \right) + \frac{\partial C}{\partial f_{\mu\nu}} \dot{f}_{\mu\nu}. 
\]  

(5.46)

The second term is needed to take the time dependence of the nondynamical metric \( f_{\mu\nu} \) into account. The requirement that \( \dot{C} \approx 0 \) could in principle lead to an extra secondary constraint or to an equation determining the lapse \( N \). It turns out that the second Poison bracket in (5.46) vanishes on the constraint surface, while the first does not [13]. The equation \( \dot{C} \approx 0 \) is therefore independent of \( N \) and provides an additional constraint, which ensures that the phase space of the theory is even dimensional. Together, these two constraints are responsible for eliminating the Boulware-Deser ghost [13].

In bimetric theory, the bracket \( \{ C(x), C(y) \} \) will also vanish weakly, since \( C \) does not depend on \( p^{ij} \). This implies that \( \dot{C} \) will be independent of \( N \), just like in massive gravity. In [13], it was argued that \( \dot{C} \approx 0 \) will therefore give rise to an extra secondary constraint in bimetric theory as well. However, \( \dot{C} \) was not computed for bimetric theory in [13]. Furthermore, it was argued in [16, 17] that \( \dot{C} \) will depend on both \( L \) and \( \partial_i L \) through the Poisson bracket \( \{ C(x), \dot{C}(y) \} \). If this was the case, \( \dot{C} \approx 0 \) would be an equation that
could be solved for $L$, rather than a constraint on the dynamical variables. One of the main aims of Paper I is to compute the time derivatives of the constraints in (5.40) explicitly and investigate if an additional secondary constraint really exists.

5.4 The geometrical metric in bimetric theory

As was explained in section 3.2, any covariant field theory contains four first class constraints whose Poisson brackets satisfy (3.25). It was noted in [18] that this can be used to identify a metric, the geometrical metric, even in theories where no metric is specified a priori. In addition, it was claimed in [35] that this should be the ”physical” metric, i.e. the metric to which matter couples, otherwise the theory is inconsistent.

In bimetric theory, the tensors $g_{\mu\nu}$ and $f_{\mu\nu}$ appear on an equal footing in the action (5.1). It is therefore not obvious what the geometrical metric of bimetric theory is, whether it’s unique and how it relates to the metrics $g_{\mu\nu}$ and $f_{\mu\nu}$. In addition, since different matter fields can couple minimally to either $g_{\mu\nu}$ or $f_{\mu\nu}$, there seems to be some tension with the idea that a unique ”physical” metric can be identified from the algebra. In [36] bimetric theory is studied in Plebanski’s [37] chiral formalism and the spatial part of the geometrical metric is identified. This metric is a very complicated function of the variables used and does not coincide with $\gamma^{ij}$ or $\phi^{ij}$. This result is difficult to interpret, partially due to the variables used.

The second aim of Paper I is to investigate this issue in the metric formulation. In order to do this, it is necessary to identify the first class constraints that generate diffeomorphisms in bimetric theory and then extract the geometrical metric from their Poisson algebra.
Chapter 6

Analysis of constraints

This chapter summarizes the results of Paper I and Paper II. Section 6.1 deals with the computation of the final secondary constraint and the absence of the Boulware-Deser ghost. In addition, the first class constraints of the theory are identified and their Poisson algebra is used to identify different geometrical metrics. Section 6.2 presents the computation of the ratio of lapses for spherically symmetric metrics.

6.1 Results of Paper I

6.1.1 The extra secondary constraint

As outlined in section 5.3, in addition to the constraints of (5.40), an extra secondary constraint is needed to ensure that the HR bimetric theory is free of the Boulware-Deser ghost and propagates the appropriate number of degrees of freedom. In dRGT massive gravity, such a constraint arises from the equation $\dot{C} \approx 0$, and it was argued in [13] that this is the case in bimetric theory as well. However, it was claimed in [16,17] that in bimetric theory, $\dot{C}$ depends on both $L$ and $\partial_i L$, which would imply that $\dot{C} \approx 0$ is not a constraint. In this section, the time derivatives of all constraints are computed. It is shown that $\dot{C}$ only depends on $L$ through an overall factor, which means that $\dot{C} \approx 0$ indeed provides a constraint. It is also shown that this is the only extra constraint that arises.

In order to investigate the possibility of an additional constraint, the time derivatives of all constraints in (5.40) must be computed, which is done by computing their Poisson brackets with the bimetric Hamiltonian.
The Poisson bracket is defined as
\[ \{A, B\} = \int d^3z \left( \frac{\delta A}{\delta \gamma_{mn}(z)} \frac{\delta B}{\delta \pi_{mn}(z)} - \frac{\delta A}{\delta \pi_{mn}(z)} \frac{\delta B}{\delta \gamma_{mn}(z)} + \frac{\delta A}{\delta \phi_{mn}(z)} \frac{\delta B}{\delta p_{mn}(z)} - \frac{\delta A}{\delta p_{mn}(z)} \frac{\delta B}{\delta \phi_{mn}(z)} \right), \tag{6.1} \]
and from the Lagrangian in (5.34), it can be seen that the Hamiltonian is given by
\[ H = -\int d^3x \left( NC + L\tilde{C} + L^iR_i \right). \tag{6.2} \]
The expressions for the time derivatives of the constraints will therefore involve Poisson brackets of constraints with each other. Computing the brackets of the five constraints in (5.40) with each other yields the result
\[ \{R_i(x), R_j(y)\} = R_j(x) \frac{\partial}{\partial y^i} \delta^3(x, y) + R_i(y) \frac{\partial}{\partial y^j} \delta^3(x, y), \tag{6.3} \]
\[ \{\tilde{C}(x), R_i(y)\} = \tilde{C}(y) \frac{\partial}{\partial y^i} \delta^3(x, y), \tag{6.4} \]
\[ \{\tilde{C}(x), \tilde{C}(y)\} = \left[ \phi^{ij}(x)R_j(x) + \phi^{ij}(y)R_j(y) \right] \frac{\partial}{\partial y^i} \delta^3(x, y), \tag{6.5} \]
\[ \{C(x), R_i(y)\} = C(y) \frac{\partial}{\partial y^i} \delta^3(x, y), \tag{6.6} \]
\[ \{C(x), C(y)\} = \left[ C(x)D^i_j n^j(x) + C(y)D^i_j n^j(y) \right] \frac{\partial}{\partial y^i} \delta^3(x, y), \tag{6.7} \]
\[ \{\tilde{C}(x), C(y)\} = C_2(x) \delta^3(x, y). \tag{6.8} \]
The quantity \( C_2 \) that appears in (6.8) is defined as
\[ C_2 = \left( R_j^{(g)} D^i_k n^k + 2m^4 \sqrt{\gamma} \bar{V}^{ki} \right) \nabla_i n^i - 2m^4 \sqrt{\gamma} \bar{\gamma}_{ij} D^i_k n^k \nabla_m U^{mn} \]
\[ + \frac{m^4}{M_f^2 \sqrt{\phi}} \left( \phi_{mn} n^k - 2 \pi_{mn} \right) W^{mn} - \frac{m^4}{M_g^2} \left( \pi_{mn} n^k - 2 \gamma_{mn} \right) U^{mn} \]
\[ + \sqrt{\gamma} \left[ \nabla_i \left( \frac{R_\xi^{(g)}}{\sqrt{\gamma}} \right) + \nabla_i \left( \frac{R_j^{(g)}}{\sqrt{\gamma}} \right) D^i_k n^k \right] n^i. \tag{6.9} \]
where the quantities $U_{mn}$, $\bar{V}_{mm}$ and $\bar{W}_{mn}$ are given by

\[
U_{mn} = 2\sqrt{\gamma} \frac{\partial (\sqrt{\gamma} U)}{\partial \gamma_{mn}},
\]

\[
\bar{V}_{mn} = \gamma^{mi} \gamma^{jn} \left[ \beta_1 \frac{1}{\sqrt{x}} \phi_{ik} (D^{-1})^k_j + \beta_2 \left( \phi_{ik} (D^{-1})^k_j D^l_i - \phi_{ij} \right) + \beta_3 \sqrt{x} \left\{ \phi_{ik} D^k_j - \phi_{ij} D^k_k + \phi_{ik} (D^{-1})^k_j e_2(D) \right\} \right],
\]

\[
\bar{W}_{mn} = \left( (D^{-1})^k_j \gamma^{nj} - n^k D^n j^n j \right) \left[ \beta_1 \frac{1}{\sqrt{x}} \delta^m_k + \beta_2 \left( D^i i \delta^m_k - D^m_k \right) + \beta_3 \sqrt{x} \left( e_2(D) \delta^m_k - D^i i D^m_k + D^m_i D^l_k \right) \right].
\]

The bracket (6.7) is computed in [13], while the rest are computed in Paper I. Naively, the bracket in (6.8) also seems to contain terms proportional to derivatives of $\delta$-functions, but the calculations carried out in Paper I show that they all cancel. Note that the brackets (6.3) - (6.7) vanish upon imposing (5.40), while (6.8) does not. It follows that the time derivatives of the five constraints are

\[
\dot{C}(x) = \{C(x), H\} \approx L(x) C_2(x),
\]

\[
\dot{\tilde{C}}(x) = \{\tilde{C}(x), H\} \approx -N(x) C_2(x),
\]

\[
\dot{R}_i(x) = \{R_i(x), H\} \approx 0.
\]

Note that, contrary to the claim of [16,17], $\dot{C}$ does not depend on $\partial_i L$. This is due to the fact that the bracket in (6.8) does not contain derivatives of $\delta$-functions. From the above equations it can be seen that $R_i$ is automatically preserved in time. Furthermore, in order for $C$ and $\tilde{C}$ to be preserved, we must impose

\[
C_2 = 0,
\]

since the lapses are always nonzero. Since $C_2$ is independent of lapses and shifts, equation (6.16) provides us with an extra secondary constraint on the dynamical variables$^1$.

In order for the extra constraint to be preserved, $\dot{C}_2$ is required to vanish on the constraint surface defined by (5.40) and (6.16). In order to compute this time derivative we require the Poisson brackets of $C_2$ with

\footnotesize
\begin{footnotesize}
$^1$This constraint has later been rederived using a different method in [34].
\end{footnotesize}
the constraints that appear in the Hamiltonian. Starting with the bracket
\{C_2(x), R_i(y)\}, we can rewrite it as

\begin{equation}
\{C_2(x), R_i(y)\} = \int d^3z \{\{\tilde{C}(x), C(z)\}, R_i(y)\}
= \int d^3z \left(\{\tilde{C}(x), \{C(z), R_i(y)\}\} - \{C(z), \{\tilde{C}(x), R_i(y)\}\}\right),
\end{equation}

where we have used (6.8) as well as the Jacobi identity. Using (6.4), (6.6)
and (6.8) it then follows that

\begin{equation}
\{C_2(x), R_i(y)\} = C_2(y) \frac{\partial}{\partial y^i} \delta^3(x, y),
\end{equation}

which implies that this bracket vanishes on the constraint surface. Now
consider the Poisson brackets of $C_2$ with $\tilde{C}$ and $C$. Using the Jacobi identity,
it is possible to show that these brackets contain no derivatives of delta
functions, as was done in [36]. It is therefore possible to express these
brackets as

\begin{align}
\{\tilde{C}(x), C_2(y)\} &= \Omega_f(x) \delta^3(x, y), \\
\{C(x), C_2(y)\} &\approx \Omega_g(x) \delta^3(x, y),
\end{align}

where $\Omega_f$ and $\Omega_g$ are functions of the phase space variables\(^2\). The exact
form of these functions is not known for the most general case, and is likely
very complicated, but it’s not needed here. Using these brackets, as well
as the Hamiltonian in (6.2), it follows that

\begin{equation}
\dot{C}_2 \approx NL \Omega_g + L \Omega_f.
\end{equation}

Note that on the constraint surface, $\dot{C}_2$ is independent of $L^i$, since the
bracket (6.18) vanishes weakly. As noted previously, $\dot{C}_2$ must vanish on the
constraint surface, which implies that

\begin{equation}
N \approx WL, \quad W \equiv -\frac{\Omega_f}{\Omega_g}.
\end{equation}

In other words, the condition that the constraint (6.16) is preserved in time
provides us with a linear relation between the two lapses.

\(^2\)Detailed derivations of (6.18), (6.19) and (6.20) can be found in Paper II.
6.1.2 Degrees of freedom

It is now possible to count the number of degrees of freedom of the theory. To start with, we have a total of 24 dynamical phase space variables in the form of $\gamma_{ij}, \phi_{ij}, \pi^i_j$ and $p^i_j$, as well as the 8 nondynamical variables $N, L, n^i$ and $L^i$. The $n^i$ can be solved for using their own equations of motion (5.44), while the lapse $N$ is determined by (6.22). The six constraints (5.40) and (6.16) allows us to eliminate six of the dynamical variables, while an additional four can be removed using general coordinate transformations. The equations of motion for these then become algebraic equations that can be solved for the lapse $L$ and the shift $L^i$. Thus, we are left with 14 undetermined phase space variables, corresponding to 7 propagating modes. This is the correct number for a theory of a massive and a massless spin-2 field. In particular, the constraints eliminate the Boulware-Deser ghost, ensuring that the theory only propagates healthy modes.

6.1.3 First class constraints and the geometrical metric

Recall from section 3.2 that in any covariant field theory there are four first class constraints whose Poisson brackets with each other satisfy the algebra (3.25), and that a metric (the so called geometrical metric) can be identified from this algebra. In particular, this should be true for bimetric theory. However, as mentioned in section 5.4, it is not obvious what the geometrical metric of bimetric theory is and how it relates to the metric which matter couples to. In this section, the four first class constraints of bimetric theory are identified and the corresponding geometrical metric is extracted from their algebra. However, there is no unique way of writing down the first class constraints, which means that the geometrical metric is not unique.

From inspection of the Poisson brackets (6.3), (6.4), (6.6) and (6.18), it follows that $R_i$ are first class constraints, since all of these brackets vanish weakly. Furthermore, (6.3) is of the same form as the last line of (3.25), which shows that $R_i$ are the generators of diffeomorphisms on the spacelike hypersurface. There should be one additional first class constraint, serving as the generator of diffeomorphisms orthogonal to the hypersurface.

However, the brackets (6.19) and (6.20) don’t vanish on the constraint surface. This is shown in appendix A of Paper I, where (6.22) is written out explicitly at the linear level. From there it follows that $\Omega_f$ and $\Omega_g$ are nonzero even after the constraints are imposed, implying that $C, \tilde{C}$ and $C_2$ are all second class. Despite this, a linear combination of them could still be first class [22]. This turns out to be the case, and this combination can
be found in the following way. Recall from above that the Hamiltonian density of a covariant theory should be of the form (3.26) after second class constraints have been imposed. Note that if we impose (6.22) on the bimetric Hamiltonian (6.2), the Hamiltonian density takes the form

$$H = -LR - L^iR_i,$$

$$R \equiv WC + \tilde{C}.$$ (6.23)

Since $R_i$ have already been identified as the generators of diffeomorphisms on the hypersurface, $R$ is a promising candidate for our final first class constraint. From (6.4) - (6.8) it follows that the Poisson brackets of $R$ with $\tilde{C}$, $C$ and $R_i$ vanish on the constraint surface. The bracket with $C_2$ can be evaluated using (6.19) and (6.20),

$$\{R(x), C_2(y)\} = \{W(x), C_2(y)\}C(x) + W(x)\{C(x), C_2(y)\}$$

$$\approx W(x)\Omega_g(x)\delta^3(x, y) + \Omega_f(x)\delta^3(x, y) = 0,$$ (6.24)

where, in the last step, the definition of $W$ in (6.22) has been used. This bracket also vanishes weakly, showing that $R$ is indeed first class. Apart from the four first class constraints there are two independent second class ones. From (6.4) - (6.8) it follows that

$$\{R(x), R_i(y)\} = R(y)\frac{\partial}{\partial y^i}\delta^3(x, y),$$ (6.25)

$$\{R(x), R(y)\} = \left[\phi^{ij}(x)R_j(x) + \phi^{ij}(y)R_j(y)\right] \frac{\partial}{\partial y^i}\delta^3(x, y),$$ (6.26)

which again is of the same form as (3.25), showing that $R$ is the generator of diffeomorphisms orthogonal to the hypersurface.

In (6.23), $L$ and $L^i$ appears as the Lagrange multipliers corresponding to lapse and shift, while $\phi^{ij}$ raises the index of $R_i$ in (6.26). This leads us to conclude that $f_{\mu\nu}$ is the geometrical metric. However, recall that $g_{\mu\nu}$ and $f_{\mu\nu}$ appears in the action on an equal footing, as is evident from the exchange symmetry in (5.3). The fact that $f_{\mu\nu}$ is singled out is a result of our choice of variables. Specifically, in the Lagrangian (5.34) we have eliminated $N^i$ using (5.31), ensuring that $L^i$ and the lapses appear as Lagrange multipliers. It is just as possible to use (5.31) to eliminate $L^i$ and express the Lagrangian as

$$\mathcal{L} = \pi^{ij}\dot{\gamma}_{ij} + p^{ij}\dot{\phi}_{ij} + NB + L\tilde{B} + N^i\mathcal{R}_i,$$ (6.27)

resulting in the constraints

$$\mathcal{B} = 0,$$

$$\tilde{B} = 0,$$

$$\mathcal{R}_i = 0.$$ (6.28)
Here, $B$ and $\tilde{B}$ are linear combinations of $C$, $\tilde{C}$ and $R_i$, so these constraints are equivalent to the ones in (5.40). It is then possible to repeat the analysis of section 6.1.1 in these variables. The requirement that $B$ and $\tilde{B}$ are preserved in time gives rise to an additional secondary constraint, the preservation of which results in an equation for the lapses. We use this to solve for $L$ in terms of $N$, and impose this solution on the Hamiltonian, which reduces to
\begin{equation}
\mathcal{H} = -N\tilde{\mathcal{R}} - N^i R_i, \tag{6.29}
\end{equation}
where $\tilde{\mathcal{R}}$ is a linear combination of $B$ and $\tilde{B}$. It turns out that $\tilde{\mathcal{R}}$ is first class\(^3\) and that its Poisson brackets with itself and $R_i$ are given by
\begin{equation}
\{\tilde{\mathcal{R}}(x), \mathcal{R}_i(y)\} = \tilde{\mathcal{R}}(y) \frac{\partial}{\partial y^i} \delta^3(x,y), \tag{6.30}
\end{equation}
\begin{equation}
\{\tilde{\mathcal{R}}(x), \tilde{\mathcal{R}}(y)\} = \left[\gamma^{ij}(x)R_j(x) + \gamma^{ij}(y)R_j(y)\right] \frac{\partial}{\partial y^i} \delta^3(x,y). \tag{6.31}
\end{equation}
These are of the form (3.25) and the index of $\mathcal{R}_i$ is raised by $\gamma^{ij}$ in (6.31). This, together with the appearance of $N$ and $N^i$ as Lagrange multipliers in (6.29) lets us identify $g_{\mu\nu}$ as the geometrical metric.

These results shows us that the HKT metric is not unique. This can be understood as follows: if a theory has four first class constraints $\tilde{\mathcal{R}}_i$ and $\tilde{\mathcal{R}}$ satisfying the Poisson algebra (3.25), then these constraints are the generators of diffeomorphisms on and orthogonal to the spacelike hypersurface, respectively. However, the notion of orthogonality is metric dependent. If the first class constraints can be expressed in such a way that they satisfy the diffeomorphism algebra, then the metric identified from that algebra is such that $\tilde{\mathcal{R}}$ generates diffeomorphisms orthogonal to the hypersurface with respect to that metric.

In the case of bimetric theory, the normal vector to the spatial hypersurface w.r.t $g_{\mu\nu}$ is in general not the same as the one w.r.t $f_{\mu\nu}$. The first class constraint $\mathcal{R}$ generates diffeomorphisms in the normal direction w.r.t $f_{\mu\nu}$, and $\tilde{\mathcal{R}}$ in the normal direction w.r.t $g_{\mu\nu}$. Together with the generators $\mathcal{R}_i$, either one can be used to parametrize the full set of spacetime diffeomorphisms. For either choice, the corresponding geometrical metric appears in the diffeomorphism algebra. The geometrical metric therefore has no direct physical meaning, it is merely related to how we choose to express the diffeomorphism generators.

\(^3\)The $\tilde{\mathcal{R}}$ is in fact a linear combination of $\mathcal{R}$ and $\mathcal{R}_i$ which could in principle have been found without expressing the Hamiltonian in terms of $N^i$.\(^4\)
As was noted in chapter 5, a matter field can couple minimally to either \( g_{\mu\nu} \) or \( f_{\mu\nu} \). The metric that couples to matter is the gravitational metric, i.e. test particles will follow its geodesics. The matter Lagrangian is linear in the lapse and shift and while it gives rise to extra terms in the constraint equations, it does not change the structure of the constraint algebra. Hence, the above results regarding the geometrical metric are still valid when the theory is coupled to matter. In particular, note that the gravitational metric and the geometrical metric need not coincide, contrary to the claim of [35]. It is, for instance, possible to couple matter to \( g_{\mu\nu} \) while expressing the Hamiltonian in terms of \( L \) and \( L^i \) (as above), resulting in \( f_{\mu\nu} \) as the geometrical metric.

### 6.2 Results of Paper II

With the results of Paper I, all of the bimetric equations of motion in the \( 3+1 \) formalism are known. They are the evolution equations for the dynamical variables \( \gamma_{ij} \), \( \phi_{ij} \), \( \pi^i \) and \( p^i \), the constraints on the dynamical variables, (5.40) and (6.16), and the constraints on the shifts (5.44) and lapses (6.22). These equations were originally derived in the ADM formalism, as has been reviewed above. With the exception of (6.22), they have all been rederived in [34] using a different method. In this approach, the starting point is the covariant bimetric field equations (5.4) and (5.5), as well as the Bianchi constraints (5.9). The bimetric field equations are projected onto the spacelike hypersurface, as well as onto the direction normal to it, which results in the evolution equations for the dynamical variables as well as the constraints (5.40) and (5.44). Projecting the Bianchi constraints on the the direction normal to the hypersurface yields the constraint equation (6.16). The two approaches are illustrated in figure 6.1.

Given the equations of motion in the \( 3+1 \) formalism it is in principle possible to solve initial value problems in bimetric theory. The fields would then be specified on some initial hypersurface, subject to the constraints, and evolved using the evolution equations. However, while it was shown in Paper I that the lapses are related through (6.22), where their ratio \( W \) is purely a function of the dynamical variables, the explicit form of \( W \) was not computed. Without knowing this, solving initial value problems is not possible. In principle one might be able to arrive at an explicit expression for \( W \) by computing the Poisson brackets (6.19) and (6.20). In practice, due to the complicated form of \( C_2 \), such a calculation would be extremely lengthy, particularly for the most general case. If instead we were to use the approach of [34], there is no obvious way to arrive at (6.22) by projecting
Figure 6.1: Comparison between the ADM analysis of Paper I (red) and the projection based approach of [34] (blue). Figure from Paper II.

The aim of Paper II is to calculate $W$ for the special case where the two metrics share the same spherical symmetry\(^4\). Restricting ourselves to this case makes the problem significantly more manageable, in part since it is one of the special cases where (5.44) can be solved explicitly for $n^i$. The most general expressions for metrics of this type are given by

$$g_{\mu \nu} dx^\mu dx^\nu = -N^2 dt^2 + A^2 (dr + N^r dt)^2 + B^2 (d\theta^2 + \sin^2 \theta d\phi), \quad \text{(6.32)}$$

$$f_{\mu \nu} dx^\mu dx^\nu = -L^2 dt^2 + \tilde{A}^2 (dr + L^r dt)^2 + \tilde{B}^2 (d\theta^2 + \sin^2 \theta d\phi), \quad \text{(6.33)}$$

in spherical polar coordinates. Here $N^r$ and $L^r$ are the radial components of the shift vectors (the angular components are zero) while $A, B, \tilde{A}$ and $\tilde{B}$ are the nontrivial components of the spatial vielbeins\(^5\). We make use of the equations of motion in the variables of [34], since these are somewhat easier to work with. Rather than the canonical momenta $\pi^{ij}$ and $p^{ij}$, the equations are then written in terms of extrinsic curvatures\(^6\), whose nonzero

\(^4\)If an isometry is shared by the two metrics, it is a symmetry of all tensor fields [38].

\(^5\)The spatial metrics can be expressed in terms of spatial vielbeins $e^a_i$ and $m^a_i$, through $\gamma_{ij} = \delta_{ab} e^a_i e^b_j$ and $\phi_{ij} = \delta_{ab} m^a_i m^b_j$, where $a, b \in \{1, 2, 3\}$.

\(^6\)The extrinsic curvatures are related to the canonical momenta through $\pi^{ij} = M_2^2 \sqrt{\gamma} \left( K^{ij} - \gamma^{ij} K \right)$ and $p^{ij} = M_1^2 \sqrt{\phi} \left( \tilde{K}^{ij} - \phi^{ij} \tilde{K} \right)$. 
components in the spherically symmetric case are denoted by

\[ K_1 \equiv K^{r}_r, \quad K_2 \equiv K^{\theta}_\theta = K^{\phi}_\phi, \quad (6.34) \]

\[ \tilde{K}_1 \equiv \tilde{K}^{r}_r, \quad \tilde{K}_2 \equiv \tilde{K}^{\theta}_\theta = \tilde{K}^{\phi}_\phi. \quad (6.35) \]

The equations are also expressed in terms of the mean shift vector, defined by

\[ q^i \equiv N^i - N D^i_j n^j = L^i + L n^i, \quad (6.36) \]

where the second equality follows from (5.31). After having imposed spherical symmetry, we compute the time derivative of \( C^2 \). For convenience we work with \( C^b \equiv C^2 / \sqrt{7} \), which is a scalar rather than a scalar density. Since \( C^b \) depends on \( n^i \), we require an evolution equation for \( n^i \). Thanks to the spherical symmetry, such an equation can be obtained by projecting the Bianchi constraints (5.9) on the hypersurface. The requirement that \( \partial_t C^b \approx 0 \) leads to the equation

\[ F_1 \partial_t \tilde{A} + F_2 \partial_t \tilde{B} + F_3 \partial_t \tilde{K}_1 + F_4 \partial_t \tilde{K}_2 + F_5 \partial_t A + F_6 \partial_t B + F_7 \partial_t K_1 + F_8 \partial_t K_2 + F_9 \partial_t \partial_r \tilde{B} + F_{10} \partial_t \partial_r B + F_{11} \approx 0, \quad (6.37) \]

where \( F_i \) are some field dependent coefficients which do not contain time derivatives. By imposing the evolution equations for the spatial metrics and the extrinsic curvatures this can be rewritten as

\[ G_0 + G_1 N + G_2 L + G_3 \partial_r N + G_4 \partial_r L + G_5 \partial_r^2 N + G_6 \partial_r^2 L \approx 0. \quad (6.38) \]

Here, there are no time derivatives and the coefficients \( G_i \) are independent of the lapses. It turns out that

\[ G_0 = \partial_r (q^r C^b) \approx 0. \quad (6.39) \]

\( G_3, G_4, G_5 \) and \( G_6 \) also vanish. This should not come as a surprise, since the analysis of the Poisson brackets in section 6.1.1 shows that the time derivative of \( C^b \) should not contain any derivatives of the lapses. Equation (6.38) then reduces to

\[ G_1 N + G_2 L \approx 0. \quad (6.40) \]

Comparing this with (6.21) and (6.22) we observe that the ratio of the lapses is given by

\[ W = -\frac{G_2}{G_1}. \quad (6.41) \]

Since \( G_1 \) and \( G_2 \) are calculated explicitly, we therefore arrive at an expression for \( W \) in the spherically symmetric case. The full expressions for \( G_1 \) and \( G_2 \) are given in Paper II, but due to their length they are omitted here.
Chapter 7

Partial masslessness

In this chapter the question of whether a nonlinear partially massless (PM) theory exists as a special case of the Hassan-Rosen bimetric theory is investigated. In section 7.1 and 7.2, the linear PM theory and the candidate nonlinear PM bimetric theory are presented. Section 7.3 summarizes the results of Paper III, which shows that the PM symmetry cannot be extended beyond cubic order in the action for a theory of two interacting spin-2 fields.

7.1 The linear PM theory

Recall from section 4.1 that the Fierz-Pauli linear theory of a massive spin-2 field propagating on an Einstein background is described by the action

\[ S_{FP} = \frac{M_p^2}{2} \int d^4x \sqrt{-g} \left[ \varphi_{\mu\nu} \bar{E}^{\mu\nu\rho\sigma} \varphi_{\rho\sigma} + \Lambda \left( \varphi^{\mu\nu} \varphi_{\mu\nu} - \frac{1}{2} \varphi^2 \right) \right. 
\]

\[ \left. - m_{FP}^2 \left( \varphi^{\mu\nu} \varphi_{\mu\nu} - \varphi^2 \right) \right]. \quad (7.1) \]

If the background is taken to be an Einstein-de Sitter space, this theory is qualitatively different from the FP theory in flat space due to the presence of the Higuchi bound at \( m_{FP}^2 \geq \frac{2}{3} \Lambda \). If the Fierz-Pauli mass lies below this bound the helicity zero mode of the massive spin-2 field becomes a ghost, whereas the theory propagates five healthy degrees of freedom for a mass above the bound [39, 40]. If the mass saturates the Higuchi bound, \( m_{FP}^2 = \frac{2}{3} \Lambda \), the theory possesses an additional gauge symmetry. Specifically, for this value of the mass the action (7.1) is invariant under the linear gauge
transformation
\[ \delta \varphi_{\mu\nu} = \left( \bar{\nabla}_\mu \bar{\nabla}_\nu + \frac{\Lambda}{3} \bar{g}_{\mu\nu} \right) \xi, \tag{7.2} \]

where \( \xi \) is a local gauge parameter \([19, 41–46]\). As a result, the theory has four degrees of freedom when the Higuchi bound is saturated. \( \varphi_{\mu\nu} \) is then referred to as a partially massless spin-2 field. Partially massless representations of fields with spin greater than or equal to two are permitted by the representation theory of the de Sitter group \( SO(1, 4) \) (for a discussion, see for instance \([47]\)).

Even at the linear level, the PM gauge symmetry is ruined by coupling \( \varphi_{\mu\nu} \) to matter, unless the matter fields are conformally invariant.

### 7.2 The candidate nonlinear PM bimetric theory

It was shown in \([48–50]\) that the linear partially massless theory can be extended to cubic order in \( \varphi_{\mu\nu} \) while leaving the gauge symmetry intact. This begs the question if it can be extended to even higher orders, and if it is possible to construct a fully nonlinear PM theory. If such a theory exists, it is natural to expect it to occur as a special case of the nonlinear massive gravity and bimetric theories described in earlier chapters.

An attempt to find a nonlinear PM theory within the class of dRGT massive gravity theories described by \((4.14)\) was made in \([51]\). The de Sitter metric was used as reference metric and it was shown that a scalar degree of freedom is not propagating in the decoupling limit, given a particular choice of the \( \beta_n \) parameters. This would seem to indicate the existence of a gauge symmetry in the decoupling limit. However, as was shown in \([52]\), the PM theory cannot be extended to quartic order in \( \varphi_{\mu\nu} \) if the reference metric is fixed. Similar no-go results were found in \([53–55]\) for one or several PM fields. This rules out nonlinear PM theories based on massive gravity. In contrast, their argument does not rule out nonlinear PM bimetric theory, where both metrics are dynamical. The candidate PM bimetric theory is described below.

In section 5.1 we linearized bimetric theory around proportional Einstein-de Sitter backgrounds. Around these backgrounds, the fluctuations can be diagonalized into a massless mode \( h_{\mu\nu} \) and a massive mode \( \varphi_{\mu\nu} \). In terms of these modes, the linearized bimetric action is given by \((5.25)\). Since the \( \varphi_{\mu\nu} \)-terms in this linearized action are of the Fierz-Pauli form, they will be invariant under a linear PM transformation

\[ \delta \varphi_{\mu\nu} = \left( \bar{\nabla}_\mu \bar{\nabla}_\nu + \frac{\Lambda}{3} \bar{g}_{\mu\nu} \right) \xi, \tag{7.3} \]
if the Fierz-Pauli mass saturates the Higuchi bound. The terms involving
the massless mode are on the other hand only invariant under linearized
diffeomorphisms. Hence, the PM transformation acting on the massless
fluctuation can at most give rise to a term that resembles a coordinate
transformation, $\delta h_{\mu \nu} \sim \bar{\nabla}_\mu \bar{\nabla}_\nu \xi$. Hence, up to a GCT, we can write $\delta h_{\mu \nu} = 0$. From this it follows that the transformations of $\delta g_{\mu \nu}$ and $\delta f_{\mu \nu}$ are given by

$$
\Delta(\delta g_{\mu \nu}) = a \left( \bar{\nabla}_\mu \bar{\nabla}_\nu + \frac{\Lambda g}{3} \bar{g}_{\mu \nu} \right) \xi, \quad (7.4)
$$

$$
\Delta(\delta f_{\mu \nu}) = b \left( \bar{\nabla}_\mu \bar{\nabla}_\nu + \frac{\Lambda g}{3} \bar{g}_{\mu \nu} \right) \xi, \quad (7.5)
$$

where the constants $a$ and $b$ can be written in terms of $\alpha$ and $c$. The linearized bimetric theory differs from the Fierz-Pauli theory since the background metric $\bar{g}_{\mu \nu}$ is dynamical, rather than imposed by hand. In [20], a unique candidate nonlinear PM bimetric theory was found by demanding compatibility of (7.4) and (7.5) with the dynamical nature of the background. This approach is reviewed below.

The first step is to note that, since $g_{\mu \nu}$ and $f_{\mu \nu}$ are dynamical, the split into background and fluctuation is not unique. It is always possible to redefine the backgrounds by transferring a part of the fluctuations to it, so that

$$
g_{\mu \nu} = \bar{g}_{\mu \nu} + \delta g_{\mu \nu} = \bar{g}'_{\mu \nu} + \delta g'_{\mu \nu}, \quad f_{\mu \nu} = \bar{f}_{\mu \nu} + \delta f_{\mu \nu} = \bar{f}'_{\mu \nu} + \delta f'_{\mu \nu}. \quad (7.6)
$$

From this it follows that if the theory has a nonlinear PM symmetry that reduces to the infinitesimal transformations (7.4) and (7.5) around proportional backgrounds, it should be possible to move these transformations from the fluctuations to the backgrounds and end up with some other consistent background solutions. For a generic local gauge parameter $\xi(x)$, the new background solutions $\bar{g}'_{\mu \nu}$ and $\bar{f}'_{\mu \nu}$ will not be proportional. As noted above, this means that it is not possible to diagonalize the fluctuations into a massless and massive mode. In order to avoid this complication, we restrict ourselves to constant gauge transformations, $\xi = \xi_0$. When these transformations are moved to the backgrounds, they give us the new backgrounds

$$
\bar{g}'_{\mu \nu} = \bar{g}_{\mu \nu} + \frac{a \Lambda g \xi_0}{3} \bar{g}_{\mu \nu}, \quad \bar{f}'_{\mu \nu} = \bar{f}_{\mu \nu} + \frac{b \Lambda g \xi_0}{3} \bar{g}_{\mu \nu}. \quad (7.7)
$$

Since $\bar{f}_{\mu \nu} = c^2 \bar{g}_{\mu \nu}$, the new backgrounds are also proportional, $\bar{f}'_{\mu \nu} = c' \bar{g}'_{\mu \nu}$, where $c' \neq c$. The difference between $c$ and $c'$ is infinitesimal, but since the
transformations (7.4) and (7.5) are integrable\(^1\), they can be integrated to generate a \(c'\) that differs from \(c\) by a finite amount.

Recall from section 5.1 that, for generic \(\beta_n\), \(c\) is determined by the background equations of motion, specifically by equation (5.17). This means that \(c'\) cannot be a solution to (5.17), since it differs from \(c\), implying that \(g'_{\mu\nu}\) and \(f'_{\mu\nu}\) are not valid background solutions. The transformations that generate them are therefore not symmetries of the nonlinear theory. Hence, a necessary condition for a nonlinear PM bimetric theory is that (5.17) leaves \(c\) undetermined. Using (5.13) and (5.14), equation (5.17) can be written

\[
\beta_1 + (3\beta_2 - \alpha^2\beta_0) c + 3 \left( \beta_3 - \alpha^2\beta_1 \right) c^2 \\
+ \left( \beta_4 - 3\alpha^2\beta_2 \right) c^3 - \alpha^2\beta_3 c^4 = 0,
\]

which only leaves \(c\) undetermined if the coefficient of every power of \(c\) vanishes, that is if

\[
\beta_0 = 3\alpha^{-2}\beta_2, \quad \beta_4 = 3\alpha^2\beta_2, \quad \beta_1 = \beta_3 = 0.
\]

These values are often referred to as the PM parameters. If they are inserted in (5.13) and (5.24), it follows that \(m_{FP}^2 = \frac{2}{5}\Lambda_g\), i.e. the mass of the massive fluctuation saturates the Higuchi bound and the linearized theory possesses the usual linear PM symmetry. Note that this was not imposed, but followed from demanding invariance of the background equations.

Since we have restricted the gauge parameter \(\xi\) to be constant, we are not dealing with the full gauge group. However, as we have seen, even demanding invariance of the background equations under the subgroup of constant transformations is enough to determine all \(\beta_n\) except one, which sets the scale of the Fierz-Pauli mass. If the theory is to be invariant under PM transformations with an arbitrary gauge parameter, it must in particular be invariant under transformations where the parameter is constant. Hence, (7.9) gives us a unique candidate nonlinear PM bimetric theory.

While [20] arrived at a unique candidate for a PM bimetric theory, there was no proof that the PM symmetry really did extend to nonlinear order. In addition, it was argued in [56] that there exists no nonlinear unitary theory of a PM spin-2 field coupled to gravity. In particular, it was claimed that this argument applied to the candidate theory presented above, a claim which was disputed in [57]. This left the existence of the nonlinear PM symmetry in the candidate theory an open question. Investigating this issue is the purpose of Paper III.

\(^1\)Their integrability was proven in [20].
7.3 Results of Paper III

In this section, the existence of the nonlinear PM symmetry in the candidate theory is investigated. The approach is to expand the action order by order in the massive field and for each order check if there exists an extension of the PM transformation which renders the action invariant. In principle this approach could never prove complete nonlinear invariance, but the no-go result of [52] for massive gravity, as well as the argument of [56], indicate that something might go wrong at the quartic order, providing motivation to investigate if this is the case. As we extend the PM transformation, our only condition is that it can contain at most second derivative terms.

7.3.1 Rewriting the bimetric action

We start by expanding the metrics $g_{\mu\nu}$ and $f_{\mu\nu}$ around proportional de Sitter backgrounds and diagonalizing the perturbations into a massless and a massive mode, as in (5.18), (5.19) and (5.20). It is now possible to define a new metric, $\tilde{g}_{\mu\nu}$, by adding the massless fluctuation $h_{\mu\nu}$ to the background metric

$$\tilde{g}_{\mu\nu} \equiv \bar{g}_{\mu\nu} + h_{\mu\nu}. \quad (7.10)$$

The metrics $g_{\mu\nu}$ and $f_{\mu\nu}$ can now be expressed in terms of this new metric and the massive fluctuation

$$g_{\mu\nu} = \tilde{g}_{\mu\nu} - \alpha^2 \varphi_{\mu\nu}, \quad f_{\mu\nu} = c^2 \tilde{g}_{\mu\nu} + \varphi_{\mu\nu}, \quad (7.11)$$

from which it follows that the bimetric action (5.1), with $\beta_n$ given by (7.9), can be written as

$$S = (1 + \alpha^2 c^2)M_g^2 \int d^4x \left[ \sqrt{-\tilde{g}} \left( R(\tilde{g}) - 2\Lambda_g \right) + \sum_{n=2}^{\infty} \mathcal{L}_n \right], \quad (7.12)$$

where the Lagrangian density $\mathcal{L}_n$ depends on $n$ powers of $\varphi_{\mu\nu}$. Here we have resummed the massless fluctuation $h_{\mu\nu}$ to all orders, and we end up with the action of a spin-2 field $\varphi_{\mu\nu}$ coupled to the metric $\tilde{g}_{\mu\nu}$ in a nonminimal way. The relative coefficients in (5.19) and (5.20), which were chosen to diagonalize the action at quadratic order around a de Sitter background, also ensures that no terms linear in $\varphi_{\mu\nu}$ appear in the action when it is expanded around the metric $\tilde{g}_{\mu\nu}$, hence the absence of an $\mathcal{L}_1$-term in (7.12).
The quadratic Lagrangian density, $L_2$, is given by

$$L_2 = \frac{\alpha^2}{2c^2} \sqrt{-\tilde{g}} \left[ \varphi_{\mu\nu} \tilde{\xi}^{\mu\nu\rho\sigma} \varphi_{\rho\sigma} + \frac{2\Lambda g}{3} \varphi_{\mu\nu} \varphi^{\mu\nu} - \frac{\Lambda g}{6} \varphi^2 
+ \left( \tilde{G}^{\mu\nu} + \Lambda g \tilde{g}^{\mu\nu} \right) \left( -\frac{1}{2} \tilde{g}_{\mu\nu} \varphi^{\rho\sigma} \varphi_{\rho\sigma} + \frac{1}{4} \tilde{g}_{\mu\nu} \varphi^2 
+ 2 \varphi^\rho_{\nu} \varphi_{\mu\rho} - \varphi \varphi_{\mu\nu} \right) \right].$$

(7.13)

Here $\tilde{G}_{\mu\nu}$ is the Einstein tensor of $\tilde{g}_{\mu\nu}$ and $\tilde{\xi}^{\mu\nu\rho\sigma}$ is given by

$$\tilde{\xi}^{\mu\nu\rho\sigma} = -\frac{1}{2} \left( g^{\rho\nu} \tilde{\nabla}^\sigma \tilde{\nabla}^\mu + g^{\rho\mu} \tilde{\nabla}^\sigma \tilde{\nabla}^\nu - g^{\rho\mu} \tilde{g}^{\sigma\nu} \tilde{\nabla}^2 - g^{\rho\sigma} \tilde{\nabla}^\mu \tilde{\nabla}^\nu 
- \tilde{g}^{\mu\nu} \tilde{\nabla}^\rho \tilde{\nabla}^\sigma + \tilde{g}^{\mu\nu} \tilde{g}^{\rho\sigma} \tilde{\nabla}^2 \right),$$

(7.14)

where $\tilde{\nabla}_\mu$ is the covariant derivative compatible with $\tilde{g}_{\mu\nu}$. If we expand $L_2$ around the de Sitter background $\bar{g}_{\mu\nu}$, the last two lines vanish and we recover the Lagrangian density of the Fiertz-Pauli action with the FP-mass saturating the Higuchi bound, consistent with (5.25). Our analysis also makes use of $L_3$ and $L_4$, whose explicit expressions are quite lengthy and hence relegated to appendix A.

### 7.3.2 Extending the gauge symmetry

In this section we attempt to find the PM gauge symmetry of the action (7.12) order by order in the massive field. We let $\delta^{(n)} \mathcal{O}$ denote the terms in the PM transformation of the field $\mathcal{O}$ depending on $n$ powers of $\varphi_{\mu\nu}$, so that schematically

$$\delta^{(n)} \mathcal{O} \sim \varphi^n \tilde{\nabla} \varphi \tilde{\nabla} \xi.$$

(7.15)

where $\xi$ is again a local gauge parameter. We also define

$$S^{(n)} = (1 + \alpha^2 c^2) M^2 \int d^4 x L_n,$$

(7.16)

and let $S^{(0)}$ be the Einstein-Hilbert action of $\tilde{g}_{\mu\nu}$. Our approach will be to, for each order in $\varphi_{\mu\nu}$, first find the transformations on a de Sitter background and then extend them to all orders in $h_{\mu\nu}$.

The zeroth order variation of the action is given by

$$\delta_{\xi}^{(0)} S = \int d^4 x \frac{\delta S^{(0)}}{\delta \tilde{g}_{\mu\nu}} \delta_{\xi}^{(0)} \tilde{g}_{\mu\nu}.$$

(7.17)
Since $S^{(0)}$ is the Einstein-Hilbert action, it is only invariant under diffeomorphisms. In order for $\delta_{\xi}^{(0)} S$ to vanish, the zeroth order PM transformation of $\tilde{g}_{\mu\nu}$ must therefore resemble a diffeomorphism, i.e.

$$\delta_{\xi}^{(0)} \tilde{g}_{\mu\nu} = \lambda_1 \tilde{\nabla}_\mu \tilde{\nabla}_\nu \xi,$$

where $\lambda_1$ is given by

$$\lambda_1 = -\left(\frac{\alpha^2 c^2 - 1}{2c^2}\right).$$

The normalization has been chosen for later convenience. This result was also derived in [57].

The first order variation of the action reads

$$\delta_{\xi}^{(1)} S = \int d^4 x \left( \frac{\delta S^{(2)}}{\delta \varphi_{\mu\nu}} \delta_{\xi}^{(0)} \varphi_{\mu\nu} + \frac{\delta S^{(0)}}{\delta \tilde{g}_{\mu\nu}} \delta_{\xi}^{(1)} \tilde{g}_{\mu\nu} \right).$$

When this is put on a de Sitter background the second term vanishes due to the background equations of motion. Demanding that the first order variation vanishes in this background yields

$$\delta_{\xi}^{(0)} \varphi_{\mu\nu}|_{dS} = \left( \tilde{\nabla}_\mu \tilde{\nabla}_\nu + \frac{\Lambda_g}{3} \tilde{g}_{\mu\nu} \right) \xi,$$

which means that we recover the standard linear PM transformation (7.2), as we should. This gauge invariance can be extended away from the de Sitter background by demanding that $\delta_{\xi}^{(1)} \tilde{g}_{\mu\nu}$ cancels the extra terms arising from the last two lines in (7.13). The resulting transformations are

$$\delta_{\xi}^{(0)} \varphi_{\mu\nu} = \left( \tilde{\nabla}_\mu \tilde{\nabla}_\nu + \frac{\Lambda_g}{3} \tilde{g}_{\mu\nu} \right) \xi,$$

$$\delta_{\xi}^{(1)} \tilde{g}_{\mu\nu} = -\frac{\alpha^2}{2c^2} \left( 2\tilde{\nabla}_{(\mu} \varphi_{\nu)} - \tilde{\nabla}_\rho \varphi_{\mu\nu} \right) \tilde{\nabla}^\rho \xi.$$

These transformations were also derived in [56]. The first order transformation of $\tilde{g}_{\mu\nu}$ is defined up to terms proportional to $\tilde{\nabla}_{(\mu} \zeta_{\nu)}$, where $\zeta_{\mu}$ depends linearly on $\varphi_{\mu\nu}$. These are field dependent diffeomorphisms and do not affect our results.

The second order variation of the action is

$$\delta_{\xi}^{(2)} S = \int d^4 x \left( \frac{\delta S^{(3)}}{\delta \varphi_{\mu\nu}} \delta_{\xi}^{(0)} \varphi_{\mu\nu} + \frac{\delta S^{(2)}}{\delta \varphi_{\mu\nu}} \delta_{\xi}^{(1)} \varphi_{\mu\nu} \right)$$

$$+ \frac{\delta S^{(2)}}{\delta \tilde{g}_{\mu\nu}} \delta_{\xi}^{(0)} \tilde{g}_{\mu\nu} + \frac{\delta S^{(0)}}{\delta \tilde{g}_{\mu\nu}} \delta_{\xi}^{(2)} \tilde{g}_{\mu\nu} \right).$$

(7.24)
As before, the last term vanishes on the de Sitter background. Requiring that \( \delta^{(2)} S \) is zero in this background therefore allows us to determine the first order transformation of the massive field. The result is

\[
\delta^{(1)} \xi |\dd S = -\lambda_1 \left( \tilde{\nabla}_{(\mu} \varphi_{\nu)} \tilde{\nabla}^\rho \xi - \frac{1}{2} \tilde{\nabla}_\rho \varphi_{\mu\nu} \tilde{\nabla}^\rho \xi - \frac{\Lambda g}{3} \varphi_{\mu\nu} \xi \right), \tag{7.25}
\]

up to terms of the form \( \delta^{(0)} \xi \varphi_{\mu\nu} \) where \( \tilde{\xi} = \varphi \xi \). These are field dependent PM transformations which do not affect our results. This result can be extended away from the de Sitter background in the same manner as before, resulting in the transformations

\[
\delta^{(1)} \varphi_{\mu\nu} = -\lambda_1 \left( \tilde{\nabla}_{(\mu} \varphi_{\nu)} \varphi_{\nu} - \frac{1}{2} \tilde{\nabla}_\nu \varphi_{\mu\nu} \tilde{\nabla}^\nu \xi - \frac{\Lambda g}{3} \varphi_{\mu\nu} \xi \right), \tag{7.26}
\]

\[
\delta^{(2)} \tilde{g}_{\mu\nu} = -\lambda_1 \frac{\alpha^2}{c^2} \left( \tilde{\nabla}_{(\mu} \varphi_{\nu)} \varphi_{\nu} - \frac{1}{2} \tilde{\nabla}_\nu \varphi_{\mu\nu} \tilde{\nabla}^\nu \xi \right) \varphi_{\sigma\rho} \tilde{\nabla}^\sigma \tilde{\nabla}_\rho \xi, \tag{7.27}
\]

under which \( \delta^{(2)} S \) vanishes. After performing a nonlinear field redefinition and taking the \( \delta^{(0)} \tilde{g}_{\mu\nu} \) transformation into account, it can be seen that these transformations match the ones derived in \([56]\).

We now consider the third order variation of the action, which is given by

\[
\delta^{(3)} S = \int d^4 x \left( \frac{\delta S^{(4)}}{\delta \varphi_{\mu\nu}} \delta^{(1)} \varphi_{\mu\nu} + \frac{\delta S^{(3)}}{\delta \varphi_{\mu\nu}} \delta^{(0)} \varphi_{\mu\nu} + \frac{\delta S^{(2)}}{\delta \varphi_{\mu\nu}} \delta^{(2)} \varphi_{\mu\nu} \\
+ \frac{\delta S^{(3)}}{\delta \tilde{g}_{\mu\nu}} \delta^{(0)} \tilde{g}_{\mu\nu} + \frac{\delta S^{(2)}}{\delta \tilde{g}_{\mu\nu}} \delta^{(1)} \tilde{g}_{\mu\nu} + \frac{\delta S^{(0)}}{\delta \tilde{g}_{\mu\nu}} \delta^{(3)} \tilde{g}_{\mu\nu} \right). \tag{7.28}
\]

Once again we start by putting this on the de Sitter background, where the last term vanishes. We then attempt to determine \( \delta^{(2)} \varphi_{\mu\nu} |\dd S \), the only undetermined quantity in the above equation, by demanding that the third order variation vanishes in this background. However, it turns out that there is no second order transformation of \( \varphi_{\mu\nu} \) which makes the variation in (7.28) vanish. In particular, there are terms of the form \( \int d^4 x \Lambda g \varphi \tilde{\nabla} \varphi \tilde{\nabla} \varphi \) that cannot be canceled by any such transformation, which was also the case in massive gravity, as shown by \([52]\).

Despite the \( \delta^{(n)} \tilde{g}_{\mu\nu} \) terms, which are not present in massive gravity, it therefore turns out not to be possible to extend the PM symmetry beyond cubic order in the action for the candidate bimetric theory. At the nonlinear level, bimetric theory therefore propagates 7 d.o.f even for the \( \beta_n \) in (7.9).
7.3.3 Generalization of the quartic action

Since the PM symmetry cannot be extended beyond cubic order for the candidate bimetric theory, it is natural to ask if there is any theory of two interacting spin-2 fields for which this is possible. In order to test this, we start by writing down the most general two derivative quartic action in a de Sitter background. Schematically, it looks like

\[ \hat{S}^{(4)} = \int d^4x \sqrt{-g} \left( \phi^2 \nabla_\phi \nabla_\phi \phi + \Lambda g \phi^4 \right), \]  

where all possible index contractions are made and all terms are given arbitrary coefficients. Note that this action will, in general, suffer from the Boulware-Deser ghost. The third order variation in (7.28) is then generalized by replacing \( S^{(4)} \), whose Lagrangian is given by (A.3), with the generic action \( \hat{S}^{(4)} \). This variation then contains two undetermined quantities, \( \hat{S}^{(4)} \) and the variation \( \delta^{(2)} \phi_{\mu\nu} \). We then investigate if there exists any combination of coefficients in (7.29) and \( \delta^{(2)} \phi_{\mu\nu}|_{dS} \) such that the third order variation of the action vanishes in the de Sitter background. The result is the same as for the candidate bimetric theory, there are terms of the form \( \int d^4x \Lambda g \phi \nabla_\phi \nabla_\phi \phi \) which cannot be canceled, regardless of the form of \( \hat{S}^{(4)} \). Our conclusion is therefore that the PM symmetry cannot be extended beyond cubic order in the action for any two derivative theory where the field content consists of two spin-2 fields.

7.3.4 Global symmetries

In this section we study the global symmetries of the candidate PM bimetric theory. The global symmetries are the isometries of the background metric \( \bar{g}_{\mu\nu} \) which also leaves the background of the massive field, \( \phi_{\mu\nu} = 0 \), invariant. For instance, consider the lowest order diffeomorphisms acting on the de Sitter background \( \bar{g}_{\mu\nu} \)

\[ \bar{g}_{\mu\nu} \rightarrow \bar{g}_{\mu\nu} + 2\nabla_{(\mu} \xi_{\nu)}. \]  

The background is left invariant by transformations satisfying \( \nabla_{(\mu} \xi_{\nu)} = 0 \), i.e. Killing’s equation. This leads to a global \( SO(1,4) \) symmetry group. We now want to investigate if this symmetry group is enhanced when the PM transformations are also included.

In [56] the global symmetries of a theory of a PM spin-2 field coupled to gravity are investigated. It is shown that, for their transformations, the global symmetry algebra is \( SO(1,5) \) to lowest order in the fields. However,
at higher orders the gauge transformations do not satisfy this algebra. Since our $\delta^{(1)}_{\xi} \varphi_{\mu\nu}$- and $\delta^{(0)}_{\xi} \bar{g}_{\mu\nu}$-transformations differ from those of [56] it is worth checking if this result also holds for our transformations.

To start with, we write down the transformations of $h_{\mu\nu}$ and $\varphi_{\mu\nu}$ to linear order in the fields. Using (7.18), (7.22), (7.23) and (7.26) and expanding around the de Sitter background yields the linear PM transformations

\[
\delta_{\xi} h_{\mu\nu} = \lambda_1 \bar{\nabla}_\mu \bar{\nabla}_\nu \xi - \frac{\lambda_1}{2} (2 \bar{\nabla}_{(\mu} h_{\nu)\rho} - \bar{\nabla}_\rho h_{\mu\nu}) \bar{\nabla}^\rho \xi
- \lambda_2 (2 \bar{\nabla}_{(\mu} \varphi_{\nu)\rho} - \bar{\nabla}_\rho \varphi_{\mu\nu}) \bar{\nabla}^\rho \xi,
\]

\[
\delta_{\xi} \varphi_{\mu\nu} = \left( \bar{\nabla}_\mu \bar{\nabla}_\nu + \frac{A_g}{3} \bar{g}_{\mu\nu} \right) \xi - \frac{1}{2} (2 \bar{\nabla}_{(\mu} h_{\nu)\rho} - \bar{\nabla}_\rho h_{\mu\nu}) \bar{\nabla}^\rho \xi
+ \frac{A_g}{3} h_{\mu\nu} \xi + \lambda_1 \frac{A_g}{3} \varphi_{\mu\nu} \xi - \frac{\lambda_1}{2} (2 \bar{\nabla}_{(\mu} \varphi_{\nu)\rho} - \bar{\nabla}_\rho \varphi_{\mu\nu}) \bar{\nabla}^\rho \xi,
\]

where

\[
\lambda_2 = \frac{\alpha^2}{2c^2}.
\]

The diffeomorphisms, denoted by $\tilde{\delta}_{\zeta}$, are given by

\[
\tilde{\delta}_{\zeta} h_{\mu\nu} = 2 \bar{\nabla}_{(\mu} \zeta_{\nu)} + \zeta^\sigma \bar{\nabla}_\sigma h_{\mu\nu} + 2 \bar{\nabla}_{(\mu} \zeta^\sigma h_{\nu)\sigma},
\]

\[
\tilde{\delta}_{\zeta} \varphi_{\mu\nu} = \zeta^\sigma \bar{\nabla}_\sigma \varphi_{\mu\nu} + 2 \bar{\nabla}_{(\mu} \zeta^\sigma \varphi_{\nu)\sigma}.
\]

If we let $O$ denote either $h_{\mu\nu}$ or $\varphi_{\mu\nu}$, we observe that, to zeroth order in the fields, the diffeomorphisms and PM transformations form a closed algebra,

\[
[\delta_{\zeta}, \delta_{\eta}] O = \delta_{\chi} O, \quad [\delta_{\xi}, \tilde{\delta}_{\eta}] O = \delta_{\tau} O, \quad [\delta_{\xi}, \delta_{\beta}] O = \tilde{\delta}_{\theta} O,
\]

where the parameters $\chi^\mu$, $\tau$ and $\theta^\mu$ are given in terms of the other transformation parameters. Next we need to find parameters $\zeta^\mu$ and $\xi$ that leave the background $\bar{g}_{\mu\nu} = \bar{g}_{\mu\nu}$, $\varphi_{\mu\nu} = 0$ invariant. It turns out not to be possible to make the background invariant under diffeomorphisms and PM transformations independently (other than for the trivial PM transformation given by $\xi = 0$). However, as mentioned previously, the zeroth order PM transformation of $\bar{g}_{\mu\nu}$ resembles a diffeomorphism. We can therefore define the transformation

\[
\delta'_{\xi} = \delta_{\xi} - \tilde{\delta}_{\epsilon},
\]

where $\epsilon^\mu = -\frac{\lambda_1}{2} \bar{\nabla}^\mu \xi$, which by definition has the property that, to lowest order, $\delta'_{\xi} \bar{g}_{\mu\nu} = 0$. Note that the algebra (7.36) is still satisfied if $\delta_{\xi}$ is
replaced by $\delta'_{\xi}$. It follows that the background is invariant under transformations which satisfy
\[
\tilde{\delta}_{\xi} g_{\mu\nu} = 2 \nabla_{(\mu} \xi_{\nu)} = 0, 
\]
\[
\delta'_{\xi} \varphi_{\mu\nu} = \left( \nabla_{\mu} \nabla_{\nu} + \frac{A_g}{3} \bar{g}_{\mu\nu} \right) \xi = 0, 
\]
(7.38)
(7.39)
to which there are nontrivial solutions. These solutions lead to a global $SO(1,5)$ symmetry algebra, as shown in [56]. In order to investigate what happens at higher orders, we write down the $\delta'_{\xi}$ transformations to linear order in the fields. They are given by
\[
\delta'_{\xi} h_{\mu\nu} = -\lambda_2 \left( 2 \nabla_{(\mu} \xi_{\nu)} - \nabla_{\rho} \xi_{\mu\nu} \right) \nabla^\rho \xi, \quad (7.40)
\]
\[
\delta'_{\xi} \varphi_{\mu\nu} = \left( \nabla_{\mu} \nabla_{\nu} + \frac{A_g}{3} \bar{g}_{\mu\nu} \right) \xi - \frac{1}{2} \left( 2 \nabla_{(\mu} h_{\nu)} - \nabla_{\rho} h_{\mu\nu} \right) \nabla^\rho \xi 
+ \frac{A_g}{3} h_{\mu\nu} \xi + \lambda_1 \frac{A_g}{3} \varphi_{\mu\nu} - \frac{\lambda_1}{2} \left( 2 \nabla_{(\mu} \varphi_{\nu)} - \nabla_{\rho} \varphi_{\mu\nu} \right) \nabla^\rho \xi 
+ \frac{\lambda_1}{2} \nabla^\sigma \xi \nabla_{(\mu} \varphi_{\nu)\sigma} + 2 \nabla_{(\mu} \nabla^\sigma \xi \varphi_{\nu)\sigma}. \quad (7.41)
\]
For parameters $\xi$ that solve (7.39) these transformations reduce to those of [56], which means that at linear order in the fields the commutator $[\delta'_{\xi}, \delta'_{\beta}] \varphi_{\mu\nu}$ does not close. Hence, at higher orders the transformations do not satisfy the $SO(1,5)$ algebra.

The above result does not apply to the special case where $\alpha^2 c^2 = -1$. However, by inspecting (5.19) and (5.20) we note that in this case the diagonalization into a massless and a massive mode is invalid, meaning that our calculations are not trustworthy. It is worth noting that for $\alpha^2 c^2 = -1$ the bimetric field equations mimic the equations of conformal gravity to lowest order in derivatives [57,58].
Chapter 8

Summary and outlook

In this thesis, various issues regarding the Hassan-Rosen bimetric theory have been investigated. The results are summarized and discussed below.

8.1 Paper I

A complete canonical analysis of the theory was performed in the metric formulation. After a change of variables, the Lagrangian becomes linear in the lapses of both metrics, as well the shift of one metric, resulting in five constraints. It was explicitly shown that the conservation of two of these constraints in time gives rise to one additional constraint, \( C_2 = 0 \). Contrary to the claim of [16, 17], \( C_2 \) is independent of lapses and shifts, and is therefore an appropriate constraint to eliminate the ghost. In addition, the conservation of \( C_2 \) in time results in a linear relation between the lapses of \( g_{\mu\nu} \) and \( f_{\mu\nu} \). The six constraints, together with general coordinate transformations, are sufficient to ensure that the bimetric theory propagates seven degrees of freedom (corresponding to one massless and one massive spin-2 field) and is free of the Boulware-Deser ghost.

Additionally, solving for one lapse in terms of the other using the linear relation mentioned above, the Hamiltonian density takes the form

\[
\mathcal{H} = -M\mathcal{R} - M^i\mathcal{R}_i, \quad (8.1)
\]

where \( M \) and \( M^i \) are Lagrange multipliers and \( \mathcal{R} \) and \( \mathcal{R}_i \) are first class constraints. It was shown that there are at least two different ways of expressing the Hamiltonian in the form (8.1), with \( M \) and \( M^i \) being either the lapse and shift of \( g_{\mu\nu} \) or \( f_{\mu\nu} \). In each of these cases \( \mathcal{R} \) is different. However, in both cases \( \mathcal{R} \) and \( \mathcal{R}_i \) are the generators of diffeomorphisms,
since they satisfy the associated Poisson algebra. From this algebra, a metric (the so called geometrical metric) can be identified, according to the prescription in [18]. The two different ways of expressing the first class constraints described above results in either \( g_{\mu\nu} \) or \( f_{\mu\nu} \) as the geometrical metric, meaning that this metric is not unique in bimetric theory. This can be understood by noting that if \( M \) and \( M^i \) are the lapse and shift of \( g_{\mu\nu} \), \( \hat{\mathcal{R}} \) generates diffeomorphisms orthogonal to the spacelike hypersurface with respect to \( g_{\mu\nu} \), which is reflected in the constraint algebra. The same holds for \( f_{\mu\nu} \) if \( M \) and \( M^i \) are its lapse and shift. In particular, note that the geometrical metric need not necessarily coincide with the gravitational metric, since it’s possible to couple matter to \( g_{\mu\nu} \) while choosing variables so that \( f_{\mu\nu} \) appears as the geometrical metric, or the other way around.

The nonuniqueness of the geometrical metric in bimetric theory naturally raises the question if \( g_{\mu\nu} \) and \( f_{\mu\nu} \) are the only possible geometrical metrics, or if there are others. One possibility might the the geometrical mean metric, which is given by \( h_{\mu\nu} = g_{\mu\lambda} S^\lambda_{\nu} \) [15, 34]. It is possible to write the Hamiltonian on the form (8.1) with \( M \) and \( M^i \) being the lapse and shift of \( h_{\mu\nu} \). In principle one could compute the Poisson brackets to see if \( \hat{\mathcal{R}} \) satisfies the diffeomorphism algebra with the spatial part of \( h_{\mu\nu} \) appearing in the \( \{ \hat{\mathcal{R}}, \mathcal{R} \} \) bracket. This brackets turns out to be difficult to compute however, and so far we have been unable to do it. Another potential candidate for a geometrical metric is the one whose spatial part is identified in [36] using the chiral formulation of bimetric theory.

### 8.2 Paper II

In Paper I it was shown that there is a linear relation between the lapses of the two metrics, where their ratio only depends on the dynamical variables. However, due to the complicated form of the Poisson brackets, this ratio was not computed explicitly. However, in order to solve initial value problems in bimetric theory, the explicit form of the ratio of lapses is necessary. In Paper II this ratio is therefore computed for the special case where the metrics share the same spherical symmetry, which makes the equations simplify enough to be manageable. While an explicit calculation of the ratio in the most general case is likely to prove very difficult, there might be other special cases where the calculation also simplifies sufficiently. One possibility is to calculate the ratio in the \( \beta_1 \)-model, since there is possible to explicitly solve for the shift like variable \( n^i \), like in the spherically symmetric case.

Initial value problems in bimetric theory can generally only be solved
numerically. Paper II can be seen as part of a larger project to make such numerical computations possible. Other steps that have been taken involve ensuring that the constraints propagate stably [59], establishing the BSSN formulation of the field equations [60], solving the constraints to determine initial data [61], finding appropriate gauge conditions [62] and developing computational tools [63]. The result of Paper II has been used in simulations of gravitational collapse of spherically symmetric dust [61].

8.3 Paper III

A unique candidate for a nonlinear partially massless bimetric theory was found in [20]. In Paper III we express this theory in terms of the massless and massive spin-2 fluctuations and resum the massless perturbation, resulting in an action for a massive spin-2 fluctuation nonminimally coupled to a metric. This is done in order to investigate if the PM gauge symmetry can be extended beyond the cubic order in the massive field. For each order in the action, we attempt to find PM transformations of the metric and the massive field such that the action is invariant to that order. It turns out that there are no transformations that leave the quartic order action invariant. Even generalizing the quartic terms in the Hassan-Rosen action to the most general two derivative quartic terms does not allow us to circumvent this result. We therefore conclude that there is no nonlinear two derivative PM theory of two interacting spin-2 fields. We also investigate the global symmetries of the candidate PM theory and show that while these form an $SO(1,5)$ algebra at the lowest order in the fields, this algebra is not satisfied at higher orders, consistent with the result of [56].

A theory where the PM symmetry is realized to all orders is found in [64]. Motivated in part by the lowest order global $SO(1,5)$ symmetry of the candidate PM bimetric theory, the authors construct a gauge theory of the $SO(1,5)$ group. The approach is similar to the construction of general relativity and conformal gravity as gauge theories of $SO(1,4)$ and $SO(2,4)$, respectively [65,66]. In addition to two spin-2 fields, this theory contains an extra vector field which couples to them in a nonminimal way. The presence of this vector field allows the theory to circumvent the no-go result of Paper III. The spin-2 fields have an interaction term of the same form as in the Hassan-Rosen bimetric theory, with the $\beta_n$ given by (7.9), but the kinetic terms are not of the Einstein-Hilbert form. Whether this theory is ghost free or not remains an open question, which could potentially be addressed by a Hamiltonian analysis. The form of the kinetic terms does however make this highly nontrivial.
Appendix A

Cubic and quartic Lagrangian densities

The explicit expressions for the cubic and quartic terms in (7.12) are given here, with the constants $\lambda_3$ and $\lambda_4$ given by

$$\lambda_3 = \frac{\alpha^2}{2c^4} (\alpha^2 c^2 - 1), \quad \lambda_4 = \frac{\alpha^2}{32c^6} (1 - \alpha^2 c^2 + \alpha^4 c^4). \quad (A.1)$$

$$\mathcal{L}_3 = \lambda_3 \sqrt{-\tilde{g}} \left\{ -\frac{1}{2} \varphi^\rho\sigma \tilde{\nabla}_\rho \varphi \tilde{\nabla}_\sigma \varphi \varphi_\gamma \varphi_\lambda + \frac{1}{2} \varphi^\rho\sigma \tilde{\nabla}_\rho \varphi \tilde{\nabla}_\sigma \varphi - \varphi^\rho\sigma \tilde{\nabla}_\rho \varphi \tilde{\nabla}_\sigma \varphi \tilde{\nabla}_\gamma \varphi \gamma \\
- \varphi^\rho\sigma \tilde{\nabla}_\sigma \varphi_\rho \tilde{\nabla}_\gamma \varphi_\gamma + \varphi^\rho\sigma \tilde{\nabla}_\gamma \varphi_\sigma \tilde{\nabla}_\rho \varphi_\gamma - \frac{1}{4} \varphi \tilde{\nabla}_\gamma \varphi \tilde{\nabla}_\gamma \varphi \\
- \varphi^\rho\sigma \tilde{\nabla}_\gamma \varphi_\rho \tilde{\nabla}_\lambda \varphi_\gamma + \frac{1}{2} \varphi \tilde{\nabla}_\gamma \varphi \tilde{\nabla}_\lambda \varphi_\lambda \\
+ 2 \varphi^\rho\sigma \tilde{\nabla}_\sigma \varphi_\rho \tilde{\nabla}_\lambda \varphi_\gamma \gamma + \varphi^\rho\sigma \tilde{\nabla}_\gamma \varphi_\sigma \lambda \tilde{\nabla}_\varphi_\gamma \varphi_\gamma - \varphi^\rho\sigma \tilde{\nabla}_\lambda \varphi_\sigma \gamma \tilde{\nabla}_\varphi_\gamma \varphi_\lambda \\
- \frac{1}{2} \varphi \tilde{\nabla}_\gamma \varphi_\gamma \varphi_\lambda \tilde{\nabla}_\varphi_\gamma \varphi_\gamma + \frac{1}{4} \varphi \tilde{\nabla}_\lambda \varphi_\gamma \varphi_\gamma \tilde{\nabla}_\varphi_\gamma \varphi_\gamma \\
+ \frac{1}{4} \left( \tilde{R}^\rho\sigma - \frac{1}{6} \tilde{R} \tilde{g}^\rho\sigma \right) \left( 8 \varphi_\rho \gamma \varphi_\sigma \varphi_\lambda \varphi_\gamma + 2 \varphi^\rho\sigma \varphi_\gamma \varphi_\gamma \varphi_\lambda \\
- 4 \varphi_\rho \gamma \varphi_\sigma \gamma \varphi + \varphi_\rho_\sigma \varphi_\rho \varphi \right) \\
+ \frac{1}{12} \Lambda_g \left( 4 \varphi_\rho \gamma \varphi_\sigma \gamma \varphi - 8 \varphi_\sigma \gamma \varphi_\phi \varphi \right) \} \quad (A.2)$$
\[
\mathcal{L}_4 = \lambda_4 \sqrt{-\tilde{g}} \left\{ \frac{\Lambda_g}{6(1 - \alpha^2 c^2 + \alpha^4 c^4)} \left[ 4(5 - 2\alpha^2 c^2 + 5\alpha^4 c^4) \varphi_\rho \varphi^\rho \varphi_\lambda \varphi^\lambda - 4(1 + \alpha^2 c^2) \varphi_\rho \varphi^\rho \varphi_\gamma \varphi_\lambda - (1 + 6\alpha^2 c^2 + \cdots) \varphi_\rho \varphi^\rho \varphi_\mu \varphi^\mu - 4\varphi_\rho \varphi^\rho \varphi_\lambda \varphi^\lambda \right] + \frac{1}{3} \left( \tilde{R}^{\rho\sigma} - \frac{1}{8} \tilde{g}^{\rho\sigma} \right) \left( 96\varphi_\rho \gamma \varphi_\sigma \varphi_\lambda \varphi_\mu - 16\varphi_\rho \varphi_\sigma \varphi_\lambda \varphi_\mu - 48\varphi_\rho \gamma \varphi_\sigma \varphi_\lambda \varphi_\mu \right) \right\}. \tag{A.3}
\]
Bibliography


[34] M. Kocic, Geometric mean of bimetric spacetimes, 1803.09752.


Part II
Papers