Cosmic tests of massive gravity

Jonas Enander

Doctoral Thesis in Theoretical Physics

Department of Physics
Stockholm University
Stockholm 2015

Stockholm
University

Doctoral Thesis in Theoretical Physics

Cosmic tests of massive gravity

Jonas Enander

Oskar Klein Centre for
Cosmoparticle Physics

and

Cosmology, Particle Astrophysics
and String Theory

Department of Physics
Stockholm University
SE-106 91 Stockholm

Stockholm, Sweden 2015

Stockholm University

Cover image: Painting by Niklas Nenzén, depicting the search for massive gravity.

ISBN 978-91-7649-049-5 (pp. i–xviii, 1–104)

Printed by Universitetsservice US-AB, Stockholm, Sweden, 2015.

Typeset in pdfLaTeX

So inexhaustible is nature's fantasy, that no one will seek its company in vain.

Novalis, *The Novices of Sais*

Abstract

Massive gravity is an extension of general relativity where the graviton, which mediates gravitational interactions, has a non-vanishing mass. The first steps towards formulating a theory of massive gravity were made by Fierz and Pauli in 1939, but it took another 70 years until a consistent theory of massive gravity was written down. This thesis investigates the phenomenological implications of this theory, when applied to cosmology. In particular, we look at cosmic expansion histories, structure formation, integrated Sachs-Wolfe effect and weak lensing, and put constraints on the allowed parameter range of the theory. This is done by using data from supernovae, the cosmic microwave background, baryonic acoustic oscillations, galaxy and quasar maps and galactic lensing.

The theory is shown to yield both cosmic expansion histories, galactic lensing and an integrated Sachs-Wolfe effect consistent with observations. For the structure formation, however, we show that for certain parameters of the theory there exists a tension between consistency relations for the background and stability properties of the perturbations. We also show that a background expansion equivalent to that of general relativity does not necessarily mean that the perturbations have to evolve in the same way.

Key words: Modified gravity, massive gravity, cosmology, dark energy, dark matter

Svensk sammanfattning

Massiv gravitation är en vidareutveckling av den allmänna relativitetsteorin där gravitonen, som förmedlar den gravitationella växelverkan, har en massa. De första stegen till att formulera en teori för massiv gravitation togs av Fierz och Pauli 1939, men det tog ytterligare 70 år innan en konsistent teori för massiv gravitation skrevs ned. Denna avhandling undersöker de fenomenologiska konsekvenserna av denna teori, när den används inom kosmologi. Vi studerar i synnerhet kosmiska expansionhistorier, strukturformation, den integrerade Sachs-Wolfe effekten och svag linsning, samt sätter gränser på teorins parametervärden. Detta görs med hjälp av data från supernovor, den kosmiska bakgrundsstrålningen, baryonisk-akustiska oscillationer, galax- och kvasarkataloger och galaktisk linsing.

Vi visar att teorin ger kosmiska expansionshistorier, galaktisk linsning och en integrerad Sachs-Wolfe effekt som alla överensstämmer med observationer. För vissa parametrar finns dock en spänning mellan konsistensrelationer för bakgrunden och stabilitetsegenskaper hos perturbationerna. Vi visar även att en bakgrundsexpansion som är ekvivalent med den hos allmän relativitetsteori inte nödvändigtvis betyder att perturbationerna utvecklas på samma sätt.

Nyckelord: Modifierad gravitation, massiv gravitation, kosmologi, mörk energi, mörk materia

List of Accompanying Papers

Paper I Mikael von Strauss, Angnis Schmidt–May, **Jonas Enander**, Edvard Mörtsell & S.F. Hassan. *Cosmological solutions in bimetric gravity and their observational tests*, JCAP **03**, 042 (2012) arXiv:1111.1655.

Paper II Marcus Berg, Igor Buchberger, **Jonas Enander**, Edvard Mörtsell & Stefan Sjörs. *Growth histories in bimetric massive gravity*, JCAP **12**, 021 (2012) arXiv:1206.3496.

Paper III **Jonas Enander** & Edvard Mörtsell. *Strong lensing constraints on bimetric massive gravity*, JHEP **1310**, 031 (2013) arXiv:1306.1086.

Paper IV **Jonas Enander**, Adam R. Solomon, Yashar Akrami & Edvard Mörtsell. *Cosmic expansion histories in massive bigravity with symmetric matter coupling*, JCAP **01**, 006 (2015) arXiv:1409.2860.

Paper V Adam R. Solomon, **Jonas Enander**, Yashar Akrami, Tomi S. Koivisto, Frank Könnig & Edvard Mörtsell. *Cosmological viability of massive gravity with generalized matter coupling*, JCAP **04**, 027 (2015) arXiv:1409.8300.

Paper VI **Jonas Enander**, Yashar Akrami, Edvard Mörtsell, Malin Renneby & Adam R. Solomon. *Integrated Sachs-Wolfe effect in massive bigravity*, Phys. Rev. D **91**, 084046 (2015) arXiv:1501.02140.

Papers not included in this thesis:

Paper VII **Jonas Enander** & Edvard Mörtsell. *On the use of black hole binaries as probes of local dark energy properties*, Phys. Lett. B **11**, 057 (2009) arXiv:0910.2337.

Paper VIII Michael Blomqvist, **Jonas Enander** & Edvard Mörtsell. *Constraining dark energy fluctuations with supernova correlations*, JCAP **10**, 018 (2010) arXiv:1006.4638.

Acknowledgments

Academic life suffers from the fact that the results of something inherently collective, i.e. science, is presented as an individual accomplishment. This is particularly true for a doctoral thesis. My warmest gratitude therefore goes out to a lot of people. First and foremost my supervisor Edvard Mörtsell: always available for discussion, and with a helpful and relaxed attitude.

My collaborators—Yashar Akrami, Marcus Berg, Michael Blomqvist, Igor Buchberger, Fawad Hassan, Tomi Koivisto, Frank Könnig, Angnis Schmidt-May, Stefan Sjörs, Adam Solomon, Mikael von Strauss—have been invaluable in my research, not only scientifically, but also socially, and I want to thank all of them. Three of my collaborators—Angnis, Mikael and Stefan—were also my office mates for a long time, and together with my recent office mates Seméli Papadogiannakis and Raphael Ferretti they have made my life in physics so much more fun.

My stay at Fysikum has been great due to a large number of people. In general, the many persons in the CoPS group, Oskar Klein Centre, Dark Energy Working Group and IceCube drilling team, with whom I have interacted with in one way or another, have always made science a much more enjoyable endeavour. In particular—and besides my supervisor, office mates and collaborators—I have to mention all of the discussions and activities that I shared during these years with Joel, Rahman, Bo, Kjell, Ariel, Rachel, Martin, Ingemar, Tanja, Maja, Hans, Sören, Emma, Lars, Kristoffer, Olle and Kattis.

But there is a life outside of physics too. My family has been a source of constant support, for which I am deeply grateful. Love supreme also goes out to all of my friends, especially Shabane, Erik, Helena, Mattias, Niklas, Emma, Seba, Jonas, Ida, Mathias, Christian and Gül, who kept me on the right side in the conflict between work and play.

Preface

This thesis deals with the phenomenological consequences of the recent complete formulation of massive gravity. The thesis is divided into six parts. The first part is a general introduction to the current cosmological standard model, i.e. a universe dominated by dark matter and dark energy. The second part describes Einstein's theory of general relativity and certain extensions that are commonly investigated. In part three we describe, in more detail, massive gravity. Part four is the main part of this thesis, where the cosmic tests of massive gravity are presented. In part five we conclude and discuss possible future research directions. Part six contains the papers constituting this thesis.

A note on conventions: We put $c = 1$, and G is related to Planck's constant, which we keep explicit, as $M_g^2 = 1/8\pi G$. The metric convention is $(-,+,+,+)$. In the papers included in this thesis, different notational conventions for the two metrics g and f have been used. In the thesis, however, we will work with a single convention. The theory of massive gravity contains two rank-2 tensor fields g and f. We will refer to both of these as metrics, although it is only the tensor field that couples to matter that properly should be called a metric. Furthermore, the theory where f is a fixed metric is usually referred to as massive gravity (or the de Rham-Gabadadze-Tolley theory), and when f is dynamical this is dubbed massive bigravity (or the Hassan-Rosen theory). In this thesis we will not be too strict concerning this terminological demarcation, and refer to both cases as massive gravity. Bigravity will, however, always refer to the case of a dynamical f.

Contribution to papers

Paper I presents the equations of motion for the cosmological background expansion, studies their solutions and puts constraints on the parameter space of the theory using recent data. I participated in the derivation and study of the equations of motion. The main part of the paper was written by Mikael von Strauss.

Paper II derives and studies the equations of motion for cosmological perturbations. I worked on all aspects of the paper, in particular the derivation of the equations of motions and their gauge invariant formulation. Marcus Berg wrote the main part of the paper, but all authors contributed to the final version.

Paper III derives linearized vacuum solution for spherically symmetric spacetimes and presents the lensing formalism used for parameter constraints. I derived the first and second order solutions and did a lot of the analytical work. I wrote the main part of the paper, except for the data analysis which was written by Edvard Mörtsell.

Paper IV derives and analyses the equations of motion for the cosmological background expansion with a special type of matter coupling. I initiated and performed a large part of the derivation and analysis, and am the main author of the paper. The data analysis was written by Edvard Mörtsell.

Paper V studies the cosmological background solutions for the so-called dRGT formulation of massive gravity, and with a special type of matter coupling. I helped with the derivation and analysis of the equations of motion, and participated in the final formulation of the paper, which was mainly written by Adam Solomon.

Paper VI uses data from the cosmological microwave background and galaxy clustering to analyse the viability of massive gravity. I initiated the analysis, lead the numerical work and wrote the major part of the paper.

Jonas Enander
Stockholm, January 2015

Contents

Part I

Introduction

Chapter 1

Warming up with gravity

We are usually told that when we drop a ball it will fall down. Things are not so simple, however. Newtonian gravity was developed taking everyday experiences into account: the trajectory of cannon balls on Earth, the motion of the moon around the Earth and the motion of the Earth around the sun. For these type of motions, Newton succeeded in answering two questions: what kind of gravitational field does a given piece of mass create, and how does another piece of mass move in that field?

Einstein also wrestled with and gave an answer to these questions, and did so by enlarging the framework that Newton and Galileo had set up. It turned out that not only do different lumps of mass create a gravitational field, but all forms of energy do so. Even the random movements of atoms within a gas and the momentum they carry contribute to gravity. In everyday circumstances these effects are insignificant, but when considering very dense and hot objects, such as stars, or the early stages of the universe when it was dominated by light, these effects can become important.

It turns out that in the long run it is the vacuum itself that will dominate gravitationally. This might sound a bit weird; the vacuum is, by definition, empty, so how can it have anything to do with gravity? And shouldn't the interaction *between* things be more important than their interactions with nothing? Things are weirder yet; the gravitational effect of the vacuum is not to make things fall towards each other, but to make them move away from each other. Gravity becomes *repulsive.* And repulsive gravity, to the best of our current knowledge, is the future, dominant behaviour of the universe.

The reason that the vacuum gravitates is that it can have an energy density (yes, this means that if you empty a box of everything, there will still be some vacuum energy there that you can never get rid of). And it turns out that the pressure of the vacuum is negative, unlike the familiar case of a gas with positive pressure.

Gravity is thus a bit more involved than the dropping of balls would indicate. It can be both attractive and repulsive. All forms of energy, even the vacuum itself, will contribute to it. And since electric forces cancel on average due to the presence of equal amounts of positive and negative charge, gravity will dominate on scales larger than, say, one centimeter. In particular, it will determine the global behaviour of our universe.

That the vacuum dominates the long term behaviour of the universe was established in a series of observations during the end of the 20:th century. Because of this, there was a strong impulse to re-examine our understanding of gravity. Is general relativity really the correct description when it comes to the large-scale behaviour of the universe? How well-tested is general relativity, considering that the most precise experiments are carried out in the solar system? Theoretically, there were many avenues to explore when it came to extensions of general relativity. This thesis deals with one of them.

Many alternatives to general relativity that has been proposed during the last two decades were introduced on an application level, in order to address the question of the accelerated expansion of the universe. Massive gravity, which is the subject of this thesis, on the other hand, was introduced from purely theoretical consistency requirements. The original question addressed is rather simple. In general relativity, the particle that is responsible for gravitational interactions is the massless graviton. So what happens if one makes this particle massive? This scenario is well understood in the case of photons. To within great experimental precision, photons are massless. It is straightforward to write down the theory of a massive photon, and derive the observational signatures (this is exactly what has been done behind the statement "the photon is massless"; experimentalists have searched for the observational signatures of a massive photon and found none). But for the graviton, it was for a long time unclear how to write down the full theory which contained a massive graviton. The issue was resolved, however, in 2011. Building on earlier work, Fawad Hassan and Rachel Rosen, working at the Oskar Klein Centre in Stockholm, wrote down the complete theory of massive gravity [1, 2]. It was then only a question of working out the observational signatures of this theory, and testing that against data. The primary tests for any new theory of gravity are the cosmic expansion history, the amount of structure formed in the universe, the spectrum of the cosmic microwave background and bending of light around massive objects such as stars and galaxies. In this thesis, all of these tests are applied to massive gravity in order to see whether the theory can be excluded or not.

Before we go into the details of these tests, we give a brief account of the two major unsolved puzzles in cosmology: dark matter and dark energy.

Chapter 2

The universe as we know it

The science of cosmology has developed tremendously during the last hundred years. The major shift in our understanding of the universe is that it is not constant in time, but rather highly dynamical: the universe is expanding. This expansion causes the universe to cool down, which means that it must have been much hotter earlier. Since matter behaves quite differently at different temperatures—which the example of boiling water clearly shows—it is thus possible to talk about a *natural history* of the universe.

The large scale structure of the universe is to a large extent governed by two gravitational effects. On the one hand we have the expansion of space, which makes objects, such as galaxies, recede from one another. On the other hand, we have the attraction that gravity produces between two objects, which makes them move towards each other. Measuring the distribution and evolution of the large scale structure will thus give us information about the interplay between these two effects, and how we should model gravity in order to properly describe this interplay.

While the current level of the science of cosmology is truly a precision science—in large parts thanks to the use of CCD cameras—this level of precision has also cast doubts on our level of understanding of gravity. From measurements of the expansion of space and gravitational attraction we infer that the energy content of the universe to roughly 95% is unknown to us. We can see the gravitational effects of this content, but we do not know what it is. This content also has two completely different gravitational effects: Dark energy, which makes out about 70% of the total energy content, causes the expansion of the universe to accelerate. It has an repulsive effect on objects not gravitationally bound to each other, making them move apart at an increasing rate. Dark matter, on the other hand, is postulated from the fact that there appears to be "more" attractive gravity on scales ranging from galaxies to cluster of galaxies, than what one would infer from the observed baryonic content.[1] Let us review the evidence for these two unknown energy contents.

2.1 The case for more attractive gravity: dark matter

Dark matter is usually postulated to be an as of yet unobserved particle that does not interact with light. Such a particle is, by itself, not exotic; neutrinos are an example of a particle that interacts gravitationally and weakly with other particles, but not electromagnetically. The gravitational evidence for dark matter comes from the following areas:

Cosmic expansion Observation of the recession of galaxies from one another gives a constraint on the total energy content of the universe. This is further explained in section 7.1. These observations point towards the existence of a matter component which makes out roughly 25% of the total energy content. This matter component is not the same as the observed baryonic component, which only accounts for roughly 20% of the total matter content.

Cosmic microwave background Some unknown mechanism caused the universe to be in a hot state roughly 13.7 billion years ago. In this hot state all the particles that we know today intermingled: baryons, neutrinos, photons etc. The universe was completely smooth, with the exception of tiny fluctuations that would later form the seeds of galaxies. Baryons and photons interacted frequently, forming an ionized plasma. As the universe expanded, the plasma cooled down until it reached a stage where the baryons could form neutral hydrogen as the photons stopped interacting with them. These photons then started to free stream in the universe, and are referred to as the cosmic microwave background. Analyzing their temperature distribution today gives detailed information about the properties of the universe. In particular, they strongly point

[1] Baryons are particles consisting of three quarks, such as the proton and neutron. They exist in the universe mostly in the form of hydrogen, and make out roughly 5% of the energy content. By energy content we mean all forms of energy components: baryonic matter, dark matter, light, neutrinos, dark energy, curvature etc.

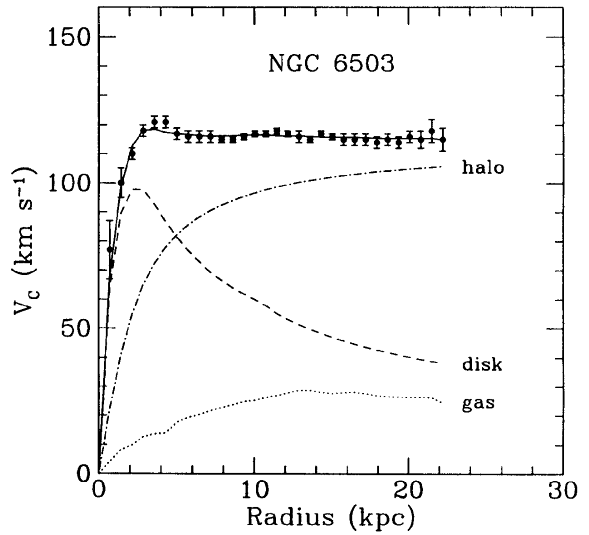

Figure 2.1. Observed rotational velocities of the galaxy NGC 6503 compared to the expected velocities as inferred from the luminous and gaseous baryonic mass. The dark matter component, in the form of a spherical halo, is needed to produce the observed rotational velocities. Plot taken from [3].

towards the existence of both dark matter and dark energy. The cosmic microwave background is described further in section 8.3 and 9.1.

Galaxy formation The fluctuations in the baryon-photon plasma described in the previous paragraph were extremely tiny at the time of decoupling. The baryonic component started to collapse, but this collapse would be far to slow to yield the galaxies and clusters of galaxies that we see today if it only involved the baryons. Therefore, there has to be a dark matter component that could start to grow earlier than the baryonic component. The growth of baryonic matter in the early universe is impeded through its coupling to light, so the dark matter component indeed has to be "dark" and not couple to light.[2]

Galaxy dynamics Galaxies are gravitationally bound systems, with stars and gas in bound orbits. Galaxies are usually gravitationally bound to

[2] A better wording would actually be "transparent", since an object is dark because of its absorption of light.

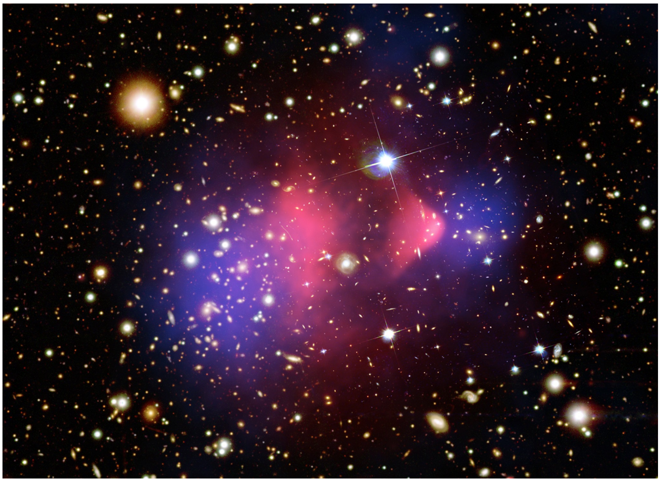

Figure 2.2. The Bullet cluster is the cluster to the right; it collided with the cluster to the left some 150 million years ago. The hot gas, shown in pink, has been displaced from the center of the clusters due to electromagnetic interaction. The dark matter distribution (blue), which is inferred from lensing, follows the cluster centers and is not affected by the interaction. Image Credit: NASA/CXC/CfA.

other galaxies in galaxy clusters. The observed motion of gas and stars in the outer regions of galaxies, and also the observed motion of galaxies in galaxy clusters, is not possible to produce through the baryonic mass distribution within the galaxy or galaxy cluster. One has to postulate the presence of more matter—once again dark matter, since it is not seen—in order to set up stronger gravitational fields. The rotation curves of gas and stars in the outer parts of spiral galaxies are nearly flat, and this can be produced with a dark matter halo.

Light bending Since matter curves the surrounding spacetime, the trajectories of light will also get curved as it passes nearby massive objects. This effect is described in detail in section 10.2. Comparing the mass distribution inferred from lensing effects around galaxies and galaxy clusters, with the observed mass distribution, gives yet another piece of evidence for the existence of dark matter.

This evidence is particularly striking in the case of e.g. the Bullet Cluster. In Fig. 2.2, the mass distribution inferred through lensing (blue) is shown for two interacting galaxy clusters, together with the mass distribution of the hot gas (pink). One clearly sees how the gas is displaced because of the interaction between the two clusters, whereas the non-interacting dark matter component resides in the center of the clusters.

It is important to note that none of the above properties tell us anything about the particle properties of the putative dark matter particle, expect that it should have non-relativistic velocities and not have any strong self-interactions. It is perfectly possible that the particle does not even exist, and that our theory of gravity is wrong, since it is basically the demand for stronger gravitational fields on galactic scales and above that calls for the introduction of a new particle. The consistent explanation for a wide range of phenomena that the postulation of a dark matter particle achieves is, however, a strong rationale for searching for it using non-gravitational means.

2.2 The case for more repulsive gravity: dark energy

One outstanding feature of the universe is that it is pretty empty, and that it appears to become more and more empty over time as it expands. The expansion rate of the universe is related to its energy content, and a major discovery in the 90:s, awarded with the Nobel Prize in 2011, was that the expansion rate of the universe is accelerating [4, 5]. A possible cause of the acceleration is an energy component, which makes out roughly 70% of the energy content of the universe, with negative pressure.

One candidate for this energy component is the vacuum itself. The vacuum is normally thought of as empty and thus without any physical influence. Gravity, at least in the manner that Einstein introduced, couples to all forms of energy, and if the vacuum has an energy density this can have a gravitational influence. That the vacuum can have an energy density makes sense from the point of view of quantum field theory. Due to Heisenberg's uncertainty principle it is impossible for something to be at perfect rest; there will always be a small residual motion present. The energy associated with this motion is called the zero-point energy. Adding up all these energies gives, within the context of quantum field theory, an energy density of the vacuum. Back-of-the-envelope calculations suggest that this energy density should be on the order of the Planck scale, whereas the observed energy density of the vacuum is some 120 orders of magnitude smaller than that.

The dark energy component gives rise to a repulsive form of gravity, unlike the dark matter component which causes the more well-known case of an attractive gravitational influence. As we saw in the previous section, there are several independent gravitational probes that point towards the existence of dark matter. The only inference of a dark energy component, on the other hand, comes from the expansion history of the universe. Dark energy does not have to be caused by the vacuum. It is possible that our theory of gravity is incorrect; gravity might be insensitive to the properties of the vacuum, and some other gravitational effect, or even new matter fields, produce the accelerated expansion.

One of the key issues concerning dark energy is whether it is constant in space and time—the true hallmark of a vacuum contribution—or if it shows any variation. A non-constant dark energy component would e.g. participate in the formation process of clusters and galaxies, as well as alter the expansion history. Current observations are consistent with a time-varying dark energy component.

The explanation of the inferred existence of dark matter and dark energy are the two outstanding problems facing cosmology today. In the next part we will look deeper into the structure of Einstein's theory of general relativity and its possible extensions. We will then see why the solution to the dark energy problem is a rather subtle affair, and whether it could be addressed in the theory of massive gravity.

Part II

Beyond Einstein

Chapter 3

How general is general relativity?

General relativity was introduced to the scientific community in November 1915, when Einstein presented his field equations to the Prussian Academy of Science [6]. The major result of his new theory of gravity—developed together with mathematicians like Marcel Grossmann and David Hilbert—was that spacetime should not be regarded as a static background arena, but instead be promoted to a dynamical entity. Matter in its various forms, e.g. rest energy, stresses, kinetic energy, curves spacetime, which in turn influences the motion of matter.

Mathematically, the curvature of spacetime is encoded in the metric $g_{\mu\nu}$, in the sense that spacetime distances are given by $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$. Matter, and its various forms of energy, is represented in the so-called energy-momentum tensor $T_{\mu\nu}$. Einstein's field equations, which relate $g_{\mu\nu}$ to $T_{\mu\nu}$, are

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{T_{\mu\nu}}{M_g^2}, \tag{3.1}$$

where $R_{\mu\nu}$ is the Ricci tensor and R the Ricci scalar. M_g^2 determines the coupling strength of gravity to matter, and is related to Newton's constant through

$$M_g^2 = \frac{1}{8\pi G}. \tag{3.2}$$

The motion of test particles in the curved spacetime is on geodesics, i.e. trajectories that extremize the spacetime distance. Einstein derived his field equation by demanding that $T_{\mu\nu}$ should satisfy

$$\nabla^\mu T_{\mu\nu} = 0, \tag{3.3}$$

which is a condition related to energy-momentum conservation. Since

$$\nabla^\mu \left(R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right) = 0 \tag{3.4}$$

identically, one thus sees that energy-momentum conservation is a consequence of the field equations.

Einstein did not believe that his field equations were the end of the story. He wanted to incorporate the other forces of nature into a unified geometric description. This turned out to be difficult, and historically the theoretical and experimental investigations of the forces of nature divided into two parts. Due to the weakness of gravitational effects on microscopic scales, the particle physics community primarily studied the electromagnetic, weak and strong interactions and developed a theoretical framework, i.e. quantum field theory, suited for these interactions. The gravitational force was instead primarily studied by astronomers and cosmologist; theoretical developments crossed roads with the increased amount of astronomical data that poured in due to technological advances after the second world war. Unifying the treatment of gravity with that of the other forces is not only an experimental problem; the theoretical framework of gravity is very different from that of quantum field theory. Even though the electromagnetic and weak forces could eventually be joined into an electroweak description, there is as of yet no unification of gravity with the other forces.

Today there are both theoretical and observational reasons why one wants to go beyond Einstein's original description of gravity. The observational reasons—i.e. dark matter and dark energy—were described in the previous chapter. The theoretical reasons are threefold: 1) The relationship between the vacuum and gravity is hard to reconcile with the quantum description of vacuum fluctuations (this problem is also related to dark energy), 2) spacetime singularities seem to be ubiquitous and 3) the quantization of general relativity does not lead to a theory valid at all energy scales. In the next three sections we will briefly discuss these points, and in section 3.4 we describe possible extensions of general relativity.

3.1 The cosmological constant problem

Adding a term $\Lambda g_{\mu\nu}$ to Einstein's field equations does not spoil the identity (3.4), and energy-momentum conservation is thus still a consequence of the field equations. In terms of the expansion history of the universe, such a constant will cause the expansion to accelerate. It is therefore a candidate for dark energy. The possible sources for Λ are three-fold:[1] 1) The cosmological constant term Λ in the gravitational Lagrangian ($\mathcal{L} \sim R - 2\Lambda$), which is a geometrical term describing spacetime curvature in the absence of sources, 2) constant potential terms coming from symmetry breaking phase transitions in the early universe (for example the electroweak phase transition) and 3) zero-point fluctuations of quantum fields. These zero-point fluctuations correspond to the energy density of the vacuum.

The cosmological constant term Λ is a free parameter. Constant potential terms are also free in the sense that one can add any constant term to the potentials, but there will always be a shift in the potential minimum during phase transitions, which should affect the expansion history. Zero-point fluctuations of quantum fields generically give a too large value compared to what is observed, but this value can always be subtracted by a constant term in the Lagrangian. Such a subtraction, however, requires a high degree of fine tuning.

Given that Einstein's theory of gravity is correct, all of these three sources combine to produce an effective cosmological constant. If the observed accelerated expansion is due to the cosmological constant, it is measured to be

$$\Lambda \sim 10^{-122} \ell_P^{-2}, \tag{3.5}$$

where ℓ_P is the Planck length (equal to 1.61×10^{-23} cm in anthropocentric units).

An intriguing property concerning the value of the vacuum energy density is that since it is constant throughout the history of the universe, its effects are only becoming discernible as the radiation and matter energy densities have become diluted enough. Furthermore, in the future the universe will be completely dominated by the vacuum energy density and approach an asymptotic de Sitter universe. The situation is depicted in Fig. 3.1. In a dark energy dominated future, different regions will become causally disconnected, i.e. even in an infinite time span, light will

[1] A good review of the cosmological constant problem can be found in [7].

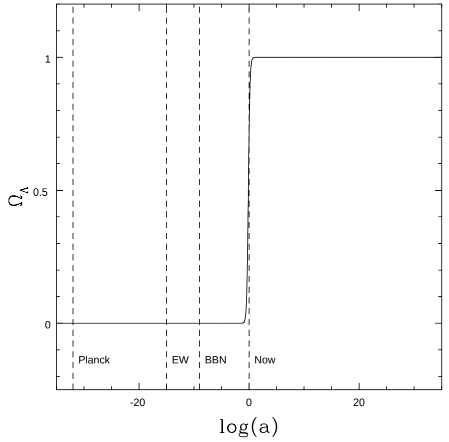

Figure 3.1. The contribution of the fractional energy Ω_Λ to the total energy content of the universe, as a function of the scale factor. Indicated are the Planck era, electroweak phase transition and time of big bang nucleosynthesis. Plot taken from [8].

not be able to go from one region to another. It will thus be impossible to reconstruct the natural history of the universe. In some sense then, we live in a privileged time where the matter and vacuum energy densities are of roughly equal size. This is usually referred to as the "cosmic coincidence problem", even though there is no consensus that it truly represents a physical problem.

The cosmological constant problem is in reality a host of problem:

- Given that the vacuum has an energy density, how does gravity couple to it?

- Why is there such a high degree of fine-tuning required to reconcile theoretical expectations with the observed value of the cosmological constant (if dark energy is due to a cosmological constant).

- Why does gravity seem to be insensitive to shifts in the constant potential terms that occurred in the early universe?

All of these problems are of course connected, and they strongly suggest that whatever theory of gravity that will extend general relativity, it has to shed new light on the relationship between gravity and the vacuum.

3.2 Singularities

General relativity uses the framework of a curved spacetime to describe gravity. Trajectories of particles in free fall occur on geodesics on that spacetime, and the position of the particle along its trajectory can be described by the proper time. It can occur, however, that for certain spacetimes these geodesics are not complete, in the sense that the geodesics can not be continued on the spacetime, even though the entire range of the parameter describing its trajectory has not been covered [9]. Such geodesic incompleteness signals the presence of a singularity. A well-known example is the singularity inside the Schwarzschild black hole, which can be reached in a finite time. It can be shown that these singularities exist under rather generic conditions. They are problematic in the sense that they signal a breakdown of the spacetime description of gravity, and suggest that close to these points one has to use some other framework than general relativity for describing gravitational interactions.

3.3 Renormalizability

One striking fact about physics is that it is possible to describe natural phenomena occurring at a certain length scale without having to care about the details at smaller scales. We can, for example, study the behaviour of water waves without taking the detailed interactions of the water molecules into account. In the context of particle physics, this means that we can describe particle interactions at a given energy even in the absence of a complete high-energy theory [10].

The exact relationship between an effective theory formulated for some given energy range, and its underlying high-energy completion, is encapsulated in the concept of renormalizability. Theories are classified as either renormalizable or non-renormalizable. A renormalizable theory does not contain an energy scale which signals its breakdown, i.e. where it needs to be replaced by some high-energy theory. A non-renormalizable theory, on the other hand, can only be treated as a low-energy effective theory. It has a built-in energy scale, which gives a limit to its regime of applicability. QCD is an example of the former, and general relativity an example of the latter. This means that the quantized version of general relativity can not be the fundamental theory of gravity (which also includes quantum mechanical effects). Its applicability as a quantum theory breaks down at the Planck scale. This is a strong theoretical hint that general relativity is only an effective, low-energy theory, and is thus not the final word on our understanding of gravity.

3.4 Theoretical avenues beyond Einstein

We saw in the previous section that general relativity suffers from certain drawbacks: it's not a renormalizable upon quantization, does not lead to singularity-free spacetimes and the relationship between gravity and the vacuum is unclear. Also, the inferred existence of dark matter might be because of our lack of understanding of gravity. Given these problems, it is natural to ask what the viable extensions of general relativity are. In 1965, Weinberg showed that general relativity is the *unique* Lorentz covariant theory of an interacting massless spin-2 particle [11]. The uniqueness property of general relativity was further investigated by Lovelock in [12, 13], where it was proven that Einstein's field equations are the unique, local, second order equations of motion for a single metric $g_{\mu\nu}$ in four dimensions. This means that modifications of general relativity must entail something of the following:[2]

1. Include higher derivatives than second order in the equations of motion.
2. Introduce new dimensions.
3. Use more fields than just $g_{\mu\nu}$.
4. Introduce non-local interactions.
5. Break Lorentz symmetry.
6. Abandon the metric description of gravity.

The first option is problematic, due to instabilities that arise when higher order derivatives are included. The existence of these instabilities was shown in the mid 19:th century by Ostrogradsky [15], and are most easily seen in the Hamiltonian picture. It turns out that under certain general conditions, higher order equations of motion leads to phase space orbits that are unconstrained, which means that negative energies can be reached. Coupling fields with higher order equations of motion to

[2]There exists several excellent reviews of modified gravity theories. In this chapter we have followed [14].

fields that have a Hamiltonian unbounded from below means that the latter fields can acquire arbitrarily large energies due to interactions with the former fields. Adding higher order curvature terms, such as R^2 or $R_{\mu\nu}R^{\mu\nu}$, to the Lagrangian yields higher order equations of motion for $g_{\mu\nu}$, which potentially can avoid the problematic Ostrogradsky instability [16]. These higher order terms are expected if one regards general relativity as an effective field theory, with the Ricci scalar being the dominant term at low energies. These higher order interactions have been studied as a possible explanation of dark energy [17].

The second option introduces new spatial dimensions that either have to be so small so that their effects are not seen in everyday life, or they have to be so large and only restricted to gravitational interactions so that their effects are only seen at cosmological distances. In braneworld models all particles are confined to a brane with three spatial dimensions, whereas gravity also has interactions in a higher-dimensional bulk. The gravitational interactions will appear four dimensional up to a crossover scale r_c, and beyond that they will be modified. The modifications depend on the precise bulk and brane setup [18].

The third option, to add more fields, is the avenue taken in massive gravity. Of course, one inevitably has to ask if the addition of more fields really is a modification of general relativity, or if it is just new matter fields one is adding. In massive gravity the extra field is $f_{\mu\nu}$. This means that one has two rank-2 tensor fields present in the theory, and it is an open question which one of them should be promoted to the metric used to measure spacetime distances. This is further discussed in chapter 5.

The remaining options—non-local interactions, breaking of Lorentz symmetry and abandoning the metric description—are rather exotic and has appeared in a variety of forms in the literature [20, 21, 22]. We shall have nothing further to say about them in this thesis.

Part III

Bimetric gravity

Chapter 4

The Hassan-Rosen theory

In this chapter we will introduce massive bigravity, also known as the Hassan-Rosen theory. In the first section we describe how general relativity is the theory of a massless spin-2 particle.[1] We then describe, in section 4.2, what the theory of a massive spin-2 particle should look like around a Minkowski background. In section 4.3 we discuss the road from the particle description in a Minkowski framework to the complete non-linear theory. We pay attention to how it is done in general relativity and the potential pitfalls when using a massive spin-2 particle. We finally state the full Hassan-Rosen theory in section 4.4. The question of how to couple matter to gravity, and its impact on the equivalence principles, is discussed in chapter 5, and important dualities, symmetries and parameter choices are described in chapter 6.

4.1 Massless spin-2

The Lagrangian for general relativity is given by

$$\mathcal{L} = -\frac{M_g^2}{2}\sqrt{-g}R + \mathcal{L}_m, \tag{4.1}$$

where M_g^2 is the coupling constant between matter (given by the matter Lagrangian $\mathcal{L}_m$) and gravity. To look at the spin-2 structure of general

[1] We base our exposition on [23] and [24].

relativity, we expand the Lagrangian around flat space by decomposing the metric as

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{4.2}$$

where $|h_{\mu\nu}| \ll 1$. The Lagrangian then becomes

$$\begin{aligned}\mathcal{L} = &- \frac{M_g^2}{4}\left(-\frac{1}{2}\partial_\lambda h_{\mu\nu}\partial^\lambda h^{\mu\nu} + \partial_\mu h_{\nu\lambda}\partial^\nu h^{\mu\lambda} - \partial_\mu h^{\mu\nu}\partial_\nu h + \frac{1}{2}\partial_\lambda h \partial^\lambda h\right) \\ &- \frac{1}{4}h_{\mu\nu}T^{\mu\nu},\end{aligned} \tag{4.3}$$

where $h = \eta^{\mu\nu}h_{\mu\nu}$, and the stress-energy tensor $T^{\mu\nu}$ is defined through

$$T_{\mu\nu} \equiv \frac{2}{\sqrt{-g}}\frac{\delta S_m}{\delta g^{\mu\nu}}. \tag{4.4}$$

The equations of motion are obtained by varying with respect to $h_{\mu\nu}$:

$$\mathcal{E}^{\alpha\beta}_{\mu\nu}h_{\alpha\beta} = -\frac{T_{\mu\nu}}{M_g^2}, \tag{4.5}$$

where

$$\mathcal{E}^{\alpha\beta}_{\mu\nu} = \left(\eta^{(\alpha}_\mu\eta^{\beta)}_\nu - \eta^{\alpha\beta}\eta_{\mu\nu}\right)\Box - \eta^\beta_\mu\partial^\alpha\partial_\nu - \eta^\beta_\nu\partial^\alpha\partial_\mu + \eta_{\mu\nu}\partial^\alpha\partial^\beta + \eta^{\alpha\beta}\partial_\mu\partial_\nu. \tag{4.6}$$

Applying ∂^μ on both sides gives

$$\partial^\mu T_{\mu\nu} = 0, \tag{4.7}$$

which is the flat space version of $\nabla^\mu T_{\mu\nu} = 0$. Energy-momentum conservation is "built into" general relativity from the beginning; we will see later how this gets modified in massive gravity. If we make a small shift of the coordinates $x^\mu \to x^\mu + \xi^\mu$, the field $h_{\mu\nu}$ transform as

$$h_{\alpha\beta} \to h_{\alpha\beta} + \partial_\alpha\xi_\beta + \partial_\beta\xi_\alpha. \tag{4.8}$$

The equations of motion are left invariant under this transformation, which is due to the reparametrization invariance of general relativity.

To show that the quadratic Lagrangian (4.3) describes a massless spin-2 field, we choose the Lorenz gauge where

$$\partial^{\mu}\left(h_{\mu\nu}-\frac{1}{2}\eta_{\mu\nu}h\right)=0. \tag{4.9}$$

The trace reversed equations of motion becomes

$$\Box h_{\mu\nu}=-\frac{1}{M_g^2}\left(T_{\mu\nu}-\frac{1}{2}\eta_{\mu\nu}T\right). \tag{4.10}$$

The solution for $h_{\mu\nu}$ is

$$h_{\mu\nu}(x)=-\frac{1}{M_g^2}\int d^4y G_{\mu\nu}^{\alpha\beta}(x-y)\,T_{\alpha\beta}(y), \tag{4.11}$$

where

$$G_{\mu\nu}^{\alpha\beta}(x-y)=\left(\delta^{\alpha}_{(\mu}\delta^{\beta}_{\nu)}-\frac{1}{2}\eta_{\mu\nu}\eta^{\alpha\beta}\right)G(x-y) \tag{4.12}$$

is the graviton propagator, and

$$G(x-y)=\frac{1}{\Box}\delta^{(4)}(x-y)=\int\frac{d^4p}{(2\pi)^4}\frac{1}{-p^2}e^{ip(x-y)} \tag{4.13}$$

is the standard scalar propagator. Of the ten components of $h_{\mu\nu}$, the Lorenz gauge (4.9) fixes four. In vacuum, the equations of motion are simply $\Box h_{\mu\nu}=0$ in this gauge, and there is still a residual gauge freedom left, since one can perform transformations for which $\Box\xi^{\mu}=0$. This can be used to remove another four components. A standard choice is the transverse-traceless gauge, in which

$$h_{0\mu}=0,\quad h=0,\quad \partial^{\mu}h_{\mu\nu}=0. \tag{4.14}$$

In this gauge there are only two propagating degrees of freedom left, which upon quantization exactly correspond to the two degrees of freedom of a massless spin-2 particle, with helicities ± 2 (the helicity-2 nature of the interaction can be inferred by studying the transformation properties of the transverse-traceless part of $h_{\mu\nu}$ under spatial rotations).

4.2 Making the graviton massive

The quadratic Lagrangian (4.3) describes a massless spin-2 field. In 1939 Fierz and Pauli added a mass term in order to describe massive spin-2 fields [25, 26]. The mass term is

$$\mathcal{L}_{FP} = -\frac{m^2}{8}\left(h_{\mu\nu}h^{\mu\nu} - h^2\right). \tag{4.15}$$

The equations of motion are now given by

$$\mathcal{E}^{\alpha\beta}_{\mu\nu}h_{\alpha\beta} - m^2\left(h_{\mu\nu} - \eta_{\mu\nu}h\right) = -\frac{1}{M_g^2}T_{\mu\nu}. \tag{4.16}$$

Comparing with the equation (4.5) for the massless field, we have two important differences. First of all, the equations of motion are no longer invariant under the gauge transformation

$$h_{\alpha\beta} \to h_{\alpha\beta} + \partial_\alpha\xi_\beta + \partial_\beta\xi_\alpha. \tag{4.17}$$

The massive spin-2 Lagrangian thus breaks the reparametrization invariance that the massless spin-2 Lagrangian contained. Secondly, source conservation is no longer automatically implied by the equations of motion, but has to be postulated. Indeed, when acting with ∂^μ on both sides of (4.16) one gets

$$\partial^\mu h_{\mu\nu} - \partial_\nu h = \frac{1}{m^2 M_g^2}\partial^\mu T_{\mu\nu}. \tag{4.18}$$

If one assumes that the source is conserved, then $\partial^\mu h_{\mu\nu} - \partial_\nu h = 0$, and plugging this back into the equations of motion gives a constraint on h:

$$h = -\frac{T}{3M_g^2 m^2}. \tag{4.19}$$

Using this once again in the equations of motion, they can be written in the following form:

$$\left(\Box - m^2\right)h_{\mu\nu} = -\frac{1}{M_g^2}\left(T_{\mu\nu} - \frac{1}{3}\eta_{\mu\nu}T + \frac{1}{3m^2}\partial_\mu\partial_\nu T\right). \tag{4.20}$$

In vacuum, we are thus left with the following equations:

$$h = 0, \tag{4.21}$$

$$\partial^\mu h_{\mu\nu} = 0, \tag{4.22}$$

$$\left(\Box - m^2\right)h_{\mu\nu} = 0. \tag{4.23}$$

The first two equations removes five components of $h_{\mu\nu}$, whereas the third equation give the equation of motion for the remaining five components.

These components are the five propagating degrees of freedom for the massive spin-2 field, corresponding to helicities ± 2, ± 1 and 0, as compared to the two degrees of the massless field.

The propagator for the massive spin-2 field can be attained by inverting (4.20):

$$h_{\mu\nu} = -\frac{1}{M_g^2}\int d^4y G_{\mu\nu}^{\alpha\beta}\left(x-y;m^2\right)T_{\alpha\beta}. \tag{4.24}$$

Here

$$G_{\mu\nu}^{\alpha\beta}\left(x-y;m^2\right) = G\left(x-y;m^2\right)\left(\delta_{(\mu}^{\alpha}\delta_{\nu)}^{\beta} - \frac{1}{3}\eta_{\mu\nu}\eta^{\alpha\beta}\right) \tag{4.25}$$

is the massive graviton propagator, and

$$\begin{aligned} G\left(x-y;m^2\right) &= \frac{1}{\Box - m^2}\delta^{(4)}\left(x-y\right) \\ &= \int \frac{d^4p}{(2\pi)^4}\frac{1}{-p^2-m^2}e^{ip(x-y)} \end{aligned} \tag{4.26}$$

is the propagator for a massive scalar. We see that in the limit $m^2 \to 0$

$$\lim_{m^2\to 0} G\left(x-y;m^2\right) = G\left(x-y\right) \tag{4.27}$$

but

$$\lim_{m^2\to 0} G_{\mu\nu}^{\alpha\beta}\left(x-y;m^2\right) \neq G_{\mu\nu}^{\alpha\beta}\left(x-y\right) \tag{4.28}$$

due to the factor of 1/3 in (4.25). There is thus, in the linearized framework, no smooth limit to general relativity as the mass of the graviton goes to zero. This is known as the van Dam-Veltman-Zakharov (vDVZ) discontinuity, named after its original discoverers [27, 28] (see also [29]).

The culprit of the vDVZ-discontinuity can be traced to the helicity-0 mode of the massive spin-2 field. This is most easily seen through a Stückelberg analysis, where the broken gauge invariance in the massive theory is restored by introducing auxiliary fields that carry the right transformation properties. This allows for a smooth $m^2 \to 0$ limit, and by studying the residual fields and their interactions one can isolate the physical difference between the massive and massless theories. For example, doing a Stückelberg analysis of the Proca Lagrangian (i.e. the Lagrangian for a massive photon), one sees that the helicity-0 mode, which carries the new longitudinal polarization, effectively decouples in the $m^2 \to 0$ limit, and one is therefore left with the original massless Maxwell Lagrangian together with a non-interacting scalar field. Doing a similar analysis for the Fierz-Pauli action one sees that there is a residual coupling left between the helicity-0 mode and the trace of the stress-energy tensor. The $m^2 \to 0$ limit is thus an interacting tensor-scalar theory.

Since the trace of the stress-energy tensor vanishes for light but not for matter, the bending of light around a massive object will be different in the tensor-scalar theory as compared to general relativity. More precisely, one can show that there is a 25% discrepancy between the massive and massless theory, irrespective of how small the graviton mass is. Measuring such a discrepancy lies well within the capabilities of current observational techniques, so at first sight, one would think that the theory of massive gravity is ruled out. It turns out, however, that things are a bit more subtle. Whereas it is correct that such a discrepancy exists, there also exists a radius wherein non-linear effects have to be taken into account. This radius was identified by Vainshtein in 1972, and is therefore known as the Vainshtein radius [30]. Solar system observations lie well inside such a radius. Vainshtein postulated that there could exist a non-linear Vainshtein mechanism that produces a smooth limit to general relativity in the limit of vanishing graviton mass. The observational consequences of the vDVZ-discontinuity for bending of light outside the Vainshtein radius is the subject of Paper III. This is further discussed in chapter 10.

4.3 The graviton vs. gravity

We saw that the field equations (4.5) implied conservation of energy-momentum $\partial_\mu T^{\mu\nu} = 0$. This can not be the complete story, however, since the interaction between gravity and matter can remove energy from the matter sources and radiate it away through gravitational waves. Also, the full equations of motion for $T^{\mu\nu}$ must include the field $h_{\mu\nu}$, which is not the case of $\partial_\mu T^{\mu\nu} = 0$. One therefore need to include higher order terms in $h_{\mu\nu}$ in the action in order to arrive at a consistent theory of gravity. This means, in particular, that self-interactions of the gravitational field are introduced. This iterative procedure leads uniquely to the Einstein-Hilbert action (4.1), up to boundary terms [31, 32].

When trying to add non-linear terms to the Fierz-Pauli action one has to ensure, at each step, that the constraints that ensured five propagating degrees of freedom is not lost. A problematic sixth degree of freedom with a negative kinetic term is usually referred to as a ghost. Such a degree of freedom signals a pathology of the theory, since other fields can acquire arbitrarily large energies from the field with negative kinetic energy. It also leads to negative probabilities upon quantization, which is unphysical. That such a field would show up for generic higher-order interactions was shown by Boulware and Deser in [33]. The pathological sixth degree of freedom is therefore referred to as the Boulware-Deser ghost.

The Fierz-Pauli action was written down using flat space as the background metric. For the fully non-linear theory, one has to introduce a new rank-2 field $f_{\mu\nu}$ in order to create non-derivative terms with $g_{\mu\nu}$. These terms will form the potential, and the terms have to be carefully constructed in order to avoid the Boulware-Deser ghost. If the rank-2 field $f_{\mu\nu}$ transforms as a tensor, the general covariance of the theory is restored. It is also possible to give dynamics to $f_{\mu\nu}$.

The non-linear completion of the Fierz-Pauli action was studied in a series of papers [34, 35, 36, 37]. The first correct potential in a certain limiting case, and using a flat reference metric, was written down by de Rham, Gabadadze and Tolley (dRGT) in [38]. The full non-linear action, together with a proof that it was ghost-free, was written down by Hassan and Rosen in [1, 39]. This was immediately generalized to arbitrary $f_{\mu\nu}$ in [40] and to a bimetric framework, where $f_{\mu\nu}$ was also given dynamics, in [2, 41]. The possibility of having several interacting, dynamical rank-2 fields was shown in [42]. The theory with a non-dynamical reference metric $f_{\mu\nu}$ is usually referred to as the dRGT theory. For dynamical $f_{\mu\nu}$, the theory is usually called Hassan-Rosen theory (and also massive bigravity, or bimetric massive gravity). The potential term constructed out of $g_{\mu\nu}$ and $f_{\mu\nu}$ is the same in both cases.

4.4 The Hassan-Rosen action

The Hassan-Rosen action, without matter couplings, is

$$S = \int d^4x \left[-\frac{M_g^2}{2}\sqrt{-g}R_g - \frac{M_f^2}{2}\sqrt{-f}R_f + m^4\sqrt{-g}\sum_{n=0}^{4}\beta_n e_n\left(\sqrt{g^{-1}f}\right)\right]. \tag{4.29}$$

Here

$$e_0(X) = 1, \qquad e_1(X) = [X], \qquad e_2(X) = \frac{1}{2}\left([X]^2 - \left[X^2\right]\right),$$
$$e_3(X) = \frac{1}{6}\left([X]^3 - 3\left[X^2\right][X] + 2\left[X^3\right]\right), \qquad e_4(X) = \det X, \tag{4.30}$$

where $[X] = \mathrm{Tr}X$. The action contains two Einstein-Hilbert terms for g and f, and an interaction potential which depends on the square root of $g^{-1}f$, defined such that

$$\sqrt{g^{-1}f}\sqrt{g^{-1}f} = g^{-1}f. \tag{4.31}$$

Five independent parameters β_n characterize the potential. A priori they can have any value, and ultimately they have to be constrained by comparing the predictions of the Hassan-Rosen theory with observations.

In the next chapter we describe possible matter couplings that one could add to this action. Depending on the type of coupling the equations of motion will obviously look different. In the specific phenomenological applications, where we use different type of couplings, we will give the equations of motion for the specific metric ansätze under consideration. Here we will instead just state the general equations of motion in vacuum:

$$R_{\mu\nu}(g) - \frac{1}{2}g_{\mu\nu}R(g) + \frac{m^4}{M_g^2}\sum_{n=0}^{3}(-1)^n\beta_n g_{\mu\lambda}Y^{\lambda}_{(n)\nu}\left(\sqrt{g^{-1}f}\right) = 0, \tag{4.32}$$

$$R_{\mu\nu}(f) - \frac{1}{2}f_{\mu\nu}R(f) + \frac{m^4}{M_f^2}\sum_{n=0}^{3}(-1)^n\beta_{4-n} f_{\mu\lambda}Y^{\lambda}_{(n)\nu}\left(\sqrt{f^{-1}g}\right) = 0. \tag{4.33}$$

Here the Y matrices are given by

$$Y_{(0)}(X) = I, \qquad Y_{(1)}(X) = X - I\cdot e_1(X),$$
$$Y_{(2)}(X) = X^2 - X\cdot e_1(X) + I\cdot e_2(X),$$
$$Y_{(3)}(X) = X^3 - X^2\cdot e_1(X) + X\cdot e_2(X) - I\cdot e_3(X), \tag{4.34}$$

where I is the identity matrix. Taking the divergence of these two equations and using the Bianchi identity given in (3.4) leads to the following two constraints:

$$\frac{m^4}{M_g^2}\nabla_g^\mu \sum_{n=0}^{3} (-1)^n \beta_n g_{\mu\lambda} Y^\lambda_{(n)\nu}\left(\sqrt{g^{-1}f}\right) = 0, \tag{4.35}$$

$$\frac{m^4}{M_f^2}\nabla_f^\mu \sum_{n=0}^{3} (-1)^n \beta_{4-n} f_{\mu\lambda} Y^\lambda_{(n)\nu}\left(\sqrt{f^{-1}g}\right) = 0. \tag{4.36}$$

These two equations are not independent of one another, which is a consequence of the reparametrization invariance of the action. Furthermore, they do not provide any new information as compared to the original equations of motion (4.32-4.33), but they are still useful since they can present the content of the equations of motion in a more manageable way.

The equations of motion contain in total 7 parameters: five β-parameters and 2 coupling constants. It is possible to do various rescalings of the two metrics and these parameters in order to reduce the number of free parameters. This is described in chapter 6.

Chapter 5

Adding matter

When the muon was discovered in 1936 Nobel laureate I.I. Rabi exclaimed "who ordered that?", since it was believed that the proton, neutron and electron would be enough to describe the sub-atomic world. In a similar way we can respond "who ordered a second metric?" when we see that the construction of a theory of a massive graviton immediately introduces a second rank-2 tensor.

Given that two dynamical rank-2 tensors fields g and f are present in massive bigravity, we are confronted with the question of how to couple matter to these two fields. In general relativity, *all* matter couples to gravity in the same way (which is usually called the weak equivalence principle) through a term $\sqrt{-g}\mathcal{L}_{\mathrm{m}}$ in the action. Since the rods and clocks that are used to measure spacetime distances couple to g, one usually refers to g as the metric. But now we have two fields, and coupling matter to both of them would create an ambiguity; which one of g and f should we promote to a metric that determines the spacetime structure?[1]

As if these conceptual difference were not enough, there are also theoretical consistency issues. The interaction potential between g and f was carefully crafted to avoid the so-called ghost problem: a degree of freedom with negative Hamiltonian that upon quantization gives rise to negative probabilities. The matter coupling could ruin the constraint that is necessary for the theory to be well-behaved.

Different couplings have been proposed in the literature, and in the next section we will state the most common ones. Not all of them are ghost-free, but this does not necessarily mean that they are excluded a

[1]Even though it is only the tensor field that couples to matter that should properly be called a metric, we will use the somewhat incorrect language of referring to both g and f as metrics.

priori, since they could potentially still be considered to be effective field theories below the energy threshold where the ghost appears. In section 5.2 we will discuss how the couplings affect the different equivalence principles closely related to the formulation of general relativity.

5.1 Matter couplings

In this section we go through some common couplings to matter. The matter Lagrangians will depend on each of the two metrics or some combination thereof. The matter fields are denoted by Φ.

Singly coupling The most straightforward coupling is to just couple either g or f to matter, through a term $\sqrt{-g}\mathcal{L}_{\mathrm{m}}\left(g_{\mu\nu},\Phi\right)$ or $\sqrt{-f}\mathcal{L}_{\mathrm{m}}\left(f_{\mu\nu},\Phi\right)$. This is referred to as singly coupling (the introduction and ghost-free status of this couplings was performed in the first formulation of the theory, [2]). This coupling breaks the symmetry between g and f (see section 6.1, however, for a possible resolution), and one is left wondering why it was, say, g instead of f that should couple to matter. The advantage with coupling to only one of the rank-2 tensors is that this tensor is then immediately promoted to *the* metric. This coupling is used in Paper I, Paper II, Paper III and Paper VI.

Doubly coupling If one wants to couple both g and f to matter, there are basically two options. The first is to couple g and f to different matter sectors:

$$S_{\mathrm{m}} = \int d^4x\sqrt{-g}\mathcal{L}_1\left(g_{\mu\nu},\Phi_g\right) + \int d^4x\sqrt{-f}\mathcal{L}_2\left(f_{\mu\nu},\Phi_f\right). \tag{5.1}$$

Here $\mathcal{L}_1$ and $\mathcal{L}_2$ need not be of the same functional form. This type of coupling will be ghost-free [2], and its phenomenology was explored in [44]. Another option is to couple both g and f to the same matter content:

$$S_{\mathrm{m}} = \int d^4x\sqrt{-g}\mathcal{L}_{\mathrm{m}}\left(g_{\mu\nu},\Phi\right) + \int d^4x\sqrt{-f}\mathcal{L}_{\mathrm{m}}\left(f_{\mu\nu},\Phi\right). \tag{5.2}$$

This type of coupling is conceptually problematic, since neither g nor f can be considered to be the metric. Instead, it turns out that one can form a combination of them that will be a metric, but this metric will depend on not only the position but also the velocity of the observer [45]. Furthermore, this coupling is not ghost-free [46, 47].

Effective coupling Another coupling, proposed in [46], is to introduce an effective metric, formed out of g and f in the following way:

$$g^{\text{eff}}_{\mu\nu} \equiv \alpha^2 g_{\mu\nu} + 2\alpha\beta g_{\mu\alpha} X^{\alpha}_{\nu} + \beta^2 f_{\mu\nu}. \tag{5.3}$$

We remind the reader that $X = \sqrt{g^{-1}f}$. One can show that gX is a symmetric tensor, and under the interchange $g \leftrightarrow f$ and $\alpha \leftrightarrow \beta$, the effective metric remains invariant (these dualities are described in more detail in section 6.1). The matter coupling is now

$$S_{\text{m}} = \int d^4x \sqrt{-g^{\text{eff}}} \mathcal{L}_{\text{m}} \left(g^{\text{eff}}_{\mu\nu}, \Phi \right). \tag{5.4}$$

The advantage with the effective metric is that matter couples symmetrically do g and f, and there is thus only *one* metric. This coupling does have a ghost, as shown in [48] (and also discussed in [49, 50, 51, 52, 53]), but the coupling could potentially still be used effectively below the energy threshold of the ghost. This coupling is used in Paper IV and Paper V.

Massless coupling The spin-2 content of massive bigravity can be analyzed through a decomposition of the metrics when linearizing it around a Minkowski background. One then finds, as expected, that the theory contains two massless helicity-2 modes and two massive helicity-2, two massive helicity-1 and one massive helicity-0 mode. The theory thus contains, in total, seven propagating degrees of freedom. Given this structure it is natural to ask whether one can reformulate the theory so that it contains one massless and one massive part, and if it is, then, possible to only couple the massless part to matter. This was investigated in [43]. Linearizing g and f as

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \frac{1}{M_g}\delta g_{\mu\nu}, \qquad f_{\mu\nu} = c^2 \bar{g}_{\mu\nu} + \frac{c}{M_f}\delta f_{\mu\nu}, \tag{5.5}$$

where c is a proportionality constant between the two background metrics, the massless and massive modes are given by

$$\delta G_{\mu\nu} = \frac{1}{\sqrt{c^2 M_f^2/M_g^2 + 1}} \left(\delta g_{\mu\nu} + c\frac{M_f}{M_g}\delta f_{\mu\nu} \right), \tag{5.6}$$

$$\delta M_{\mu\nu} = \frac{1}{\sqrt{c^2 M_f^2/M_g^2 + 1}} \left(\delta f_{\mu\nu} - c\frac{M_f}{M_g}\delta g_{\mu\nu} \right). \tag{5.7}$$

$\delta G_{\mu\nu}$ will satisfy the equations of motion for a massless field, and $\delta M_{\mu\nu}$ the equations of motion for a massive field. A possible non-linear generalization of the massless field $\delta G_{\mu\nu}$ is

$$G_{\mu\nu} = g_{\mu\nu} + \frac{M_f^2}{M_g^2} f_{\mu\nu}. \tag{5.8}$$

There exists many different non-linear extensions of the massive mode $\delta M_{\mu\nu}$. One such example is

$$M_{\mu\nu} = g_{\mu\alpha} X^{\alpha}_{\nu} - c g_{\mu\nu}. \tag{5.9}$$

One could then couple matter to only the massless part of the theory, through

$$S_{\mathrm{m}} = \int d^4x \sqrt{-G} \mathcal{L}_{\mathrm{m}} \left(G_{\mu\nu}, \Phi \right). \tag{5.10}$$

In [43] it was shown that this type of coupling reintroduces the ghost. From the analysis of the massive and massless decomposition, one sees that if one couples to only g and f a given source will excite both the massive and massless states.

5.2 Equivalence principles

Historically, the equivalence principles have played an important role in the formulation of both Newtonian gravity and general relativity. Galileo first pointed out that all bodies fall at the same rate under the influence of gravity. This fact was used as an assumption in Newton's formulation of gravity, and explained by Einstein as the effect of geodesic motion in a curved spacetime.

The modern formulation of the equivalence principles are as follows (see [14, 54] for a review):

1. All bodies with the same initial position and velocity follow the same trajectories.

2. Local non-gravitational experiments are velocity independent.

3. Local non-gravitational experiments are independent of position and time.

These principles give rise to a metric formulation of gravity, where trajectories are given by geodesics, and local non-gravitational experiments are described in the framework of special relativity. Point one is usually called the weak equivalence principle. Point two and three assumes that the constant of nature, such as the electromagnetic coupling strength, mixing angles etc, are not space and time dependent. When point two and three are also applied to gravitational experiments, they are referred to as the strong equivalence principle. A time-varying Newton's constant is an example of a breaking of the strong equivalence principle.

Since the Hassan-Rosen theory contains two tensor fields, and allows for several types of matter couplings, one naturally wonders what happens with the equivalence principles. When coupling all matter to the same tensor field, irrespective of this field is g or f or a combination thereof, the weak equivalence principle will hold, since all matter will travel on geodesics defined with respect to that tensor field.[2] Non-gravitational experiments will still be independent of velocity, position and time, since one can always transform the field that matter couples to into a local Lorentzian frame. Gravitational experiments can, however, look rather different. Coupling the same matter to g and f in the same way is equivalent to coupling matter to a so-called Finsler metric, which not only depends on the observers spacetime position but also velocity. This would break the strong equivalence principle. Also, Newton's constant can, effectively, look different at different lengthscales, which would also break the strong equivalence principle. We thus conclude that it is possible to break the strong equivalence principle in the Hassan-Rosen theory, while keeping the weak equivalence principle intact.

[2] A caveat occurs if one only couples matter to one of the metrics, say g, and considers f to be some kind of matter field. Since f then couples differently to the metric as compared to other matter fields, this would imply a violation of the weak equivalence principle. This is, however, somewhat of a terminological question, since one could equally well consider f to be a part of the gravity sector.

Chapter 6

Dualities and symmetries

6.1 Mapping the theory into itself

When using the effective metric described in section 5.1, the action is

$$\begin{aligned} S = &-\frac{M_g^2}{2}\int d^4x\sqrt{-\det g}R\left(g\right) - \frac{M_f^2}{2}\int d^4x\sqrt{-\det f}R\left(f\right) \\ &+ m^4\int d^4x\sqrt{-\det g}\mathcal{V}\left(\sqrt{g^{-1}f};\beta_n\right) \\ &+ \int d^4x\sqrt{-\det g_{\text{eff}}}\mathcal{L}_m\left(g_{\text{eff}},\Phi\right). \end{aligned} \tag{6.1}$$

Here the effective metric g_{eff} is defined in (5.3). The singly couplings can be attained by putting either α or β to zero. Not all of the parameters in the action are independent: It is possible to rescale the parameters $M_g, M_f, m^2\beta_i, \alpha$ and β together with $g_{\mu\nu}$ and $f_{\mu\nu}$ to effectively remove certain parameters. The interaction potential satisfies

$$\sqrt{-\det g}\mathcal{V}\left(\sqrt{g^{-1}f};\beta_n\right) = \sqrt{-\det f}\mathcal{V}\left(\sqrt{f^{-1}g};\beta_{4-n}\right), \tag{6.2}$$

and using this property the action will be invariant as

$$g_{\mu\nu}\leftrightarrow f_{\mu\nu}, \qquad M_g\leftrightarrow M_f, \qquad \alpha\leftrightarrow\beta, \qquad \beta_n\to\beta_{4-n}. \tag{6.3}$$

Under these scalings the effective metric will also remain invariant. The meaning of these transformations is that there exists a *duality* in the action, which maps one set of solutions of the theory, for a given set of parameters, to another solution of the theory, with another set of parameters. Besides the theoretical interest, this is also of practical interest when performing parameter scans, since only a part of the parameter space has to be investigated.

It turns out that not all of the parameters M_g, M_f, α, β are physically meaningful. Due to the scaling properties (6.3) one can either use an effective coupling constant M_{eff} together with the ratio β/α as the physically independent parameters (and since M_{eff} can be absorbed in the matter definition, only β/α is important), *or* rescale away α and β and use M_g and M_f (and, once again, since one of them can be absorbed in the matter definition, only their ratio matters). Both of these approaches are physically equivalent, and one choice can be mapped to another (this is described in detail in appendix A of Paper IV). The advantage of the first choice is that there is only one coupling constant between gravity and matter, and the single coupling limit is made explicit by either sending the ratio β/α to zero or infinity. The advantage with the other choice is that there is no ambiguity when defining the effective metric (since it does not contain any free parameters), and M_g and M_f have a straightforward interpretation of how strong the respective metrics couple to matter. The single coupling limit is not as transparent any more, however. Also, β/α only appears in the matter sector, whereas M_f/M_g appears in both the matter sector and in the interaction terms.

Notice that due to the mapping (6.3) one can consider the case of coupling only g to matter as equivalent to coupling only f to matter, but using another set of β_i parameters.

6.2 Special parameters

When coupling to matter through the effective metric, the theory has in total six parameters: the five β_i:s together with the ratio β/α or M_f/M_g. There is a famous saying by von Neumann: "With four parameters I can fit an elephant, and with five I can make it wiggle its trunk." What we can do with the elephant using six parameters is beyond the scope of this thesis, but clearly one has to consider if there are special parameter cases that are worth special attention. We use the rescaling described in the previous section, where β/α is a free parameter rather than M_f/M_g.

Vacuum energy The meaning of the different β_i parameters will depend on the type of coupling. To beginning with, let us discuss which of the parameters corresponds to a cosmological constant. In general relativity, the cosmological constant appears in the action as a numerical factor in front of the volume determinant $\sqrt{-g}$. Since the matter Lagrangian couples to gravity through the same term, the vacuum energy associated with matter will appear from a gravitational point of view as a cosmological constant. The vacuum energy and cosmological constant are thus indistinguishable in general relativity.

In massive bigravity, the vacuum energy is still associated with zero-point matter fluctuations, and thus contribute a constant piece to the matter Lagrangian. But since matter can couple in different ways to g and f, the vacuum energy contribution can not straightforwardly be identified with a cosmological constant contribution. Furthermore, the meaning of the cosmological constant is ambiguous. If, as in general relativity, it is the term that comes in front of the volume determinant that is identified with the cosmological constant, then it is β_0 and β_4 that corresponds to a cosmological constant for g and f, respectively. If, however, one looks at the effective contribution that one has to cancel in order to have flat space backgrounds, all of the β_i parameters will contribute.

Going back to the case of the vacuum energy, it will renormalize the β_0 or β_4 term when coupling matter only to g or f, respectively. When $\alpha, \beta \neq 0$ in the effective metric, the vacuum energy will renormalize all β_i terms, which can be considered a drawback when using the effective coupling. This is described in more detail in section 7.4.

Partially masslessness The parameter choice

$$\beta_1 = \beta_3 = 0, \qquad \beta_0 = 3\beta_2 = \beta_4 \tag{6.4}$$

is particularly intriguing since it gives rise to a new gauge symmetry at the linear level, which is conjectured to hold non-linearly [55, 56, 57, 58, 59, 60, 61]. This symmetry would remove the helicity-0 mode of the massive spectrum. This mode is responsible for the problematic vDVZ-discontinuity and fifth forces, and the theory would, in this sense, be more "well-behaved". This gauge symmetry have been proven at the linear level around proportional backgrounds, but the complete non-linear symmetry has not been found. The imposition of these parameter values will have different effects depending on the type of coupling, as described further in section 7.4.

Maximally symmetric model For the parameter values

$$\beta_0 = \beta_4, \qquad \beta_1 = \beta_3, \qquad \alpha = \beta \tag{6.5}$$

the theory is maximally symmetric in the sense that a solution will be mapped to itself under the transformation $g_{\mu\nu} \leftrightarrow f_{\mu\nu}$, $\beta_n \to \beta_{4-n}$, $\alpha \leftrightarrow \beta$ [58]. There is thus no distinction between the two metrics for these parameter values. In the case of singly coupling, the parameter choice $\beta_0 = \beta_4$ and $\beta_1 = \beta_3$ means that it does not matter if one couples g or f; the equations of motion are the same.

Part IV

Cosmic phenomenology

Chapter 7

Expansion histories

In this chapter we describe tests of massive gravity using cosmic expansion histories. We start by discussing the meaning of cosmological expansion and observational tests thereof in section 7.1. In section 7.2 we look at the setup with a non-dynamical reference metric. After showing that it generically is hard to get viable cosmologies in this setup, we turn to the case of massive bigravity with a single matter coupling in section 7.3, and a symmetric matter coupling in section 7.4.

7.1 Probing the universe to zeroth order

On length scales larger than roughly 100 Mpc the universe is statistically homogenous and isotropic. This means that it looks—on average—the same at each point and in each direction. It is not, however, true that the universe looks the same at each point in time. The expansion of the universe dilutes its matter content, and makes it cool down. The different physical processes that occur at different temperatures makes the universe highly asymmetric in time. The expansion of the universe is encoded in the scale factor $a(t)$; the fractional change in distances between time t_1 and t_2 is given by $a(t_2)/a(t_1)$.

The line element for a homogeneous and isotropic universe is

$$ds^2 = -N^2(t)dt^2 + a^2(t)\delta_{ij}dx^i dx^j. \tag{7.1}$$

Here, and in the following, we assume a flat spatial geometry, which is in good agreement with observations [62]. The lapse $N(t)$ can be rescaled by a time reparametrization; common choices are cosmic time, for which $N(t) = 1$, or conformal time, for which $N(t) = a(t)$. The line element 7.1

is known as the Friedmann-Lemaître-Robertson-Walker (FLRW) metric. The fractional expansion of the universe, in time units of Ndt, is given by the Hubble function[1]

$$H \equiv \frac{\dot{a}}{aN}. \tag{7.2}$$

The Hubble function H, and thus the expansion history of the universe, will be affected by the energy content of the universe. The exact form of this relationship depends on the nature of gravity, i.e. the relationship between energy content and spacetime curvature. Independent measurements of the expansion history and the energy content of the universe are thus needed to properly understand the effect of gravitational interactions on cosmological scales. Since not all of the energy content of the universe can be observed directly, the approach used in practice is a bit different: by reconstructing the expansion history one can infer, for a given theory of gravity, what the energy content has to be (which in the case of general relativity leads to the conclusion that dark matter and dark energy has to constitute a large part of the energy content). If there exists a viable background solution given the free parameters of the theory, other observations—such as structure formation, lensing, planetary orbits etc—then have to be supplemented to see if the inferred energy content is plausible. If not, the theory of gravity under scrutiny can be ruled out.

In order to determine the expansion history two distance measures are most commonly used: luminosity and angular distances. The observed flux $\mathcal{F}$ of photons and the luminosity L of an object are related by

$$\mathcal{F} = \frac{L}{4\pi d_L^2}. \tag{7.3}$$

This relationship defines the luminosity distance d_L, which thus relates the intrinsic brightness of an object to its observed brightness. It is given by

$$d_L(z) = (1+z)\int_0^z \frac{dz'}{H(z')} \tag{7.4}$$

[1]This definition means that $H = \dot{a}/a$ in cosmic time and $H = \dot{a}/a^2$ in conformal time. Sometimes H is defined as $\dot{a}/a$ also in conformal time, which means that H is then *not* a constant in a de Sitter spacetime. We therefore use the definition (7.2).

in a flat universe. Observations of the light from objects with known luminosity can thus be used to constraint the expansion history of the universe. The angular distance is defined as

$$d_A \equiv \frac{D}{\delta\theta}, \tag{7.5}$$

where D is the proper size of an object and $\delta\theta$ the observed angle in the sky. The angular distance is related to the luminosity distance through

$$d_A = \frac{d_L}{1+z}. \tag{7.6}$$

Here z is the redshift, defined through

$$z \equiv a(t)^{-1} - 1, \tag{7.7}$$

if we normalize $a(t)$ such that it is unity today.

The definition of the FLRW metric and its relationship to the luminosity and angular distances does not depend on any other properties of gravity except that it can be described by a metric. In order to use the inferred expansion history to establish the energy content of the different physical constituents of the universe—i.e. baryonic matter, light, neutrinos etc—one needs an equation, usually called the Friedmann equation, that relates H and the energy density ρ. In general relativity, for example, the Hubble function has a remarkably simple relationship to the energy content of the universe:

$$H^2 = \frac{8\pi G}{3}\rho. \tag{7.8}$$

Here ρ contains all energy components, including a vacuum energy density. The different energy components will evolve in time according to the continuity equation:

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0. \tag{7.9}$$

The pressure p will only become gravitationally important when the average kinetic energy is comparable to the rest energy of the species under consideration. For baryonic matter the average velocities are too small to produce any gravitational effect, and the pressure is therefore effectively zero. For neutrinos, on the other hand, the average velocities are close to the speed of light, and the pressure can therefore have a gravitational impact. The relationship between the pressure and the energy density is specified through an equation of state $p = p(\rho)$. This is usually parametrized as

$$p = w\rho. \tag{7.10}$$

For baryonic matter we have $w = 0$, for light $w = 1/3$ and for a constant vacuum energy density $w = -1$.

The Friedmann equation (7.8) can be parametrized as

$$H^2 = H_0^2 \left[\Omega_{\rm m}(1+z)^3 + \Omega_\gamma(1+z)^4 + \Omega_{\rm de} f(z)\right], \tag{7.11}$$

where the subscripts "m", "γ" and "de" stand for matter (both baryonic and dark), radiation and dark energy, respectively. H_0 is the value of the Hubble function today, i.e. at $z = 0$, and

$$\Omega_i = \frac{8\pi G\rho_{i,0}}{3H_0^2}, \tag{7.12}$$

where $\rho_{i,0}$ is the energy density of the i:th component today. Finally

$$f(z) = \exp\left(3\int_0^z \frac{1+w(z')}{1+z'}dz'\right) \tag{7.13}$$

gives the time evolution of a possible dark energy component. Considering a flat universe, we must have

$$\Omega_{\rm m} + \Omega_\gamma + \Omega_{\rm de} = 1. \tag{7.14}$$

$\Omega_{\rm m}$, Ω_γ and $\Omega_{\rm de}$ thus gives the fractional contributions to the energy density today.

To discern the background cosmology and put constraints on the energy components of the universe, one primarily looks at Type Ia supernovae (SNe Ia), the cosmic microwave background (CMB) and baryonic acoustic oscillations (BAO).

Type Ia supernovae In binary star systems consisting of a white dwarf that accretes gas from an accompanying star, a thermonuclear explosion takes places when the white dwarf mass approaches the Chandrasekhar mass $M \sim 1.4M_\odot$ (the exact nature of this explosion and the binary system that preceded it are unknown). The common physical origin of the SNe Ia makes the peak luminosities of the events similar. Also, the scatter can be reduced even further using empirical correlations between the peak luminosity, colour and the time evolution of the luminosity. This standardization of the observed flux allows us to use (7.3) to infer the luminosity distance and, since the luminosity distance depends on the expansion history, observations of SNe Ia at different redshifts to probe $H(z)$.

Current SNe Ia catalogues used for cosmological applications consists of several hundred objects. Measurement uncertainties are mostly limited by systematics. The highest redshifts used in reconstructions of the expansion history are $z \sim 1$, which corresponds to an age of the universe of roughly six billion years.

Cosmic microwave background As photons decoupled from the primordial plasma some 13.4 billion years ago, they started to travel almost freely in the universe, with the structure existing 300 000 years after the big bang imprinted on their current distribution. The CMB is therefore one of the primary probes in cosmology. It will be discussed more in depth in section 8.3, and here we just mention how the CMB can be used to probe the expansion history.

The region of space from which the observed decoupled photons have travelled until they were observed is called the last scattering surface. Prior to decoupling, the plasma experienced a series of oscillations; on the one hand the plasma collapsed under its own gravity, on the other hand it expanded due to the background expansion and photon pressure. These oscillations were imprinted on both the photon and baryonic content at the time of decoupling. The largest oscillation length at the time of decoupling was roughly equal to the interaction length scale at that time, i.e. how far perturbations in the plasma could have travelled since it was first formed. This length scale is referred to as the sound horizon scale at decoupling, and forms a "standard ruler" that can be used to probe the evolution of distances from the time of decoupling until today.

Comparing the size of the sound horizon at decoupling and the angle this size subtends in the CMB today, one can, from (7.5), infer an angular distance to the last scattering surface. Since the angular distance depends on the expansion history, observation of the angular size of the sound horizon in the CMB thus allows for constraints on $H(z)$ and the energy content of the universe.

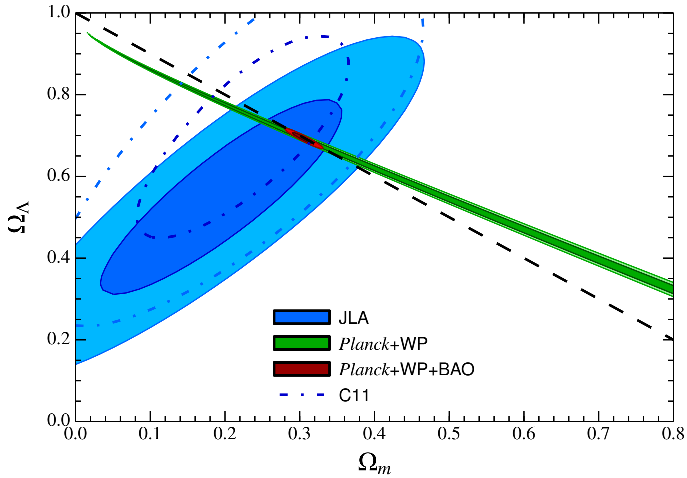

Figure 7.1. 68% and 95% confidence contours for $\Omega_{\rm m}$ and Ω_Λ. The equation of state of dark energy is assumed to be -1, and spatial curvature is allowed for. JLA is the SNe Ia data set, compiled in [64], C11 the SNe Ia data set compiled in [65], Planck+WP means Planck temperature and WMAP polarization measurements. The black dashed line signifies a flat universe. One sees that CMB and BAO measurements alone strongly favours a flat universe. Plot taken from [64].

Baryon acoustic oscillations The baryonic distribution at the time of decoupling is imprinted on the large scale galaxy distribution today. In particular, the standard ruler that the sound horizon size at decoupling implies will also show up in the baryonic structure at later times. Measuring this length scale at different redshifts is thus a probe of the expansion history.

A galaxy catalogue will contain both angular positions and redshifts, which can be combined into a "distance average" through the following quantity:

$$D_V(z) = \left[d_A(z)^2 \frac{z}{H(z)}\right]^{1/3}. \tag{7.15}$$

The factor of $1/H(z)$ comes from the fact that locally $dr = dz/H(z)$. The extra factor of z is purely conventional. Baryon acoustic oscillations allows for measurements of the ratio between the sound horizon close to decoupling and $D_V(z)$. By observing how this ratio evolves between different redshifts one can constrain the expansion history.

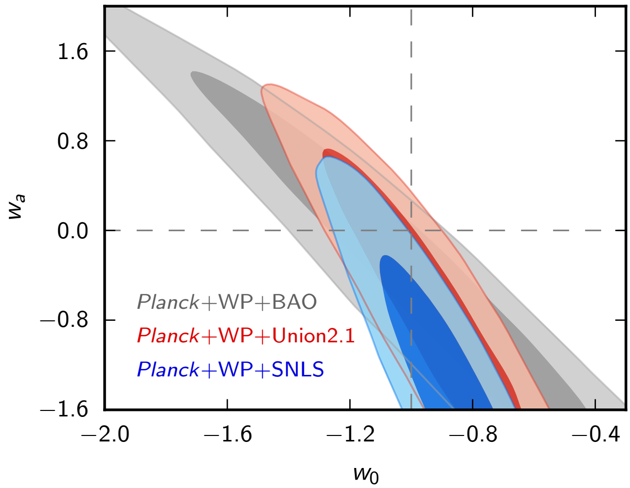

Figure 7.2. 68% and 95% confidence contours for w_0 and w_a, which parametrize the equation of state of dark energy. The form of the parametrization is given in (7.17). Union 2.1 [66] and SNLS [64] are SNe Ia datasets. The dashed line corresponds to a cosmological constant. Plot taken from [62].

The ΛCDM model Combining measurements of SNe Ia, BAO and the CMB gives constraints on the different constituents of the universe and their properties. Of primary interest is the Hubble constant H_0 (the value of the Hubble function today), the total matter content Ω_m and the dark energy content Ω_de. Furthermore, one outstanding question is to know the equation of state of dark energy. As of writing, the CMB measurements by the Planck satellite provide the most stringent constraints on these parameters [62]. A joint analysis of SNe Ia, BAO and the CMB, performed in [64], gives

$$\Omega_\mathrm{m} = 0.314 \pm 0.020, \qquad w_\mathrm{de} = -0.994 \pm 0.069, \qquad H_0 = 67.32 \pm 1.98. \tag{7.16}$$

These fits assume a flat universe and constant dark energy equation of state w_de. In Fig. 7.1 the individual constraints from SNe Ia, CMB and CMB+BAO are shown for a universe that allows curvature. One sees that the CMB and BAO measurements alone are enough to constrain the dark matter and dark energy components, favouring a flat universe.

It is only one model, the cosmological constant, that predicts a constant dark energy of state. One should therefore, in principle, fit the observational data to a time-varying equation of state. A common parametrization of the time variation of the equation of state of dark energy is

$$w_{\rm de}\left(z\right) = w_0 + w_a \frac{z}{1+z}. \tag{7.17}$$

In Fig. 7.2 constraints on w_0 and w_a are shown. One sees that current constraints on the properties of the dark energy components are weak; while data is consistent with dark energy being a constant vacuum energy component, it also allows for significant dynamics.

7.2 Cosmology with a non-dynamical reference metric

The original formulation of massive gravity, usually referred to as de Rham-Gabadadze-Tolley (dRGT) theory, does not contain any kinetic term for the second metric f. Instead, f is fixed to a certain form, usually the flat Minkowski metric, by hand. The equations of motion thus only determine g.

The search for cosmological solutions for this setup, with matter coupling to g only, showed that the theory did not accommodate flat or closed FLRW solutions [67]. While it is possible to have open FLRW solutions [68, 69], these were shown to have instabilities [70, 71, 72, 73]. The introduction of the effective metric, defined in (5.3), made it possible to couple parts of the matter sector to g, and another part to the effective metric. Reference [46], in which the effective metric was first introduced, also explored the possibility of viable cosmologies using this coupling. The perturbative stability of this setup was proved in [74]. It was shown in Paper V, however, that only highly contrived setups, with a fine-tuned scalar field potential, would yield a universe similar to that of ΛCDM. This provides a strong incentive to let the f metric become dynamical, since this leads to viable expansion histories using the standard techniques of cosmology.

In the setup with a fixed reference metric and matter coupling to the effective metric, given by

$$g^{\rm eff}_{\mu\nu} = \alpha^2 g_{\mu\nu} + 2\alpha\beta g_{\mu\alpha} X^\alpha_\nu + \beta^2 \eta_{\mu\nu}, \qquad X^\mu_\nu = \left(\sqrt{g^{-1}\eta}\right)^\mu_{\ \nu}, \tag{7.18}$$

the equations of motion are

$$\begin{aligned}(X^{-1})^{(\mu}{}_{\alpha}G^{\nu)\alpha}+m^2\sum_{n=0}^{3}(-1)^n\beta_n g^{\alpha\beta}(X^{-1})^{(\mu}{}_{\alpha}Y^{\nu)}_{(n)\beta}\\ =\frac{\alpha}{M_{\rm Pl}^2}\det{(\alpha+\beta X)}\left(\alpha(X^{-1})^{(\mu}{}_{\alpha}T^{\nu)\alpha}+\beta T^{\mu\nu}\right).\end{aligned} \tag{7.19}$$

These equations were derived in [50]. Since X^{μ}_{ν} is not symmetric, a further simplification of the equations of motion is not possible. Using the FLRW ansatz (7.30) for g,

$$ds_g^2=-N^2(t)dt^2+a^2(t)\delta_{ij}dx^idx^j, \tag{7.20}$$

the effective metric becomes

$$ds^2_{g_{\rm eff}}=-N^2_{\rm eff}(t)dt^2+a^2_{\rm eff}(t)\delta_{ij}dx^idx^j, \tag{7.21}$$

where

$$N_{\rm eff}=\alpha N+\beta,\qquad a_{\rm eff}=\alpha a+\beta. \tag{7.22}$$

The Bianchi constraint is

$$m^2M_{\rm Pl}^2a^2P(a)\dot{a}=\alpha\beta a^2_{\rm eff}p\dot{a}, \tag{7.23}$$

where

$$P(a)\equiv\beta_1+\frac{2\beta_2}{a}+\frac{\beta_3}{a^2}. \tag{7.24}$$

We see that the possibility of a time-varying scale factor is highly restricted. For example, if we assume the standard constant equation of state $p=w\rho$, then (7.23) becomes

$$m^2M_{\rm Pl}^2a^2P(a)=\alpha\beta w a^2_{\rm eff}\rho. \tag{7.25}$$

ρ will only be a function of a (or, equivalently, $a_{\rm eff}$), and unless the left hand side has the same functional form for a as the right hand side, eq. (7.25) is not consistent with a time variation of a. What this tells us is that the pressure has to depend on the lapse, and not only the scale factor, in order to have a time varying a. Physically, this corresponds to two different scenarios. A perfect fluid could be affected by the global expansion, and it would then depend on the scale factor. But it could also be affected by internal dynamics, such as the decay time of its constituent particles. It would then also depend on its proper time coordinate, which is $N_{\text{eff}}dt$. In this case, the pressure would also depend on the lapse.

One way to have a lapse-dependent pressure is to use a scalar field. The stress-energy tensor for the scalar field is

$$T^{\mu\nu} = \nabla^{\mu}_{\text{eff}}\chi\nabla^{\nu}_{\text{eff}}\chi - \left(\frac{1}{2}\nabla_{\alpha}\chi\nabla^{\alpha}_{\text{eff}}\chi + V(\chi)\right)g^{\mu\nu}_{\text{eff}}, \tag{7.26}$$

where $\nabla^{\mu}_{\text{eff}} \equiv g^{\mu\nu}_{\text{eff}}\nabla^{\text{eff}}_{\nu}$ and $V(\chi)$ is the potential of the scalar field. The density and pressure associated to this scalar are

$$\rho_{\chi} = \frac{\dot{\chi}^2}{2N^2_{\text{eff}}} + V(\chi), \qquad p_{\chi} = \frac{\dot{\chi}^2}{2N^2_{\text{eff}}} - V(\chi). \tag{7.27}$$

As long as the potential is not independent of the lapse and $\dot{\chi} \neq 0$, it can be shown that this setup allows for a dynamical a.

In a realistic setup, however, one should also include other forms of matter. In Paper V it was shown that in order to get sensible cosmologies, where the effective Hubble function does not display a pathological behaviour, one needs highly contrived potentials for the scalar field. A dynamical a is thus possible, but at the price of a fine-tuned potential. Since standard techniques of cosmology, i.e. using only perfect fluids to model the expansion history, is not applicable in the dRGT framework with a effective metric, it is therefore natural to consider the scenario where the second metric f is also dynamical.

7.3 Coupling matter to one metric

We now turn to the theory of massive bigravity to see how it modifies the Friedmann equation (7.8). In this section we will couple only g to matter through a term $\sqrt{-g}\mathcal{L}_{\text{m}}$ in the Lagrangian. This setup has also been studied in [75, 76, 77, 78]. The equations of motion become, using the scaling described in section 6.1,

$$R_{\mu\nu}(g) - \frac{1}{2} g_{\mu\nu} R(g) + m^2 \sum_{n=0}^{3} (-1)^n \beta_n g_{\mu\lambda} Y^{\lambda}_{(n)\nu} \left(\sqrt{g^{-1}f}\right) = \frac{T_{\mu\nu}}{M_g^2}, \tag{7.28}$$

$$R_{\mu\nu}(f) - \frac{1}{2} f_{\mu\nu} R(f) + m^2 \sum_{n=0}^{3} (-1)^n \beta_{4-n} f_{\mu\lambda} Y^{\lambda}_{(n)\nu} \left(\sqrt{f^{-1}g}\right) = 0. \tag{7.29}$$

For the background cosmology, we use the following ansatz:[2]

$$ds_g^2 = -N_g^2 dt^2 + a_g^2 \delta_{ij} dx^i dx^j, \tag{7.30}$$
$$ds_f^2 = -N_f^2 dt^2 + a_f^2 \delta_{ij} dx^i dx^j, \tag{7.31}$$

The Hubble functions for the two metrics are defined as

$$H_g \equiv \frac{\dot{a}_g}{N_g a_g}, \qquad H_f \equiv \frac{\dot{a}_f}{N_f a_f}. \tag{7.32}$$

Since we only couple g to matter, it is a_g and H_g that are the relevant functions when it comes to observables.

It turns out that the ratio between the two scale factors is an important function, and we denote it by r:

$$r \equiv \frac{a_f}{a_g}. \tag{7.33}$$

Defining the following parameter combinations:

$$B_0(r) \equiv \beta_0 + 3\beta_1 r + 3\beta_2 r^2 + \beta_3 r^3, \tag{7.34}$$
$$B_1(r) \equiv \beta_1 r^{-3} + 3\beta_2 r^{-2} + 3\beta_3 r^{-1} + \beta_4, \tag{7.35}$$

the two Friedmann equations become

$$3H_g^2 = \frac{\rho}{M_g^2} + m^2 B_0, \tag{7.36}$$
$$3H_f^2 = m^2 B_1. \tag{7.37}$$

[2] Notice that different notations and conventions are used in the different papers. Throughout this thesis, we will use the same notation, however.

Here ρ contains all energy components besides a constant vacuum energy density, which is degenerate with the β_0-term. The continuity equation for the energy density still holds:

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0. \tag{7.38}$$

The Bianchi constraint gives

$$m^2\left(\beta_1 + 2\beta_2 r + \beta_3 r^2\right)\left(N_f \dot{a}_g - N_g \dot{a}_f\right) = 0. \tag{7.39}$$

This constraint yields two possible solution branches. In the first branch, obtained by putting the first parenthesis to zero, we get an algebraic condition for a constant r. This branch is therefore called the algebraic branch. This branch can be shown to yield solutions that are identical to general relativity, with the addition of an effective cosmological constant that the depends on the β_i parameters. The massive fluctuations also vanish in the algebraic branch.

In the second branch, we instead get

$$\frac{N_f}{N_g} = \frac{\dot{a}_f}{\dot{a}_g}. \tag{7.40}$$

This branch is called the dynamical branch, since r can have a time evolution. By combining (7.36) and (7.37), we get the following polynomial equation for r:

$$\beta_3 r^4 + (3\beta_2 - \beta_4) r^3 + 3(\beta_1 - \beta_3) r^2 + \left(\frac{\rho}{M_g^2 m^2} + \beta_0 - 3\beta_2\right) r - \beta_1 = 0 \tag{7.41}$$

The background equations are thus solved by using (7.41) to determine r in terms of ρ, and then using (7.36) to determine H_g.

For an expanding universe, eq. (7.41) has two important limits: the late universe when $\rho \to 0$, and the early universe when $\rho \to \infty$. For the late universe, we see that r will approach a constant, which means that both the g- and f-metric will approach a de Sitter form in the future. For the early universe, there are two possible asymptotic behaviours for r. One is that r goes to zero, i.e. $a_f \ll a_g$, which means that we will recover the standard form of general relativity for H_g. The other possibility is that r goes to infinity, i.e. $a_f \gg a_g$, in the early universe. Notice that these respective scenarios tells us how the two scale factors evolve with respect to one another in the early universe.

Since (7.41) is a quartic polynomial, the general solution with arbitrary β_i parameters is highly unwieldy. A comprehensive analysis was performed in [78, 77], whereas Paper I focused on a few interesting parameter choices. Let us here focus on one of them. For $\beta_1 = \beta_3 = 0$, we have that

$$r^2 = \frac{3\rho/\left(m^2 M_g^2\right) + \beta_0 - 3\beta_2}{\beta_4 - 3\beta_2} \tag{7.42}$$

and

$$H_g^2 = \frac{\beta_4}{\beta_4 - 3\beta_2}\frac{\rho}{3M_g^2} + \frac{m^2\left(\beta_0\beta_4 - 9\beta_2^2\right)}{3\left(\beta_4 - 3\beta_2\right)}. \tag{7.43}$$

We see that the Hubble function will evolve exactly as in general relativity, with an effective cosmological constant and a rescaled Planck mass for ρ. This rescaling of the effective Planck mass is an interesting phenomenon that could alleviate the need for dark matter for the expansion history. In order to show this, one needs, however, spherically symmetric solution in order to see what the local Planck mass becomes.

A full statistical analysis of the allowed β_i parameter values was presented in [77]. For the minimal β_i models, where only one of the β_i is non-vanishing, only the case of β_0 and β_1 could give good fits to data. The case of β_0 is obvious, since this is degenerate with a ΛCDM universe. β_1 gives an effective, and observationally acceptable, dark energy component. The β_2 and β_3 cases were observationally excluded. The case of only β_4 non-vanishing is automatically excluded, since this only gives a cosmological constant contribution for f, which does not give a ΛCDM-like behaviour for the expansion history. When including more than two or more non-vanishing β_i parameters, almost all cases produced good fits to data (the exceptions being when either β_2 and β_4, or β_3 and β_4 were the only non-vanishing parameters).

7.4 Coupling matter to two metrics

When coupling matter to the effective metric

$$g^{\text{eff}}_{\mu\nu} = \alpha^2 g_{\mu\nu} + 2\alpha\beta g_{\mu\alpha}X^\alpha_\nu + \beta^2 f_{\mu\nu}, \qquad X^\mu_\nu = \left(\sqrt{g^{-1}f}\right)^\mu_{\ \nu}, \tag{7.44}$$

the equations of motion become [50] (again using the scalings described in section 6.1)

$$(X^{-1})^{(\mu}{}_{\alpha}G_g^{\nu)\alpha}+m^2\sum_{n=0}^{3}(-1)^n\beta_n g^{\alpha\beta}(X^{-1})^{(\mu}{}_{\alpha}Y^{\nu)}_{(n)\beta}$$
$$=\frac{\alpha}{M_{\text{eff}}^2}\sqrt{\frac{\det g_{\text{eff}}}{\det g}}\left(\alpha(X^{-1})^{(\mu}{}_{\alpha}T^{\nu)\alpha}+\beta T^{\mu\nu}\right), \tag{7.45}$$

$$X^{(\mu}{}_{\alpha}G_f^{\nu)\alpha}+m^2\sum_{n=0}^{3}(-1)^n\beta_{4-n} f^{\alpha\beta}X^{(\mu}{}_{\alpha}\hat{Y}^{\nu)}_{(n)\beta}$$
$$=\frac{\beta}{M_{\text{eff}}^2}\sqrt{\frac{\det g_{\text{eff}}}{\det f}}\left(\alpha T^{\mu\nu}+\beta X^{(\mu}{}_{\alpha}T^{\nu)\alpha}\right). \tag{7.46}$$

Using the ansatz (7.30) and (7.31) for the g and f metric, the effective metric becomes

$$ds_{\text{eff}}^2=-N^2dt^2+a^2\delta_{ij}dx^idx^j, \tag{7.47}$$

where

$$N=\alpha N_g+\beta N_f, \tag{7.48}$$
$$a=\alpha a_g+\beta a_f. \tag{7.49}$$

Since matter couples to the effective metric, it is the expansion rate of the effective metric that is observed, rather than that of g or f. We are therefore interested in the behaviour of the effective Hubble function

$$H\equiv\frac{\dot{a}}{Na}. \tag{7.50}$$

The Bianchi-like constraint becomes

$$\left[m^2\left(\beta_1a_g^2+2\beta_2a_ga_f+\beta_3a_f^2\right)-\frac{\alpha\beta a^2p}{M_{\text{eff}}^2}\right](N_f\dot{a}_g-N_g\dot{a}_f)=0. \tag{7.51}$$

Just as in the case (7.39) of singly coupling, we get two possible branches. Unlike in the case of singly coupling, however, where the algebraic branch gave rise to expansion histories identical to that of ΛCDM, the phenomenology when using the effective coupling is much richer. The possible expansion histories were explored in appendix C of Paper IV. Focusing once again on the dynamical branch, for which

$$\frac{N_f}{N_g} = \frac{\dot{a}_f}{\dot{a}_g}, \tag{7.52}$$

the equations that determine H and r become

$$H^2 = \frac{\rho}{6M_{\text{eff}}^2}\left(\alpha + \beta r\right)\left(\alpha + \beta r^{-1}\right) + \frac{m^2\left(B_0 + r^2 B_1\right)}{6\left(\alpha + \beta r\right)^2}, \tag{7.53}$$

$$0 = \frac{\rho}{M_{\text{eff}}^2}\left(\alpha + \beta r\right)^3\left(\alpha - \beta r^{-1}\right) + m^2\left(B_0 - r^2 B_1\right). \tag{7.54}$$

Here ρ are all matter components, and they satisfy the continuity equation with respect to the effective metric:

$$\dot{\rho} + 3\frac{\dot{a}}{a}\left(\rho + p\right) = 0. \tag{7.55}$$

At first sight, (7.53) looks like the ordinary Friedmann equation, but with a modulation of the effective Planck mass and a dynamical dark energy component, both of which depend on r. The interpretation is not as straightforward, however, since one can use (7.54) to rewrite (7.53) in various ways.

In the early universe $\rho \to \infty$, and from (7.54) we must have either $r \to \beta/\alpha$ or $r \to -\alpha/\beta$. One can show that the later case is observationally excluded. In the late universe, when $\rho \to 0$, one possibility is that r goes to a constant r_c which solves

$$\beta_3 r_c^4 + \left(3\beta_2 - \beta_4\right) r_c^3 + 3\left(\beta_1 - \beta_3\right) r_c^2 + \left(\beta_0 - 3\beta_2\right) r_c - \beta_1 = 0. \tag{7.56}$$

Another possibility is that $r \to \infty$ in a way such that (7.54) is still satisfied. Disregarding the latter case, r will always be finite, starting from β/α and ending at r_c. The late time asymptotic behaviour for such a universe is that of a de Sitter spacetime, with cosmological constant

$$\Lambda = \frac{m^2\left[\beta_1 + \left(\beta_0 + 3\beta_2\right) r_c + 3\left(\beta_1 + \beta_3\right) r_c^2 + \left(3\beta_2 + \beta_4\right) r_c^3 + \beta_3 r_c^4\right]}{2r_c\left(\alpha + \beta r_c\right)^2}. \tag{7.57}$$

For parameter choices such that $r_c = \beta/\alpha$ in the solution of (7.56), r is constant at all times. From (7.52) it then follows that the two metrics are conformally related, $f_{\mu\nu} = r_c^2 g_{\mu\nu}$. Unlike the case in the singly coupled setup where conformally related metrics implied de Sitter, this does not need to hold when using the effective coupling.

Partially massless gravity The parameter choice $\beta_1 = \beta_3 = 0$, $\beta_0 = 3\beta_2 = \beta_4$ is special, since a new gauge symmetry arises at the linear level, which is believed to also exist at all orders (as described in section 6.2). In the case of singly coupling, it is not possible to impose the partially massless parameter condition in the presence of matter sources. This no longer holds true when using the effective coupling. For this parameter choice, we have that

$$H^2 = \frac{\alpha^2+\beta^2}{3M_{\rm eff}^2}\rho + \frac{m^2\beta_0}{3(\alpha^2+\beta^2)}, \tag{7.58}$$

and $r = \beta/\alpha$. The expansion rate is thus identical to that of ΛCDM, but with a renormalized matter coupling. It was also shown in [50] that structure formation will look identical to general relativity when r is a constant. This theory therefore stands out as an intriguing extension of general relativity, with deviation possibly only appearing at the non-linear level.

Vacuum energy As mentioned in section 6.2, the vacuum energy will renormalize all β_i parameters when using the effective coupling. This comes from the fact that matter loops will give rise to a term $\sqrt{-\det g_{\rm eff}}\Lambda_v$ in the Lagrangian. Since

$$\sqrt{-\det g_{\rm eff}} = \sqrt{-\det g}\det\left(\alpha + \beta X\right) = \sqrt{-\det g}\sum_{n=0}^{4}\alpha^{4-n}\beta^n e_n\left(X\right), \tag{7.59}$$

the β_i parameters will get contributions

$$m^2\beta_n = \Lambda_v\alpha^4\left(\frac{\beta}{\alpha}\right)^n \tag{7.60}$$

from matter loops. If all β_i parameters are generated from these matter loops, the Friedmann equation becomes

$$H^2 = \frac{\left(\alpha^2+\beta^2\right)\rho}{M_{\rm eff}^2} + \frac{\left(\alpha^2+\beta^2\right)\Lambda_v}{3}. \tag{7.61}$$

This is equivalent to a ΛCDM universe.

A disadvantage with the effective coupling as compared to singly coupling is that the values of *all* β_i parameters have to be fine-tuned, due

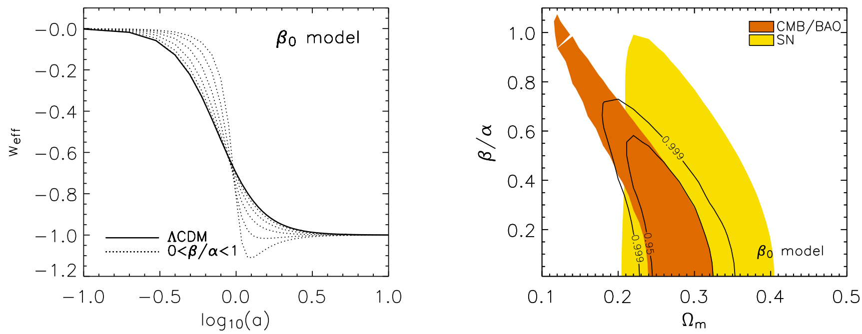

Figure 7.3. *Left panel:* Effective equation of state of the $w_{\rm eff}$, defined in eq. (7.62), for the ΛCDM model (solid line) and β_0 minimal model (dotted line), for different values of β/α. As $\beta/\alpha \to 0$, the β_0 minimal model approaches the ΛCDM model. In all cases, $\Omega_m = 0.3$. *Right panel:* Confidence contours for Ω_m and β/α for the β_0 model. Shaded regions correspond to individual constraints by SNe Ia and CMB/BAO (95% confidence level), solid lines are combined constraints (95% and 99.9% confidence level). The plots are taken from Paper IV.

to the renormalization from matter loops. Also, since one can not unambiguously isolate the so-called self-acceleration of the model, i.e. an acceleration of the expansion that is not due to a vacuum energy component, it makes it hard to motivate the introduction of the effective coupling from the perspective of the cosmological constant problem.

Observational constraints In Paper IV observational constraints on the minimal β_i-models were studied. Minimal here means that only one of the β_i parameters is non-zero. The duality property of the solutions, described in section 6.1, implies that only β_0, β_1 and β_2 have to be studied. Since a constant r yields expansion histories identical to ΛCDM, for $\beta/\alpha = \{0, 1/\sqrt{3}, 1\}$ the β_0, β_1 and β_2 cases, respectively, will give excellent fits to data.

A combined, effective equation of state can be defined as

$$w_{\rm eff} = -1 - \frac{1}{3}\frac{d\log H^2}{d\log a}. \tag{7.62}$$

In the left panel of Fig. 7.3 we show how $w_{\rm eff}$ differs, as a function of β/α, from that of a ΛCDM universe for the β_0 minimal model. We see that the minimal β_0-model approaches the ΛCDM model both in the past and

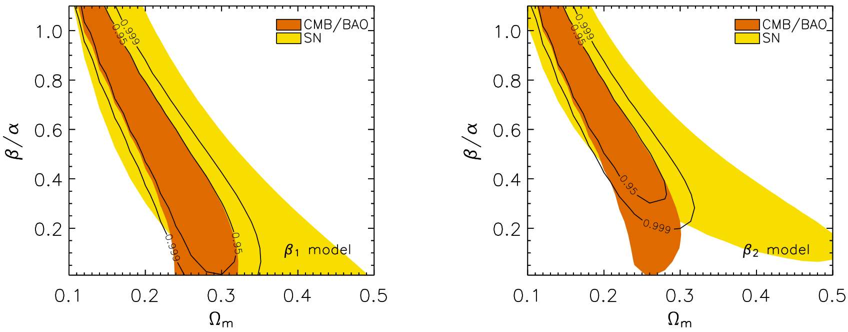

Figure 7.4. Confidence contours for Ω_m and β/α for the β_1 and β_2 minimal models. Shaded regions correspond to SNe Ia and CMB/BAO constraints (95% confidence level). Solid lines are combined constraints (95% and 99.9% confidence level). The plots are taken from Paper IV.

the future, as expected from the asymptotic behavior of (7.54). In the right panel of Fig. 7.3, we show the allowed parameter space when using SNe Ia and CMB/BAO constraints (shaded regions, with 95% confidence level). The combined constraints are shown as solid lines, with 95% and 99.9% confidence levels. We see that for $\beta/\alpha = 0$, the β_0 model predicts $\Omega_m \sim 0.3$, which is expected since this value of β/α corresponds to the ΛCDM model. Ω_m is here defined as

$$\Omega_m \equiv \frac{\alpha^2 \rho_0}{3M_{\text{eff}}^2 H_0^2}. \tag{7.63}$$

In Fig. 7.4 constraints for the β_1 and β_2 models are given. An interesting feature of these models is that a higher value of β/α allows for a lower value of Ω_m. The reason is that β/α will introduce prefactors in front of the matter contribution. Increasing β/α increases this prefactor, and thus allows for a decreased matter content.

The effective coupling allows for viable cosmologies in the minimal β_i models. In the case of singly coupling, this was only possible for the β_0 and β_1 case. This is clearly an advantage with the effective coupling.

Chapter 8

Structure formation

In the previous chapter we studied how massive gravity modified the relationship between the energy content of the universe and the expansion history. The expansion of the universe is, of course, not the end of the story in cosmology. Gravitational attraction will produce bound structures, such as galaxies, that are not affected by the expansion internally. How such bound structures are formed depends on the strength of gravity at different length scales. In order to study this effect, one assumes a background FLRW metric, and perturbations on top of that background. Perturbations exist both in the metric and in the matter content. By Fourier transforming the perturbation fields, one can study the growth of each perturbation wavelength separately. This allows for an comparison between the perturbation wavelength and the Hubble scale, mass length scale and length of the sound horizon.

In section 8.1-8.3 we give a background to structure formation in general relativity. In section 8.4-8.6 we discuss how perturbations evolve in massive gravity.

8.1 Probing the universe to first order

In order to study perturbations on a homogenous and isotropic background, the metric can be decomposed into a part that is of a FLRW form, denoted by $\bar{g}_{\mu\nu}$, and a perturbation $h_{\mu\nu}$, in the following way:

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}. \tag{8.1}$$

The components of $h_{\mu\nu}$ can in turn be decomposed into scalar, vector and tensor parts (defined by their transformation properties under a coordinate change) as

$$h_{00} = -N^2 E, \tag{8.2}$$
$$h_{i0} = aN\left(\partial_i F + G_i\right), \tag{8.3}$$
$$h_{ij} = a^2\left(A\delta_{ij} + \partial_i\partial_j B + \partial_j C_i + \partial_i C_j + D_{ij}\right), \tag{8.4}$$

where

$$D_{ij} = D_{ji}, \quad \partial_i C_i = \partial_i G_i = \partial_i D_{ij} = D_{ii} = 0. \tag{8.5}$$

E, F, A and B are scalars, G_i and C_i vectors and D_{ij} is a tensor. At linear level, the scalar, vector and tensor parts do not mix with each other, and one can study them separately. In the rest of this thesis, we will focus on the scalar sector, since this is related to the growth of matter in the universe.

The scalar part of the stress-energy tensor is:

$$T^0_0 = -\rho(1+\delta), \tag{8.6}$$
$$T^i_0 = -(\rho+p)v^i, \tag{8.7}$$
$$T^0_i = (\rho+p)(v_i + \partial_i F_g), \tag{8.8}$$
$$T^i_j = (p+\delta p)\delta^i_j + \Sigma^i_j. \tag{8.9}$$

Here $\delta = \delta\rho/\rho$, $v^i \equiv dx^i/dt$ and $\Sigma^l_l = 0$. We also define $\theta \equiv \partial_i v^i$, which will be used in the equations of motion for the perturbations. For pressureless matter, we have $p = \delta p = \Sigma^l_l = 0$.

The equations of motion for the perturbations, in the case of general relativity, are given schematically by

$$\delta\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right) = \frac{\delta T_{\mu\nu}}{M_g^2}. \tag{8.10}$$

The perturbations in the metric and stress-energy tensor can be written in terms of their Fourier components as (ϕ symbolizes the different fields)

$$\phi\left(\vec{x}, t\right) = \int \frac{d^3k}{\left(2\pi\right)^3}\tilde{\phi}\left(k, t\right)e^{i\vec{k}\cdot\vec{x}}. \tag{8.11}$$

Because of the isotropy of the background, we can decompose the Fourier components as

$$\tilde{\phi}(\vec{k},t) = \sum_n \alpha_n(\vec{k})\phi_{nk}(t)\,. \tag{8.12}$$

Here $\alpha_n(\vec{k})$ is a time-independent amplitude that depends on the direction, while $\phi_{nk}(t)$, which gives the time evolution of the perturbations, only depends on the magnitude $|\vec{k}|$. n labels the different solutions (e.g. growing and decaying modes), and since we are looking at linear perturbations $\tilde{\phi}$ will be a superposition of the solutions. In the following we will use ϕ to signify both $\phi(\vec{x},t)$ and $\tilde{\phi}(\vec{k},t)$; it should be clear from the context if we are discussing the fields in position space or Fourier space.

But before we discuss the perturbations, we have to briefly mention the coordinate freedom available in their description.

8.2 To gauge or not to gauge

General relativity is a covariant theory, which means that the equations of motion look the same in all coordinate systems. Under an infinitesimal coordinate transformation $x^\mu \to x^\mu + \xi^\mu$, $h_{\mu\nu}$ will transform as

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu. \tag{8.13}$$

Looking at the scalar components of $h_{\mu\nu}$, we can decompose ξ^μ as $\xi^0 = \epsilon^0$ and $\xi^i = \partial_i \epsilon$. The scalar perturbations will then transform as

$$\Delta A = -2\frac{\dot{a}}{a}\epsilon^0, \qquad \Delta B = -2\epsilon, \tag{8.14}$$

$$\Delta E = -2N^{-1}\partial_t\left(N\epsilon^0\right), \qquad \Delta F = \frac{N}{a}\epsilon^0 - \frac{a}{N}\dot{\epsilon}. \tag{8.15}$$

This means that the scalar perturbations by themselves can not carry all physical information, since one can in principle choose a coordinate system where some of them vanish. In order to isolate the actual physics, one therefore has two options: Either fix a coordinate system completely and do all computations in that system (where care has to be taken so that there is no residual coordinate freedom left), or form coordinate invariant combinations of the scalar perturbations, and work only with those. Choosing a coordinate system is usually referred to as fixing a gauge, and coordinate invariant variables are usually called gauge invariant variables.

A popular gauge choice is the Newtonian gauge, in which $F = B = 0$. Equivalently, one can define the gauge invariant combination

$$\Psi = -\frac{1}{2}A + \frac{\dot{a}}{N}\left(\frac{a}{2N}\dot{B} - F\right), \tag{8.16}$$

$$\Phi = \frac{1}{2}E - \frac{1}{N}\partial_t\left[a\left(\frac{a}{2N}\dot{B} - F\right)\right]. \tag{8.17}$$

Ψ and Φ coincide with $-A/2$ and $E/2$ in the Newtonian gauge. One can therefore consistently work in the Newtonian gauge, and then promote the final results into gauge invariant quantities. Besides the computational advantage, the fields Ψ and Φ will also coincide with the ordinary Newtonian potential in the limit of weak fields and low velocities.

8.3 Growth of structure in general relativity

In this section we review the main features of the growth of structure as described by general relativity.[1] The evolution of perturbations in the universe is primarily governed by two things: the dominant energy component of the background (i.e. radiation, matter or dark energy), and the couplings between the different constituents of the perturbations. Dark matter, for example, does not couple electromagnetically whereas neutrinos only couple weakly. Baryons and photons will couple to each other when the universe had a temperature above roughly 3000 K, since photons then had enough energy to ionize the hydrogen atoms. As the universe cooled down, the photons eventually stopped interacting at a high enough rate with baryons to form a plasma. Instead, the photons started to propagate freely in the universe. The interaction rate between baryons and photons depend on the temperature of the photons, and the transition from a regime of high interaction to basically no interactions, the so called time of recombination[2], is of great importance in cosmology.

Let us first look at the case of the perturbation of a single fluid[3], and then discuss the rather involved case of perturbations of many fluids. The

---

[1] There exists many textbooks dealing with this topic. In this section we follow [81].

[2] The term *recombination* is truly a misnomer: it refers to the fact that protons and electrons could combine to form neutral hydrogen without being dissociated through high-energetic photons. But this was the *first* time of combination in the cosmic history, and this era should properly be called *combination*. At a later stage the first stars once again ionized the hydrogen; this era is properly called *reionization*.

[3] The term fluid here refers to any of the physical components available for an idealized fluid description, i.e. photons, baryons, dark matter, neutrinos, scalar fields etc. The fluid will be decomposed into a homogeneous and isotropic background, and

equations of motion for the perturbations of a single fluid in momentum space (a $'$ signifies a time derivative with respect to conformal time) are

$$k^2\Phi + 3\frac{a'}{a}\Phi' + 3\frac{a'^2}{a^2}\Phi = -4\pi G a^2 \delta\rho, \tag{8.18}$$

$$\Phi' + \frac{a'}{a}\Phi = -4\pi G a^2 \left[(\rho + p)\frac{\theta}{k^2}\right], \tag{8.19}$$

$$\Phi'' + 3\frac{a'}{a}\Phi' + \left(2\frac{a''}{a} - \frac{a'^2}{a^2}\right)\Phi = 4\pi G a^2 \delta p. \tag{8.20}$$

For the background we could define an equation of state w through $p = w\rho$. For the perturbations we can similarly define an equation of state as

$$u_s^2 = \frac{\delta p}{\delta\rho}. \tag{8.21}$$

This is called the sound speed of the perturbation. u_s^2 does not have to be the same as w, as happens e.g. in the case of matter perturbations in a dark energy dominated background. Assuming for the moment that $u_s^2 = w$ and combining eqs. (8.18) and (8.19), we get

$$\Phi'' + 3\frac{a'}{a}\left(1 + u_s^2\right)\Phi' + k^2 u_s^2 \Phi = 0. \tag{8.22}$$

This equation determine the evolution of Φ, and once the potential has been determined the energy density and velocity of the perturbation can be read off from (8.18) and (8.19). The last term in (8.22) depends on the wavenumber k of the perturbation. Comparing this with the second term, we see that when

$$\lambda \gg u_s H^{-1}, \tag{8.23}$$

where $\lambda = a/(2\pi k)$ is the physical wavelength of the perturbation, we can neglect the last term in (8.22). $u_s H$ is the sound horizon of the mode, which gives the characteristic length for causal interactions of the perturbations. For matter, which has $u_s = 0$, it is always zero, whereas for light, which has $u_s = 1/\sqrt{3}$, it is of the order of the cosmological horizon.

---

perturbations on top of that. The idealized fluid approximation, which uses energy densities, pressures, velocities and anisotropic stresses to describe the properties of a physical system, breaks down for the photon-baryonic plasma, and one has to resort to a more detailed description using Boltzmann's equations in this case.

Solving (8.22) for perturbations with wavelengths larger than the sound horizon we find one constant solution for Φ and one solution that is decaying. The decaying solutions are usually neglected: it is assumed that the universe was initially very homogenous, and since the decaying solutions were growing in the past, including them would lead to strong initial inhomogeneities.

From eq. (8.18) the density contrast, defined by

$$\delta_i \equiv \frac{\delta\rho_i}{\rho_i}, \tag{8.24}$$

will also be a constant for modes for which $\lambda \gg H^{-1}$.[4] These modes do not produce any structure, since, colloquially speaking, there have not been enough time for any significant gravitational self-interactions to occur.

The superhorizon modes will eventually become smaller than the horizon since H is an decreasing function of time. The behaviour of the subhorizon modes, for which

$$\lambda \gg u_s H \tag{8.25}$$

will behave very differently for relativistic and non-relativistic components. For a relativistic component (in a relativistic background), one can show, using once again eqs. (8.18) and (8.20), that

$$\Phi = -3\Phi_{(i)}\frac{\cos\left(ku_s\eta\right)}{\left(ku_s\eta\right)^2}, \tag{8.26}$$

$$\delta = -3\delta_{(i)}\cos\left(ku_s\eta\right). \tag{8.27}$$

$\Phi_{(i)}$ and $\delta_{(i)}$ are the value of the potential and density contrast when the perturbation wavelength was larger then the horizon. We see that the potential will undergo damped oscillations and the density contrast will oscillate. This means that there is no growth of structure for a relativistic component.

[4] These modes are usually called superhorizon modes. Modes for which $\lambda \ll H^{-1}$ are called subhorizon modes. For δ, it is the cosmological horizon H^{-1} which is important. For the potential Φ, it is the sound horizon $u_s H^{-1}$. We will also refer to potential modes as super- or subhorizon, depending on if they are larger or smaller than the sound horizon.

The non-relativistic component (in a non-relativistic background) still has a constant Φ as solution, but the density of the subhorizon mode evolves as

$$\delta = -\frac{\eta^2 k^2}{6}\Phi \sim a. \tag{8.28}$$

The density contrast is thus growing linearly with the scale factor, and structure is forming. Also note that δ depends linearly on k, so that modes with shorter wavelength have a larger density contrast, or, equivalently, modes that enter the horizon earlier have had more time to grow.

Perturbations in a relativistic component in a radiation dominated background correspond to the early universe, and perturbations in a matter components in a matter dominated background correspond to the universe from the time of recombination up to a redshift of about two. At later times, the dark energy component will start to dominate the background. The potential, in such a background, will be inversely proportional to the scale factor and thus decay, whereas both the sub- and superhorizon matter perturbation will have a constant energy density contrast. This means that there is no growth of structure in such a universe. We thus see that there only exists a limited timespan in the history of the universe when structures could form.

It is also important to note that a matter component in a relativistic background, such as cold dark matter, will grow slowly (unlike the baryonic component which is coupled to radiation and therefore does not grow). When the matter component dominates the background, but before recombination, the dark matter component will grow linearly with the scale factor. This allows for large gravitational potentials to be set up by the dark matter component, which the baryonic component can fall into after recombination. Without this enhancement of the baryonic growth, it would not be possible to form non-linear structures such as galaxies.

When dealing with multiple components, the right hand sides of eqs. (8.18-8.20) have to be replaced with the total energy density, velocity and pressure. The individual components will also satisfy the following conservation equations:

$$\delta\rho_\lambda' + 3\frac{a'}{a}\left(\delta\rho_\lambda + \delta p_\lambda\right) - \left(\rho_\lambda + p_\lambda\right)\left(\theta_\lambda + 3\Phi'\right) = 0, \tag{8.29}$$

$$\left[\left(\rho_\lambda + p_\lambda\right)\frac{v_\lambda}{k^2}\right]' + 4\frac{a'}{a}\left(\rho_\lambda + p_\lambda\right)\frac{v_\lambda}{k^2} + \delta p_\lambda + \left(\rho_\lambda + p_\lambda\right)\Phi = 0. \tag{8.30}$$

Initial data for these perturbations can be divided in two main sets, which are *adiabatic* and *isocurvature* initial conditions. They correspond to two different creation mechanisms for the perturbations. For adiabatic initial conditions, there is one function ϵ that describe all the perturbations:

$$\delta\rho_\lambda = \rho'_\lambda \epsilon, \qquad \delta p_\lambda = p'_\lambda \epsilon. \tag{8.31}$$

This means that all perturbations fluctuate in the same manner, and there is thus no change in the relative number density of the perturbations. These perturbation arise when there is a single, global mechanism that is the source of the initial conditions, for example a single inflaton field.

For isocurvature initial conditions, only one component has a non-vanishing density contrast initially, whereas the other components are homogenous. These initial conditions requires that the universe was in a state of thermal non-equilibrium (otherwise only adiabatic modes, related to the overall temperature fluctuations, would be created). CMB data can discriminate between these two initial conditions, and strongly favours adiabatic modes.

Let us now briefly describe the main features of the evolution of coupled perturbations. The main components in our universe are cold dark matter (CDM), neutrinos, photons and baryons. Of these, photons and baryons will couple before recombination and has to be treated as a single fluid. After recombination, photons will freely stream in the universe which we today observe as the CMB.

The baryon-photon plasma will undergo acoustic oscillations before recombination, and there is thus no growth of structure. The CDM component, on the other hand, will grow logarithmically during the radiation dominated regime, and linearly during the matter dominated regime. When the photons decouple from the baryons at recombination, the oscillations in the plasma are imprinted in the both the photon and baryon distribution. These oscillations are seen in the CMB and as baryonic acoustic oscillations. Since the density contrast of the photons is related to the relative temperature fluctuations at the time of decoupling, observations of the CMB will directly probe the structure of the density contrast at the time of decoupling.

After decoupling, the baryonic density contrast started to grow: If there had not been a CDM component that could have grown unimpeded prior to decoupling, and thus increase the rate of baryonic growth, the baryonic density contrast today would only have been on the order of 10^{-2} (since the baryonic density contrast is proportional to the scale factor) and no galaxies or other structure could have been formed. This is one of the main reasons why a CDM component has to be postulated.

8.4 Perturbations in massive bigravity

We now turn to the investigation of the growth of structure in the Hassan-Rosen theory. We will use the setup where only g couples to matter. Both g and f are perturbed, and we will use the following form for the perturbations (mimicking the approach used in general relativity):

$$ds_g^2 = -N_g^2 \left(1 + E_g\right) dt^2 + 2N_g a_g \partial_i F_g dx^i dt + a_g^2 \left[\left(1 + A_g\right) \delta_{ij} + \partial_i \partial_j B_g\right] dx^i dx^j, \tag{8.32}$$

$$ds_f^2 = -N_f^2 \left(1 + E_f\right) dt^2 + 2N_f a_f \partial_i F_f dx^i dt + a_f^2 \left[\left(1 + A_f\right) \delta_{ij} + \partial_i \partial_j B_f\right] dx^i dx^j. \tag{8.33}$$

In order to compute the first order perturbation of the field equations, i.e.

$$\delta \left[R_{\mu\nu}(g) - \frac{1}{2} g_{\mu\nu} R(g)\right] + m^2 \sum_{n=0}^{3} (-1)^n \beta_n \delta \left[g_{\mu\lambda} Y^{\lambda}_{(n)\nu}\left(\sqrt{g^{-1}f}\right)\right] = \frac{1}{M_g^2} \delta T_{\mu\nu}, \tag{8.34}$$

$$\delta \left[R_{\mu\nu}(f) - \frac{1}{2} f_{\mu\nu} R(f)\right] + m^2 \sum_{n=0}^{3} (-1)^n \beta_{4-n} \delta \left[f_{\mu\lambda} Y^{\lambda}_{(n)\nu}\left(\sqrt{f^{-1}g}\right)\right] = 0, \tag{8.35}$$

the square root matrices have to be evaluated. For perturbations around a diagonal background, one can use the following formula:

$$\left(\sqrt{D + \varepsilon}\right)_{ij} = \sqrt{D_{ii}} \delta_{ij} + \frac{\varepsilon_{ij}}{\sqrt{D_{ii}} + \sqrt{D_{jj}}} + \mathcal{O}\left(\varepsilon^2\right). \tag{8.36}$$

Here D is a diagonal matrix and ε a small perturbation (no summation over repeated indices). Defining

$$\mathcal{D}_i \equiv \frac{A_i + E_i}{a_i^2} + H_i \left(\frac{4F_i}{a_i} - \frac{3\dot{B}_i}{N_i}\right) + \frac{2\dot{F}_i}{a_i N_i} - \frac{1}{N_i^2}\left(\ddot{B}_i - \frac{\dot{N}_i \dot{B}_i}{N_i}\right), \tag{8.37}$$

the equations of motion in the g-sector become (here we specialize to pressureless dust as matter source)

$$3H_g\left(H_gE_g-\frac{\dot{A}_g}{N_g}\right)-k^2\left(\frac{A_g}{a_g^2}+\frac{2H_gF_g}{a_g}-\frac{H_g\dot{B}_g}{N_g}\right)+\frac{m^2rP}{2}\left[3\left(A_f-A_g\right)-k^2\left(B_f-B_g\right)\right]=-\frac{\rho\delta}{M_g^2}, \tag{8.38}$$

$$\frac{\dot{A}_g}{N_g^2}-\frac{H_gE_g}{N_g}+\frac{m^2Pa_f}{\left(x+r\right)N_g}\left(xF_f-rF_g\right)=-\frac{\rho}{M_g^2}\left(\frac{\theta}{k^2}-F_g\right), \tag{8.39}$$

$$\frac{1}{N_g^3}\left[\left(-3H_gN_g^2+\dot{N}_g\right)\dot{A}_g+H_gN_g^2\dot{E}_g+\left(3H_g^2N_g+2\dot{H}_g\right)N_g^2E_g-N_g\ddot{A}_g\right]+m^2\left[\frac{Px}{2}\left(E_f-E_g\right)+Qr\left(A_f-A_g\right)\right]=0, \tag{8.40}$$

$$-\frac{\mathcal{D}_g}{2}-\frac{m^2rQ}{2}\left(B_f-B_g\right)=0. \tag{8.41}$$

In the f-sector, they become

$$3H_f\left(H_fE_f-\frac{\dot{A}_f}{N_f}\right)-k^2\left(\frac{A_f}{a_f^2}+\frac{2H_fF_f}{a_f}-\frac{H_f\dot{B}_f}{N_f}\right)-\frac{m^2P}{2r^3}\left[3\left(A_f-A_g\right)-k^2\left(B_f-B_g\right)\right]=0, \tag{8.42}$$

$$\frac{\dot{A}_f}{N_f^2}-\frac{H_fE_f}{N_f}+\frac{m^2Pa_g}{r^2N_f\left(x+r\right)}\left(rF_g-xF_f\right)=0. \tag{8.43}$$

$$\frac{1}{N_f^3}\left[\left(-3H_fN_f^2+\dot{N}_f\right)\dot{A}_f+H_fN_f^2\dot{E}_f+\left(3H_f^2N_f+2\dot{H}_f\right)N_f^2E_f-N_f\ddot{A}_f\right]-\frac{m^2}{xr^2}\left[\frac{P}{2}\left(E_f-E_g\right)+Q\left(A_f-A_g\right)\right]=0, \tag{8.44}$$

$$-\frac{\mathcal{D}_f}{2}+\frac{m^2Q}{2xr^2}\left(B_f-B_g\right)=0. \tag{8.45}$$

Here

$$P \equiv \beta_1 + 2\beta_2 r + \beta_3 r^2, \tag{8.46}$$
$$Q \equiv \beta_1 + (x + r)\,\beta_2 + x r \beta_3, \tag{8.47}$$
$$x \equiv \frac{N_f}{N_g}. \tag{8.48}$$

Due to the complexity of the equations it is not possible to solve them analytically. There exists certain approximations, however, that can be employed in their study. In Paper II, the equations were studied in the vicinity of a de Sitter background, since the equations can be solved exactly in such a background. This is described in section 8.5. The perturbations can also be studied in the early, radiation dominated, regime [82, 83]. Another possibility is to use the quasi-static approximation, in which the fields are assumed to vary at the same rate as the background. This approach was employed in [84, 85, 86]. Superhorizon modes were studied in [87]. A stability analysis was performed in [88]. Finally, one can solve the equations numerically, which was done in [84, 89]. This is used in Paper VI to study the integrated Sachs-Wolfe effect, and is described in chapter 9.

When working with the perturbations we have been using two different time coordinates. In the next section, we will study the behaviour of perturbations close to the de Sitter regime in cosmic time, i.e. $N_g = 1$ is employed. In chapter 9, where we study the ISW effect, we will be using conformal time, i.e $N_g = a_g$. In both cases we will be using gauge invariant variables, and their definition differs only in the choice of time coordinate.

Under a coordinate transformation the fields in the g and f sector will transform as:

$$\Delta A_i = -2\frac{\dot{a}_i}{a_i}\epsilon^0, \qquad \Delta B_i = -2\epsilon, \tag{8.49}$$
$$\Delta E_i = -2N_i^{-1}\partial_t\left(N_i\epsilon^0\right), \qquad \Delta F_i = \frac{N_i}{a_i}\epsilon^0 - \frac{a_i}{N_i}\dot{\epsilon}, \tag{8.50}$$

where $i = \{g, f\}$. In Paper II, the following gauge invariant variables were used:

$$\begin{aligned}
\Psi_g &= -\tfrac{1}{2}A_g + \dot{a}_g\left[\tfrac{1}{2}a\dot{B}_g - F_g\right], &\quad \Psi_f &= -\tfrac{1}{2}A_f + \tfrac{\dot{a}_f}{N_f}\left[\tfrac{a_f}{2N_f}\dot{B}_f - F_f\right],\\
\Phi_g &= \tfrac{1}{2}E_g - \partial_t\left(\tfrac{1}{2}a_g^2\dot{B}_g - a_g F_g\right), &\quad \Phi_f &= \tfrac{1}{2}E_f - \tfrac{1}{N_f}\partial_t\left(\tfrac{a_f^2}{2N_f}\dot{B}_f - a_f F_f\right),
\end{aligned}$$

$$\begin{array}{rcl} \mathcal{F} & = & F_f - \frac{a_g N_f}{a_f} F_g + \frac{1}{2}\left[\frac{N_f a_g^2}{a_f}\dot{B}_g - \frac{a_f}{N_f}\dot{B}_f\right], \\ \mathcal{B} & = & \frac{1}{2}\left(B_f - B_g\right). \end{array} \tag{8.51}$$

These variables have the advantage of being close to the commonly used gauge invariant variables Ψ and Φ, defined in eqs. (8.29) and (8.30), so that Ψ_g and Φ_g can be directly compared to the prediction of general relativity.

8.5 Leaving de Sitter

The goal of Paper II was to study how much the growth of structure changed in massive bigravity compared to general relativity, and to isolate the effects of the new scalar degree of freedom. We saw in section 4.1 that general relativity, as a theory of a massless spin-2 field, has two propagating degrees of freedom, which correspond to the helicity ± 2 modes. A massive spin-2 field has 5 degrees of freedom, which correspond to the helicity ± 2, ± 1 and 0 modes. From a particle point of view, helicities are related to the representations of the Poincaré group. In the geometrical framework of general relativity, helicity is defined in two steps: first one decomposes the metric into scalar, vector and tensor components, described in section 8.1. Then one identifies the components of the scalar perturbation that gives a canonical normalized kinetic term or, equivalently, a Klein-Gordon equation of motion. This will correspond to the helicity-0 mode, and a similar procedure identifies the ± 1 and ± 2 modes. A helicity analysis for de Sitter in dRGT, with a fixed reference metric, was performed in [90, 91].

The main framework in employed Paper II for doing the analysis of the growth of structure and helicity-0 is to start in a de Sitter regime, where the theory can be solved analytically, and then introduce a small deviation at the background level from de Sitter.

Massive and massless modes in de Sitter By defining the following combinations of the perturbations

$$\Phi_+ = \Phi_g + c^2\Phi_f, \qquad \Phi_- = \Phi_g - \Phi_f, \tag{8.52}$$

$$\Psi_+ = \Psi_g + c^2\Psi_f, \qquad \Psi_- = \Psi_g - \Psi_f, \tag{8.53}$$

where c is the proportionality factor between the two metrics, it is possible to decouple the massive and massless sector from each other (as first shown in [43]). The massless fields are then solved by

$$\Psi_+ = \Phi_+ = \frac{C_1}{c_2}a^{-1} + \frac{C_2}{c_2}a^{-3}, \tag{8.54}$$

just as in general relativity. The constant c_2 gives the overall normalization of the scale factor, as a function of cosmic time:

$$a(t) = c_2 \exp\left(H_{\mathrm{dS}}t\right), \qquad c_2 \equiv \left(\frac{1-\Omega_\Lambda}{4\Omega_\Lambda}\right)^{1/3}. \tag{8.55}$$

The energy density and velocity are given by

$$\frac{\delta\rho}{M_g^2} = -\left(12C_2H_{dS}^2 + 2k^2C_1\right)\frac{a^{-3}}{c_2^3} - 2C_2k^2\frac{a^{-5}}{c_2^5}, \tag{8.56}$$

$$\frac{\delta u}{M_g^2} = 4C_2H_{dS}\frac{a^{-3}}{c_2^3}. \tag{8.57}$$

Here δu is defined directly from the gauge-invariant stress-energy as $\delta T^0_i \equiv \partial_i \delta u$. In the singly coupling setup that we are working with, matter will be a source to both the massive and massless sector. The peculiar situation is, however, that the evolution of the matter sources is governed by the massless sector, and the massive sector thus has a fixed source term. The helicity-0 mode is given by the following combination:

$$\Pi \equiv \frac{\Phi_- + \Psi_-}{a^2} \tag{8.58}$$

It satisfies a massive Klein-Gordon equation

$$\left(\Box - M^2\right)\Pi = -J, \tag{8.59}$$

with source (a dot here signifies a derivative with respect to cosmic time)

$$J = \frac{1}{M_g^2}\left(-\delta p - \delta\rho + 2H_{dS}\delta u - 2\delta\dot{u} + \frac{1}{3}\nabla^2\chi - H_{dS}a^2\dot{\chi} - a^2\ddot{\chi}\right). \tag{8.60}$$

Here χ is the anisotropic stress defined through

$$\delta T^i_j = \partial_i\partial_j\chi, \tag{8.61}$$

the $\Box$ operator is given by

$$\Box = -\frac{\partial}{\partial t^2} - 3H_{dS}\frac{\partial}{\partial t} + \frac{\nabla^2}{a^2}, \tag{8.62}$$

and M is the mass of the helicity-0 mode, given by

$$M^2 = m^2\left(c + c^{-1}\right)\left(\beta_1 + 2\beta_2 c + \beta_3 c^2\right). \tag{8.63}$$

The fields $\mathcal{F}$ and $\mathcal{B}$ can be solved in terms of the other fields.

Solutions to eq. (8.59) are well-known in terms of Bessel and Lommel functions. Π will stay finite at all times, but the perturbations fields Φ_- and Ψ_-, which are proportional to $a^2\Pi$, will diverge in the future, unless

$$M^2 < 2H_{\rm dS}^2. \tag{8.64}$$

The parameter choice $M^2 = 2H_{\rm dS}^2$ is known as the Higuchi bound. This is a stability bound which relates the mass of the graviton to the cosmological constant in de Sitter. The bound is satisfied if $M^2 > 2H_{\rm dS}^2$. It was first discussed [92, 93, 94] and later investigated in more detail by Higuchi in [95, 96, 97] and Deser and Waldron in [98, 99, 100]. The bound gives unitarity conditions on the massive graviton upon quantization; if the bound is violated, which means that $M^2 < 2H_{\rm dS}^2$, the kinetic energy of the graviton becomes negative, which upon quantization leads to negative probabilities. Furthermore, at the bound, when $M^2 = 2H_{\rm dS}^2$, the helicity-0 mode vanishes due to a new gauge symmetry.

The parameter M is formed out of the β_i parameters, and the β_i parameters also govern the background expansion. The Higuchi bound can therefore be related to consistency conditions for the background. The model under consideration in Paper II has $\beta_1 = \beta_3 = 0$. The background expansion can be written as

$$H^2 = \frac{\rho}{3M_P^2} + H_{\rm dS}^2. \tag{8.65}$$

The effective Planck mass M_P, which does not have to be the same as that measured by local experiments, can be expressed in terms of $H_{\rm dS}$ and M, defined in (8.63), as

$$M_P^2 = M_g^2\left(1 + c^2\right)\frac{M^2 - 2H_{\rm dS}^2}{M^2 - 2(1+c^2)H_{\rm dS}^2}. \tag{8.66}$$

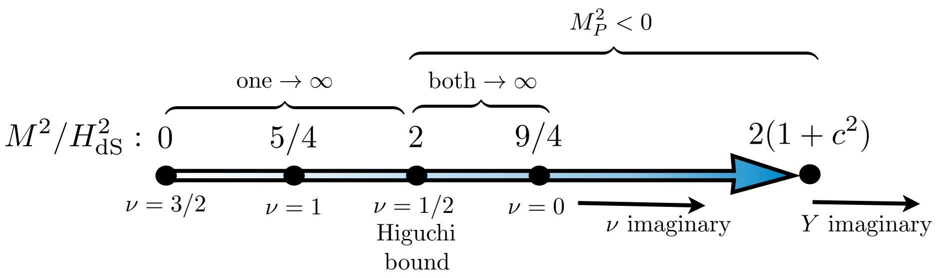

Figure 8.1. Effective Planck mass, a_f (denoted by Y in Paper II) and the helicity-0 mode as functions of the mass of the helicity-0 mode in units of $H_{\rm dS}^2$ (also shown is ν, defined in (8.68). Below the Higuchi bound, one of the solutions to the helicity-0 mode gives divergences in the metric. Above the Higuchi bound, both solutions do, and the Planck mass becomes imaginary. For $M^2/H_{\rm dS}^2 > 2(1+c^2)$, also the scale factor in the f-metric becomes imaginary.

It can be shown that a_f becomes imaginary in the early universe if $2(1+c^2)H_{\rm dS}^2 < M^2$. To avoid this, and also to avoid having a negative M_P^2, one must thus impose

$$M^2 < 2H_{\rm dS}^2. \tag{8.67}$$

We thus have a rather peculiar situation, where the Higuchi bound has to violated in order for Φ_- and Ψ_- to remain finite. Furthermore, it has to be violated in order to avoid imaginary quantities in the background. The relationship between the Higuchi bound and massive bigravity was also studied in [72, 101, 102, 89]. In Fig. 8.1 we show schematically the relationship between the background effective Planck mass, Higuchi bound and behaviour of the background solution as a function of $M^2/H_{\rm dS}^2$. The parameter ν, defined as

$$\nu^2 = \frac{9}{4} - \frac{M^2}{H_{\rm dS}^2}, \tag{8.68}$$

appears in the solution to the massive Klein Gordon equation (8.59).

Growth index Calculating the evolution of the perturbative fields in the vicinity of a de Sitter regime shows that even though the background is identical to a ΛCDM scenario, the behaviour of the perturbations differ from that predicted by general relativity. The approach employed in Paper II was to do an expansion in the parameter

$$\tilde{\epsilon} \equiv -\frac{\dot{H}}{H^2}. \tag{8.69}$$

In de Sitter, $\tilde{\epsilon} = 0$, and as one moves away from de Sitter into the matter dominated regime $\tilde{\epsilon}$ is growing. The perturbation equations can, for small $\tilde{\epsilon}$, be solved iteratively by using the analytic de Sitter solution. Writing the perturbation equations schematically as $D_{mn}\phi^n = \mathcal{J}_m$, we can expand the differential operators D_{mn}, fields ϕ^m and sources $\mathcal{J}_m$ as

$$D_{mn} = D^0_{mn} + \epsilon D^{\epsilon}_{mn}, \qquad \phi^m = \phi^m_0 + \epsilon\phi^m_{\epsilon}, \qquad \mathcal{J}_m = \mathcal{J}^0_m + \epsilon\mathcal{J}^{\epsilon}_m. \tag{8.70}$$

Here $\epsilon = 6\exp(-3H_{\mathrm{dS}}t)$, which turns out to be the relevant expansion parameter. Unfortunately, the quasi-de Sitter approach can not be extended to the current time, since $\epsilon \sim 0.5$ today. The perturbative expansion would thus break down at this point. Before that point, it was, however, shown in Paper II that even though solutions of the perturbations are the same as those in general relativity in de Sitter, there is indeed a difference as one moves into the quasi-de Sitter regime. This shows that even if the background expansion is identitical to general relativity, the perturbations does not have to be the same.

The evolution of the perturbations in the sub-horizon limit and with comparison with data were performed in [84, 86, 85], using the quasi-static approximation for subhorizon modes. This means that the fields are assumed to vary at the same rate as the background, so that $\Phi \sim H\dot{\Phi} \sim H^2\ddot{\Phi}$. In order to compare the predictions of the theory with general relativity, one often uses the following parameters:

$$\eta = \frac{\Phi}{\Psi} \tag{8.71}$$

and

$$Y = -\frac{2k^2a^2\Phi}{3H^2\Omega_m\delta_m}. \tag{8.72}$$

For general relativity, we have $\eta = Y = 1$. In Fig. 8.2, we show η and Y in for the model where only β_1 and β_4 non-zero. This model is used in Paper IV, and is described in more detail in section 9.2. The deviations of η and Y from the predictions of general relativity can clearly be seen. Current bounds on η and Y from observations are not stringent enough, however, to favor one model over another [85, 86].

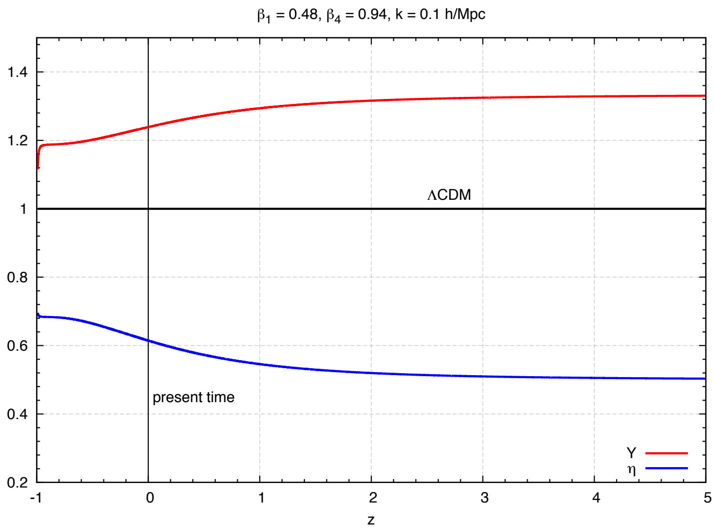

Figure 8.2. Evolution of η and Y, defined in (8.71) and (8.72), in the IBB model. The black horizontal line corresponds to the ΛCDM prediction, and the black vertical line gives the time today. Plot taken from [86].

8.6 On instabilities and oscillations

Paper II dealt with the properties of perturbations in massive bigravity in the late universe. The behaviour of the perturbations in the early universe turns out to be dramatic. In [83, 82] it was shown that there exists exponential instabilities in the radiation dominated regime, which very rapidly makes the fields become non-linear. The non-linear behaviour could be cured by a cosmological Vainshtein mechanism, or it could signal that the theory simply predicts too much structure in the early universe. The instabilities also depend on the type of matter coupling employed [103].

The nature of the instabilities were further studied in [84, 86, 89, 103, 104], and [86] showed that the parameter choice where only β_1 and β_4 are non-vanishing can lead to viable cosmologies at both the background and perturbative level (but it is claimed in [89] that higher order perturbations could be unstable for this model). This model was dubbed Infinite Branch Bigravity (IBB), since r goes to infinity in the early universe. The best-fit parameters for IBB derived in [86] using expansion histories and structure formation are $\beta_1 = 0.48^{+0.05}_{-0.16}$ and $\beta_4 = 0.94^{+0.11}_{-0.51}$. The full numerical solution to the system also showed that the fields are characterized by oscillatory behaviour around a mean value that is close to that of general relativity.

The overall conclusion concerning the perturbations in massive bigravity can be summarized as follows:

- The behaviour of the scalar perturbations differ from the case of general relativity, even if the background expansion is identical.
- There is a strong link between the Higuchi bound and the background expansion, which could invalidate certain parameter ranges.
- There exists early time instabilities for a large range of the parameter space. These instabilities could be cured by non-linear terms, or they could signal that too much structure is predicted.
- Certain parameter ranges give a viable alternative to general relativity, such as the case when only β_1 and β_4 is non-vanishing. It still has to be studied, however, whether the oscillations present in the solution for the perturbations fields leads to observational signatures.

Chapter 9

Integrated Sachs-Wolfe effect

In the two previous chapters we described how massive bigravity affects cosmic expansion histories and growth of structure in the universe. The key issue there is the relationship between the matter content and spacetime curvature. In this and the following chapter we will study how light propagates in a curved spacetime, and how this can be used to put constraints on massive bigravity.

As described in section 7.1 and 8.3, the cosmic microwave background (CMB) is one of the primary objects for observational cosmology. The improved sensitivity in measuring the CMB is shown in Fig. 9.1. In this chapter we will see how one can use observations of the CMB, together with observations of galaxies and Active Galactic Nuclei (AGN), to observationally distinguish different theories of gravity. The main aspect of this approach is the time evolution of the gravitational potentials, since they affect both the form of the CMB and the galaxy and AGN distribution. The alteration of the CMB due to the time evolution of the gravitational potentials is called the Integrated Sachs-Wolfe (ISW) effect.

9.1 Cosmic light, now and then

The primordial photons that we today observe as the CMB are primarily affected by four things: 1) the photon distribution at the time of decoupling, 2) the relative motion between us and the last scattering surface, 3) the gravitational potentials at the time of decoupling and 4) the change in the gravitational potential between decoupling and today. In terms of the temperature distribution of the photons, these four effects are given by

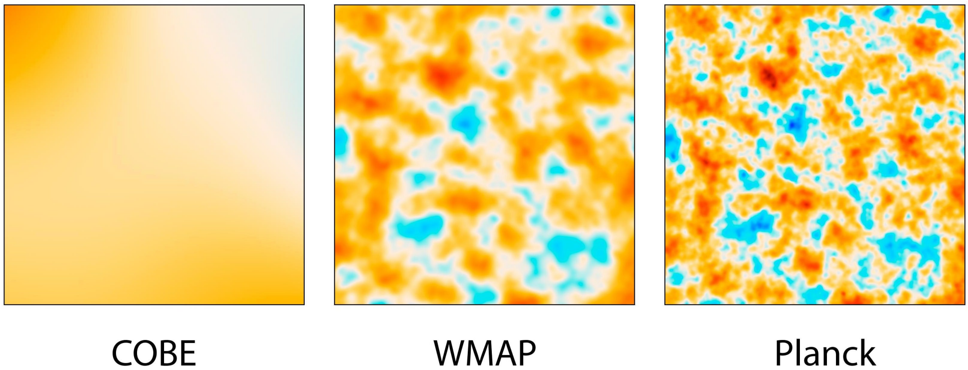

Figure 9.1. The three images show temperature fluctuations of the cosmic microwave background in a 10 square degree patch of the sky for three different satellite experiments. The COBE satellite was launched in 1989, and was the first instrument to give a full-sky map of the temperature anisotropies. The WMAP satellite was launched in 2001 and provided data until 2012. It was instrumental in the establishment of the standard model of cosmology. The Planck satellite was launched in 2009 and released its first data set in 2013, which confirmed and sharpened the results of the earlier experiments. Image courtesy: NASA/JPL-Caltech/ESA.

$$\frac{\delta T}{T_0}(\mathbf{n},\eta_0) = \frac{1}{4}\delta_\gamma(\eta_r) + \mathbf{n}\cdot\mathbf{v}(\eta_r) + \Psi_g(\eta_r) + \int_{\eta_r}^{\eta_0}\left(\Psi_g' + \Phi_g'\right)d\eta. \quad (9.1)$$

Here $\mathbf{n}$ is a unit vector pointing in the direction of observation and $\mathbf{v}$ is the relative motion between us and the last scattering surface. The third and fourth terms are responsible for the Sachs-Wolfe and Integrated Sachs-Wolfe effects [105]. The temperature fluctuation is usually represented by doing an expansion in spherical harmonics:

$$\frac{\delta T(\mathbf{n})}{T_0} = \sum_{l=1}^{\infty}\sum_{m=-l}^{m=l} a_{lm}Y_{lm}(\mathbf{n}) \quad (9.2)$$

This formula can be inverted so that the coefficient a_{lm} are determined from $\delta T/T_0$. From these coefficients one can define the angular power spectrum C_ℓ as

$$\langle a_{lm}a^*_{l'm'}\rangle = C_\ell\delta_{ll'}\delta_{mm'} \quad (9.3)$$

Here the star signifies a complex conjugate. The average is supposed to be taken over an ensemble of universes; since this is impossible in practice, one usually replaces the ensemble average by a volume average. This means that one assumes that different regions of the universe are uncorrelated, and looking at the properties of the universe in two different directions is equivalent to looking at two different realisations in an ensemble of universes. Furthermore, if the temperature fluctuations satisfy the properties of a Gaussian random field (which means that its Fourier coefficients have a Gaussian distribution and are uncorrelated with one another), C_ℓ characterize the temperature fluctuations completely.

C_ℓ has today been measured up to $\ell \sim 2500$. It is customary to define

$$\mathcal{D}_\ell \equiv T_0^2 \frac{l(l+1)}{2\pi} C_\ell. \tag{9.4}$$

$\mathcal{D}_\ell$ gives the squared amplitude of the temperature fluctuations in a decimal interval of multipoles. Furthermore, for a nearly flat initial power spectrum of the potential fluctuations—which is predicted by e.g. inflation—$\mathcal{D}_\ell$ is constant at low l in the absence of the ISW effect. In Fig. 9.2 we show $\mathcal{D}_\ell$ as measured by the Planck satellite [62]. The main feature in the spectrum is the series of oscillations, with the largest peak at $\ell \sim 200$, which are caused by the oscillations in the baryon-photon plasma prior to decoupling.

Fig. 9.2 also shows the best-fit ΛCDM model as a solid line. To compute this one needs to rely on numerical routines, since the physics of the baryon-photon plasma is rather complex. The contribution of the ISW effect to the CMB is more straightforward, however, since it only involves the time evolution of the gravitational potentials [106]. By defining the weight functions

$$K_\ell^{\dot{\Phi}}(k) = \int dz \frac{d}{dz}\left[\frac{\Phi(z)}{\Phi(0)}\right] j_\ell\left[k\chi(z)\right], \tag{9.5}$$

where j_ℓ are the spherical Bessel functions and $\chi(z) = \eta_0 - \eta(z)$, the power spectrum of the ISW effect can be derived as

$$C_\ell^{ISW} = 4\pi \int \frac{dk}{k} \Delta_\Phi^2(k) K_\ell^{\dot{\Phi}}(k) K_\ell^{\dot{\Phi}}(k). \tag{9.6}$$

The power spectrum $\Delta_\Phi^2(k)$ of the gravitational potential is given by

$$\Delta_\Phi^2(k) = 9\Omega_m^2 \delta_H^2 \left(\frac{k}{H_0}\right)^{n-1} T^2(k). \tag{9.7}$$

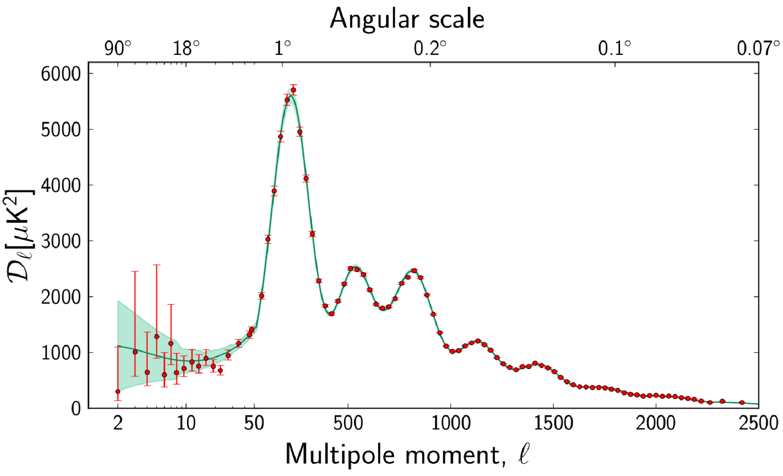

Figure 9.2. Angular power spectrum of the cosmic microwave background, as measured by the Planck satellite. $\mathcal{D}_\ell$ is defined in eq. (9.4). The late time integrated Sachs-Wolfe effect concerns the lowest multipole, $l \sim 2 - 10$, where cosmic variance and measurement uncertainties are largest. Plot taken from [62].

Here δ_H is the amplitude of the matter density contrast at horizon scales today, n is the spectral tilt which depends on the primordial conditions (e.g. inflation) and $T(k)$ is the transfer function which takes the dynamics from the primordial initial conditions up to the time of decoupling into account. In Paper VI the so-called BBKS approximation for the transfer function was used [107]. $\Delta^2_\Phi(k)$ is related to the power spectrum of matter, which will be used in the analysis in the next section, through the cosmological Poisson equation:

$$\Delta^2_m(k) = \frac{1}{9\Omega^2_m}\left(\frac{k}{H_0}\right)^4 \Delta^2_\Phi(k) \tag{9.8}$$

The ISW effect itself is subdominant as compared to the other terms in eq. (9.1). The ordinary Sachs-Wolfe effect has an amplitude on the order of $\sim 1000\ \mu\text{K}^2$, whereas the ISW effect only has an amplitude of $\sim 100\ \mu\text{K}^2$. One therefore needs to cross-correlate measurements of the CMB with other probes of the gravitational potentials.

9.2 Cross-correlating the ISW effect

Galaxies trace the underlying dark matter distribution, which in turn is related to the gravitational potential. An overdensity of galaxies in a certain direction thus corresponds to a stronger gravitational potential, and hence a larger time evolution of the potential. This also affects the photons propagating from that direction, and one can thus cross-correlate the galaxy distribution with the CMB, a method first developed in [108]. The angular power spectrum of this cross-correlation between galaxies and the ISW is given by (we are here extending the framework presented in [109])

$$C_\ell^{Tg} = C_\ell^{\dot{\Phi}g} = 4\pi \int \frac{dk}{k} \Delta_m^2(k) K_\ell^{\dot{\Phi}}(k) K_\ell^g(k). \tag{9.9}$$

The weight functions $K_\ell^{\dot{\Phi}}$ and K_ℓ^g are given by

$$K_\ell^g(k) = \int dz b(z) \frac{dN}{dz} \frac{\delta(k,z)}{\delta(k,0)} j_\ell[k\chi(z)], \tag{9.10}$$

$$K_\ell^{\dot{\Phi}}(k) = \int dz \frac{d}{dz}\left[\frac{\Phi+\Psi}{\delta(k,0)}\right] j_\ell[k\chi(z)]. \tag{9.11}$$

Here $b(z)$ is the bias function relating the galaxy or AGN tracers to the underlying matter distribution and dN/dz is the distribution function of galaxies, normalized so that $\int dz' dN/dz' = 1$. The weight functions are thus normalized against the matter distribution today, and we thus use $\Delta_m^2(k)$, given in eq. (9.8), in the angular power spectrum. An alternative approach would be to normalize against the gravitational potentials at decoupling. In this case, one would replace $\delta(k,0)$ with the gravitational potentials at decoupling, and use $\Delta_\Phi^2(k)$ together with a correction factor which takes the change of the potential between decoupling and today into account.

The predictions of general relativity for C_ℓ^{Tg} can be computed given a power spectrum of the density contrast or potential, defined either today or at the time of decoupling. In massive bigravity, three main aspects differ from general relativity: Φ is not equal to Ψ, δ does not have the same relationship to the gravitational potentials, and the time evolution of both δ and the potentials depend on k.

In Paper VI the IBB model, which has only β_1 and β_4 non vanishing, was tested using the cross-correlation of the ISW effect with galaxies and AGN. The best-fit parameters $\beta_1 = 0.48$ and $\beta_4 = 0.94$ was used, and it

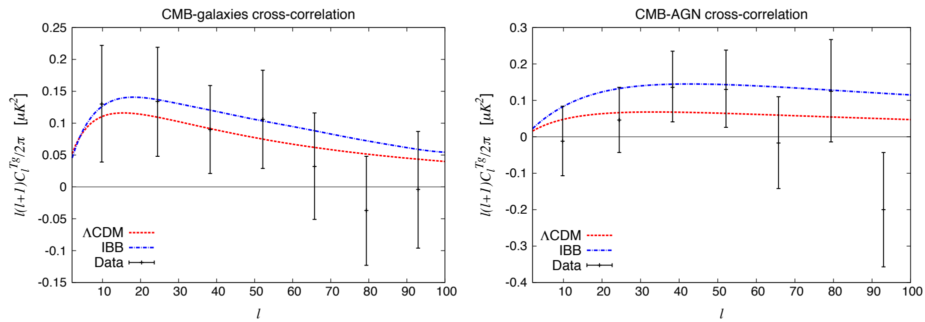

Figure 9.3. Angular power spectrum of the cross-correlation between the ISW effect and galaxies (left panel) and AGN (right panel) computed for both general relativity (ΛCDM) and massive bigravity (IBB). Data points taken from [109] and are based on the WISE survey [110] and WMAP temperature map [112]. The predictions of massive bigravity for the IBB model are larger than ΛCDM by roughly a factor of 1.5, but still consistent with data.

was assumed that $\Phi = \Psi$ at the time of decoupling (which mimics the initial conditions of general relativity).[1] The analysis was a continuation of the test of general relativity against this specific cross-correlation, which was performed in [109]. This analysis used galaxies and quasars provided by the *Wide-field Infrared Survey Explorer* (WISE), presented in [110] and further analysed in [111]. This catalogue contains more than 500 million sources distributed over the entire sky, which makes it an ideal candidate catalogue when performing a cross-correlation analysis of the CMB, using the WMAP temperature map [112]. [109] claimed a 3σ detection of the ISW effect through the cross-correlation analysis. The amplitude $\mathcal{A}$ of the detection is normalized against the expected ΛCDM correlation, and was measured to be $\mathcal{A} = 1.24 \pm 0.47$ for the galaxies and $\mathcal{A} = 0.88 \pm 0.74$ for the AGN.

A crucial step in the cross-correlation is the estimate of the bias factor, relating the galaxies and quasars to the underlying dark matter distribution and thus the gravitational potential. In [109] the bias factor was estimated using the lensing that the tracers induced on the CMB. This is possible since lensing directly probes the gravitational potentials. For galaxies they derived two possible bias factors: a constant bias $b^G = 1.41 \pm 0.15$ and a redshift dependent bias $b^G(z) = b_0^G(1+z)$, with $b_0^G = 0.98 \pm 0.10$.

[1]The numerics was solved by first forming two coupled, second order differential equations for Φ and Ψ, which could then be used to solve for the gauge invariant density contrast. This approach was first employed in [84].

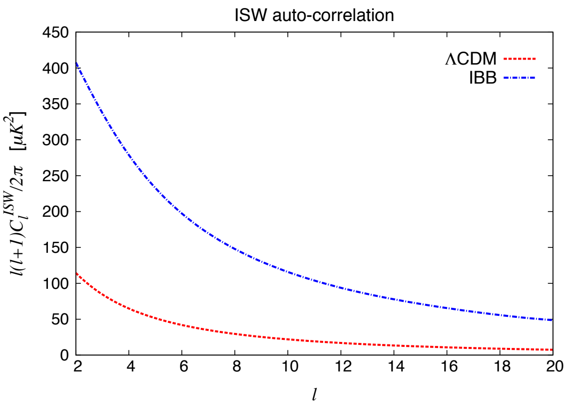

Figure 9.4. The ISW auto-correlation, defined in (9.6), as predicted by ΛCDM and IBB. Just as with the case of the cross-correlation of the ISW effect with galaxies or AGN, IBB predicts a larger amplitude as compared to ΛCDM.

For the AGN the bias was estimated as $b^A(z) = b_0^A[0.53 + 0.289(1+z)^2]$, with $b_0^A = 1.26 \pm 0.23$.

In principle, the bias factor should be independently analysed in the IBB model, using the lensing of the CMB. Such an analysis would involve a complete determination of the full CMB spectrum in massive bigravity, which is a rather daunting task. Paper VI therefore used the inferred bias of [109], assuming that lensing properties around massive objects are similar to those of general relativity for the parameter choices under consideration (this assumption is further discussed in the section 10.1).

C_ℓ^{Tg} for galaxies and AGN is shown in Fig. 9.3. The data points are imported from [109], and the predictions for both general relativity and IBB are shown as red and blue lines. The prediction of general relativity is calculated using $\Omega_m = 0.3$. The same power spectrum is used for ΛCDM and IBB, with the density contrast today at horizon scales put to $\delta_H = 7.5 \times 10^{-5}$, and using a flat initial spectrum, i.e. $n = 1$. IBB predicts a higher amplitude as compared to ΛCDM by roughly a factor of 1.5, but this is still consistent with data. Using the gravitational potential at decoupling as normalization instead of the density contrast today slightly varies the prediction of IBB. For the galaxies, there is also a slight variation in C^{Tg} when using the linear, rather than the constant, bias factor. These variations are, however, well within the data uncertainties. The auto-correlation function for the ISW effect itself is shown in Fig. 9.4. Just as for the cross-correlation, the IBB model predicts a higher amplitude as compared to ΛCDM. As stated in the introduction to this chapter, it is, however, not possible to directly measure the auto-correlation, since it is subdominant as compared to the ordinary Sachs-Wolfe effect.

The framework laid down in Paper VI for analysing the cross-correlation of the ISW with galaxies and AGN can be extended in several ways:

- Using the lensing-ISW cross-correlation allows for a direct correlation of the potentials without including the bias factor. This requires, however, a better understanding of the relationship between lensing of the CMB by intermediate sources.

- The results of Paper VI shows that the normalization scheme used matters, although it only gives rise to slight variations. It is only by analyzing multi-component perturbations prior to decoupling that the correct normalization scheme can be employed.

- Having solved that system, one can also predict the form of the CMB and the matter spectrum. This is of course the ultimate test of massive bigravity and the IBB model.

Chapter 10

Lensing

Given a modified theory of gravity there are two immediate tasks to perform: First, assume homogenous and isotropic spatial slices and then investigate the time evolution of the theory. This corresponds to looking at cosmic expansion histories, described in chapter 7. Secondly, one can assume static (i.e. time-independent) manifolds and study the spherically symmetric solutions. This corresponds to looking at black hole and star solutions, which is the subject of this chapter. Roughly speaking, then, the first approach correspond to isolating the time behaviour of the theory, and the second approach corresponds to isolating the spatial, or radial, behaviour of the theory.

Spherically symmetric solutions can be observationally probed in two ways. First of all, they will affect planetary orbits and thus modify the Keplerian dynamics that hold in the solar system. Secondly, light bending gets modified. Already in Newton's theory of gravity light bending is possible, due to the fact that the mass of a test-particle is irrelevant for its motion. This holds, in particular, for a massless particle. In Einstein's theory, light bending is due to the geodesic motion of light in a gravitational field, and its value is famously a factor of two larger than that predicted in Newton's theory. For a modified theory of gravity the amount of light bending could be lower or higher than that predicted by general relativity, and can thus be used as a observational tool to distinguish different theories of gravity.

In order to observationally test the amount of light bending a given massive object produces, one needs to know the mass of the object through an independent measurement. This can be achieved through Keplerian dynamics, i.e. observing the motion of planets around a star in the case of the bending of light that passes close to the surface of the star, or observing the motion of stars and gas in a galaxy in the case of the bending of light by galaxies.

The first study of spherically symmetric vacuum solutions in massive bigravity appeared in [113], which presented linearized solutions, and [114], which did a numerical study of the complete system. Further studies investigated the Vainshtein mechanism [115, 116], rotating black holes [117], charged black holes [118], stability issues [119, 120, 121] and uniqueness properties [122, 123].

10.1 Gravity of a spherical lump of mass

To derive vacuum solutions, we use the following spherically symmetric ansatz:

$$ds_g^2 = -V(r)^2 dt^2 + W(r)^2 \left(dr^2 + r^2 d\Omega^2\right), \tag{10.1}$$

$$ds_f^2 = -A(r)^2 dt^2 + B(r)^2 dr^2 + C(r)^2 r^2 d\Omega^2. \tag{10.2}$$

The form of ds_g^2 has been chosen to facilitate the lensing analysis. The equations of motion derived using this ansatz are too complicated to be solved analytically. For weak gravitational fields, we can, however, linearize the metric components in the following way:

$$V \simeq 1 + \delta V, \quad W \simeq 1 + \delta W, \tag{10.3}$$

$$A \simeq c\,(1 + \delta A)\,, \quad B \simeq c\,(1 + \delta B)\,, \quad C \simeq c\,(1 + \delta C)\,. \tag{10.4}$$

Here c is the proportionality constant for the background solution of the two metrics. In order to have asymptotic flatness, we must impose the following conditions on the β_i parameters and c:

$$\beta_0 + 3\beta_1 c + 3\beta_2 c^2 + \beta_3 c^3 = 0, \tag{10.5}$$

$$\beta_4 c^3 + 3\beta_3 c^2 + 3\beta_2 c + \beta_1 = 0, \tag{10.6}$$

Defining the effective mass

$$m_g^2 \equiv m^2 \left(c + c^{-1}\right) \left(\beta_1 + 2\beta_2 c + \beta_3 c^2\right), \tag{10.7}$$

the linearized solutions to the equations of motion become

$$\delta V = -\frac{GM_1}{r} - \frac{c^2 GM_2}{r} e^{-m_g r}, \tag{10.8}$$

$$\delta W = \frac{GM_1}{r} + \frac{c^2 GM_2}{2r} e^{-m_g r}, \tag{10.9}$$

$$\delta A = -\frac{GM_1}{r} + \frac{GM_2}{r} e^{-m_g r}, \tag{10.10}$$

$$\delta B = \frac{GM_1}{r} + \frac{GM_2 \left[2\left(1+c^2\right)\left(1+m_g r\right) + c^2 m_g^2 r^2\right]}{2 m_g^2 r^3} e^{-m_g r}, \tag{10.11}$$

$$\delta C = \frac{GM_1}{r} - \frac{GM_2 \left[\left(1+c^2\right)\left(1+m_g r\right) + m_g^2 r^2\right]}{2 m_g^2 r^3} e^{-m_g r}, \tag{10.12}$$

M_1 and M_2 are integration constants. We see that the solution contain two components: the massless components which go as $1/r$, and the massive, Yukawa components which go as $\exp(-m_g r)/r$. Defining $\lambda_g \equiv 1/m_g$, we see that the massive component vanishes when $r \gg \lambda_g$.

In general relativity, the linearized solutions start to become invalid as r approaches the Schwarzschild radius. In massive gravity the Vainshtein radius makes the linearized solutions invalid already for a radius which is very large compared to the Schwarzschild radius. This radius can be identified by computing the second order solutions and investigate when they become comparable to the linearized solutions. In the limit $r \ll \lambda_g$, and imposing $\beta_2 = -c\beta_3$ the Vainshtein radius becomes

$$r_V = \left[\frac{GMc^2\left(1+c^{-2}\right)^2}{m_g^2}\right]^{1/3}. \tag{10.13}$$

Here we have used a point mass object as source, so that $M_1 \sim M_2 \sim M$. Putting $c = 1$ and $r_H = 1/H_0$, we have

$$r_V \simeq 0.17 \left[\left(\frac{M}{M_\odot}\right)\left(\frac{\lambda_g}{r_H}\right)^2\right]^{1/3} \text{kpc} \simeq 3.4 \cdot 10^{-8} \left[\left(\frac{M}{M_\odot}\right)\left(\frac{\lambda_g}{r_H}\right)^2\right]^{1/3} r_H, \tag{10.14}$$

When deriving the cross-over radius between the linear and non-linear regime in [30], Vainshtein also conjectured that higher order terms could make the solutions within the Vainshtein radius close to that of general

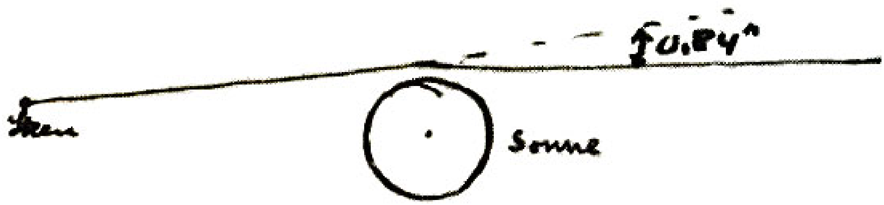

Figure 10.1. Arthur Eddington, who first measured the bending of light around the sun in 1919, made this handdrawing depicting the phenomenon. The value 0.84 arcseconds is the prediction of Newtonian gravity. Image credit: University of Louisville, Astronomy Education Wiki.

relativity. This could then produce a smooth limit to general relativity in the limit of a vanishing graviton mass, and thus remove the vDVZ discontinuity (described in section 4.2).

The existence of a Vainshtein mechanism in massive bigravity was first shown in [114, 116]. Although this was only presented in passing, and the details were not studied extensively, it means that one can assume with some confidence in this analysis that within the Vainshtein radius the prediction of the theory will be very close to that of general relativity.

10.2 Gravitational lensing

The bending of light as it passes a massive object follows from the geodesic motion in the gravitational field that the massive object sets up [124]. This is depicted in Fig. 10.1. If we write the metric as

$$ds_g^2 = -\left(1+2\Phi\right)dt^2 + \left(1-2\Psi\right)\left(dr^2 + r^2 d\Omega^2\right), \tag{10.15}$$

the velocity for massless particles (for which $ds_g^2 = 0$) is

$$v^2 = \frac{1+2\Phi}{1-2\Psi}. \tag{10.16}$$

The index of refraction, which describes the variation of the velocity of light in a medium as compared to the vacuum, is then, under the assumption that $\Phi, \Psi \ll 1$

$$n = 1/v = 1 - \Psi - \Phi. \tag{10.17}$$

By solving the geodesic equation, one can show that the deflection angle, i.e. the amount that the trajectory is shifted relative to the unperturbed trajectory, is

$$\vec{\alpha} = \int \vec{\nabla}_\perp \left(\Phi + \Psi\right) d\lambda. \tag{10.18}$$

Here λ is the parameter which describes the trajectory, and $\vec{\nabla}_\perp$ is the derivative perpendicular to the trajectory. Effectively, one can integrate over the unperturbed light path, since the deflection is small. For a point mass, we have, in general relativity, $\Phi = \Psi = -GM/r$, and

$$|\alpha_{GR}| = \frac{4GM}{b}, \tag{10.19}$$

where b is the impact parameter (i.e. the minimal distance to the lens for the unperturbed trajectory).

Point-particle solutions in massive bigravity are given by

$$\Phi = -\frac{GM}{r}(1 + \frac{4c^2}{3}e^{-m_g r}), \tag{10.20}$$

$$\Psi = -\frac{GM}{r}(1 + \frac{2c^2}{3}e^{-m_g r}), \tag{10.21}$$

Defining the lensing potential $\varphi = \left(\Phi + \Psi\right)/2$, the ratio between φ and Φ, which is the field relevant for trajectories of massive particles, is

$$\frac{\Phi}{\varphi} = \frac{1 + (4c^2)/3}{1 + c^2}, \tag{10.22}$$

as $r \ll \lambda_g$. In general relativity, this ratio is equal to unity. The factor of $4/3$ is the famous vDVZ-discontinuity, which is present in the case of massive gravity with a fixed reference metric. We see that for $c = 1$, the ratio is equal to $7/6$, and when $c \to \infty$ it becomes equal to the vDVZ value of $4/3$. When $c \to 0$, we get back the predictions of general relativity, since $\Phi \to \varphi$.

10.3 Bending light in the right way

By using objects for which the gravitational field can be probed both by massive objects and light, we can put constraints on the parameters of massive bigravity. In Paper III, elliptical galaxies, for which both the

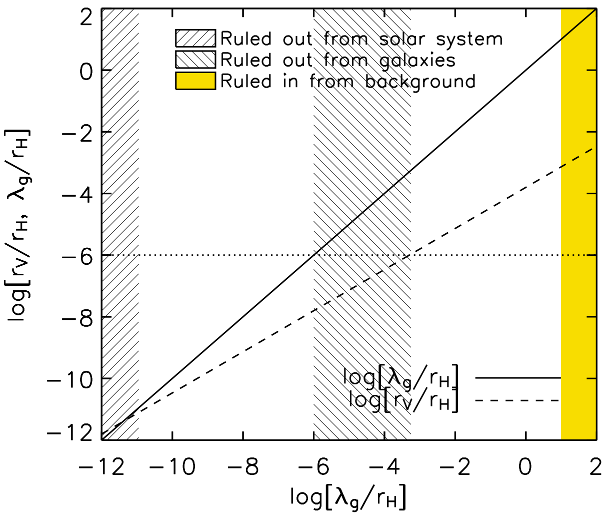

Figure 10.2. Allowed range for λ_g/r_H, where r_H is the Hubble scale today. The thick line is the Compton wavelength of the graviton, and the dashed line the Vainshtein radius for a galaxy. The dotted line corresponds to the galaxy scale, ~ 5 kpc. Values of λ_g where the galactic scale is larger than the Vainshtein radius and smaller than λ_g are excluded. Also, precis measurements in the solar system exclude values of λ_g less than roughly $10^{-11} r_H \sim 0.05$ pc. The other ranges for λ_g are allowed. In particular, the case favored by the cosmic expansion history, where $\lambda_g \sim r_H$ is viable.

velocity dispersion and deflection angle is known, were used. From the observed deflection angle the gravitational lensing potential was reconstructed for each galaxy in the dataset. This could then be used to predict the velocity dispersion for given parameter choices c and m_g (or, equivalently, λ_g), which in turn could be compared to data. The calculation of the velocity dispersion relies on the assumed matter distribution; this assumption, however, is not crucial for the parameter constraints, since both the lensed photons and velocity dispersion tracers are at a similar galactic radii.

Due to the Vainshtein mechanism one has to divide the constraints into three sets. At distances within the Vainshtein radius, the Vainshtein mechanism should ensure that the reconstruction of the potential from the deflection angles should follow along the lines of general relativity. In the region that is outside the Vainshtein radius, but within λ_g, the Yukawa terms produce deviations from general relativity. At distances outside both the Vainshtein radius and λ_g, the Yukawa terms are negligible and general relativity is once again restored. It is exactly in the parameter range where the Yukawa terms are relevant on galactic scales that we find that the theory is excluded.

In Fig. 10.2 we show the range of values that are allowed for λ_g. If one assumes that the Vainshtein mechanism does work, the lensing analysis excludes values of λ_g in the range $5\,\mathrm{kpc} \lesssim \lambda_g \lesssim 5\,\mathrm{Mpc}$. Constraints on Yukawa terms in the solar system [125], together with deflection and time delay of light in the vicinity of the sun [126], also constraints λ_g to be larger than 0.05 pc.

Part V

Summary and outlook

Summary and outlook

This thesis has explored the phenomenological consequences of a natural extension of general relativity, in which a graviton is endowed with a non-vanishing mass. Such an extension incorporates a second rank-2 tensor field, which could either be fixed or dynamical. The latter case, usually called the Hassan-Rosen theory, is the focus of all of the papers included in this thesis, with the exception of Paper V. Several possible couplings between matter and the two rank-2 tensor fields are possible, and two have been used in this work: one where matter only couples to one of the fields, and one where it couples to both in a symmetrical way. Let us now summarize our main results.

Expansion histories The cosmological expansion histories were studied in three of the papers included in this thesis. In Paper I matter coupled to one of the metrics, and the equations of motion for the background was derived and analysed. It was shown that there exists ΛCDM-like expansion histories. A similar analysis was performed in Paper IV, where matter coupled to an effective metric which was formed out of a combination of the two metrics. It was shown that this type of coupling gives expansion histories that are generically "closer" to ΛCDM than those of the singly coupling case. Paper V analysed the cosmology of the same type of effective matter coupling, but with one of the metrics non-dynamical. It was shown that only when including an additional matter field, e.g. a scalar field, with highly contrived behaviour, was it possible to get viable cosmologies.

Structure formation Paper II analysed the equations of motion for cosmological perturbations in and near a de Sitter background. In particular, it investigated the relationship between the Higuchi bound, the behaviour of the massive perturbations and the parameters of the background. It was shown that the Higuchi bound has to be violated for consistency reasons on the level of the background. Perturbations which were identical in behaviour to that of general relativity in de Sitter, were shown to behave differently as compared to general relativity when entering the quasi-de Sitter regime.

Integrated Sachs-Wolfe effect The integrated Sachs-Wolfe effect is the modification of the cosmic microwave background (CMB) due to the time evolution of the gravitational potentials from the time of decoupling until today. This effect can be measured by cross-correlating the CMB with other tracers of the gravitational potentials, such as galaxies or quasars. Paper VI analysed the cross-correlation predicted by the so-called Infinite Branch Bigravity model of the Hassan-Rosen theory, which is given by a certain parameter choice. It was shown that the cross-correlation is larger than that predicted by general relativity, but still consistent with data.

Lensing Paper III investigated the lensing predicted by the Hassan-Rosen theory in the vicinity of galaxies, and used the observed lensing to put constraints on the parameter range of the theory. Lensing occurs when light travels in the gravitational potentials created by massive objects. By combining independent measurements of the galaxy masses, together with the observed lensing, it is possible to constrain different theories of gravity. Paper III showed that the effective parameter combination that sets the Compton wavelength of the graviton can not be in the range $5\,\mathrm{kpc} \lesssim \lambda_g \lesssim 5\,\mathrm{Mpc}$, where 5 kpc is roughly the scale of galaxies. A length scale above that is allowed for due to the Vainshtein mechanism, and a length scale below that, but above the scale of the solar system, is also viable.

In the introductory part of this thesis we described two of the outstanding problems in cosmology: dark matter and dark energy. These two phenomena have so far only been observed gravitationally, and it is possible that their inferred existence is due to our lack of understanding of gravitational interactions on length scales of galaxies and beyond. The natural question, then, is whether the Hassan-Rosen theory can shed some light on dark matter and dark energy.

The problem of dark energy is in reality a host of different problems. One is to understand the cause of the observed acceleration of the expansion of the universe. If this is caused by a cosmological constant, then the problem is to explain its observed value, and reconcile it with the value suggested by quantum field theory. If dark energy is not due to a cosmological constant, then one has to explain why the cosmological constant does not contribute to the dark energy component. A new, classical theory of gravity can in principle address the following questions: 1) How does gravity couple to the vacuum? 2) Are there self-accelerating solutions that do not rely on a cosmological constant? 3) Is there any natural parameter range for these solutions? The Hassan-Rosen theory answers these questions in the following way:

1. Since several possible matter couplings are allowed for in the Hassan-Rosen theory, the relationship between gravity and the vacuum energy density is not uniquely predicted. When coupling matter to only one of the metrics, the situation is similar to general relativity in that both matter loops and a bare cosmological constant term will contribute to the effective cosmological constant. When coupling matter to an effective metric built out of the two constituent metrics, the vacuum energy density will renormalize all the parameters of the theory.

2. Once again, the answer to this question depends on the type of coupling. If coupling only one of the metrics to matter, there does exist self-accelerating solutions. If coupling both metrics, then there are accelerating solutions, but since all the parameter terms are related to the vacuum energy density, it is not possible to unambiguously refer to these solutions as self-accelerating solutions.

3. There is no natural parameter range for the solutions that give rise to an accelerated expansion; instead, one has to fix the parameters to be of the order of the Hubble scale today in order to match the predicted acceleration with the observed value.

Concerning dark matter, it is possible that the Hassan-Rosen theory could be important. We have seen that the effective Planck mass, which sets the strength of the coupling between matter and gravity, is modified for cosmological solutions. This opens up for the possibility that the local Planck mass, as inferred through e.g. Cavendish experiments, is different from the global, or cosmological Planck mass. This means that the inferred dark matter content could be smaller or larger depending on the scale that it is probed at. To answer this question, one has to study spherically symmetric solutions, important for planetary orbits, and galactic solutions, important for rotation curves, to see how the Planck mass gets normalized at each level.

The Hassan-Rosen theory suffers from exponential instabilities when it comes to growth of structure in the universe. While not necessarily ruling the theory out, this could invalidate the perturbative framework commonly used for calculating the growth rate of structure. There are also problems with the instabilities related to the Higuchi bound. The exponential instabilities depend on the type of matter coupling employed, and it might be possible to cure them completely by some proper extension of the theory. At the moment, these instabilities pose the most serious threat to the validity of the theory.

The Hassan-Rosen theory is somewhat perplexing. It is constructed as a response to the physically well-motivated question "what does the theory of a massive graviton look like?" This simple question immediately leads into a rather complex theoretical realm with two rank-2 tensor fields, instead of just one. One gets five interaction terms, and several possible matter couplings. The cosmological community quickly showed that this gives rise to a rich phenomenology, as well as possible drawbacks in terms of instabilities. There exists a parameter range for which the exponential instabilities are avoided, and which gives a phenomenology consistent with observations. On the other hand, the existence of the instabilities might be a sign that the theory is by no means the final word on gravity. The resolution to this problem through further extensions or an improved understanding of the theory will be very interesting, and most likely carry a lot of surprises.

Bibliography

[1] S. F. Hassan and R. A. Rosen, *On Non-Linear Actions for Massive Gravity*, JHEP **1107** (2011) 009 [arXiv:1103.6055 [hep-th]].

[2] S. F. Hassan and R. A. Rosen, *Bimetric Gravity from Ghost-free Massive Gravity*, JHEP **1202** (2012) 126 [arXiv:1109.3515 [hep-th]].

[3] K. Freese, *Review of Observational Evidence for Dark Matter in the Universe and in upcoming searches for Dark Stars*, arXiv:0812.4005 [astro-ph].

[4] S. Perlmutter *et al.* [Supernova Cosmology Project Collaboration], *Measurements of Omega and Lambda from 42 high redshift supernovae*, Astrophys. J. **517** (1999) 565 [astro-ph/9812133].

[5] A. G. Riess *et al.* [Supernova Search Team Collaboration], *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, Astron. J. **116** (1998) 1009 [astro-ph/9805201].

[6] A. Einstein, *Zur allgemeinen Relativitätstheorie*, Sitzungsber. Preuss. Akad. Wiss. Berlin **1915** 778-786.

[7] J. Martin, *Everything you always wanted to know about the cosmological constant problem (but were afraid to ask)*, Comptes Rendux Physique, **13** (2012) [arXiv: 1205.3365 [astro-ph.CO]]

[8] S. M. Carroll, *The Cosmological constant*, Living Rev. Rel. **4** (2001) 1 [astro-ph/0004075].

[9] J. M. M. Senovilla and D. Garfinkle, *The 1965 Penrose singularity theorem*, arXiv:1410.5226 [gr-qc].

[10] C. P. Burgess, *Quantum gravity in everyday life: General relativity as an effective field theory*, Living Rev. Rel. **7** (2004) 5 [gr-qc/0311082].

[11] S. Weinberg, *Photons and gravitons in perturbation theory: Derivation of Maxwell's and Einstein's equations*, Phys. Rev. **138** (1965) B988.

[12] D. Lovelock, *The Einstein tensor and its generalizations*, J. Math. Phys. **12** (1971) 498.

[13] D. Lovelock, *The four-dimensionality of space and the Einstein tensor*, J. Math. Phys. **13** (1972) 874.

[14] T. Clifton, P. G. Ferreira, A. Padilla and C. Skordis, *Modified gravity and cosmology*, Phys. Rep., **513** (2012) [arXiv:1106.2476 [astro.ph.CO]].

[15] M. Ostrogradsky, *Mémoires sur les équations différentielles relatives au problème des isopérimètres*, Mem. Acad. St. Petersbourg VI **4** (1850) 385

[16] R. P. Woodard, *Avoiding dark energy with 1/r modifications of gravity*, Lect. Notes Phys. **720** (2007) 403 [astro-ph/0601672].

[17] A. De Felice and S. Tsujikawa, *f(R) theories*, Living Rev. Rel. **13**, 3 (2010) [arXiv:1002.4928 [gr-qc]].

[18] R. Maartens and K. Koyama, *Brane-World Gravity*, Living Rev. Rel. **13** (2010) 5 [arXiv:1004.3962 [hep-th]].

[19] G. R. Dvali, G. Gabadadze and M. Porrati, *4-D gravity on a brane in 5-D Minkowski space*, Phys. Lett. B **485** (2000) 208 [hep-th/0005016].

[20] R. P. Woodard, *Nonlocal Models of Cosmic Acceleration*, Found. Phys. **44** (2014) 213 [arXiv:1401.0254 [astro-ph.CO]].

[21] S. Liberati and D. Mattingly, *Lorentz breaking effective field theory models for matter and gravity: theory and observational constraints*, arXiv:1208.1071 [gr-qc].

[22] K. Krasnov, *Non-metric gravity: A Status report*, Mod. Phys. Lett. A **22** (2007) 3013 [arXiv:0711.0697 [gr-qc]].

[23] M. Maggiore, *Gravitational Waves. Vol. 1: Theory and Experiments*, Oxford 2008.

[24] K. Hinterbichler, *Theoretical Aspects of Massive Gravity*, Rev. Mod. Phys. **84** (2012) 671 [arXiv:1105.3735 [hep-th]].

[25] M. Fierz and W. Pauli, *On relativistic wave equations for particles of arbitrary spin in an electromagnetic field*, Proc. Roy. Soc. Lond. A **173** (1939) 211.

[26] W. Pauli and M. Fierz, *On Relativistic Field Equations of Particles With Arbitrary Spin in an Electromagnetic Field*, Helv. Phys. Acta **12** (1939) 297.

[27] H. van Dam and M. J. G. Veltman, *Massive and massless Yang-Mills and gravitational fields*, Nucl. Phys. B **22** (1970) 397.

[28] V. I. Zakharov, *Linearized gravitation theory and the graviton mass*, JETP Lett. **12** (1970) 312 [Pisma Zh. Eksp. Teor. Fiz. **12** (1970) 447].

[29] P. Van Nieuwenhuizen, *On ghost-free tensor lagrangians and linearized gravitation*, Nucl. Phys. B **60** (1973) 478.

[30] A. I. Vainshtein, *To the problem of nonvanishing gravitation mass*, Phys. Lett. B **39** (1972) 393.

[31] S. Deser, *Selfinteraction and gauge invariance*, Gen. Rel. Grav. **1** (1970) 9 [gr-qc/0411023].

[32] T. Padmanabhan, *From gravitons to gravity: Myths and reality*, Int. J. Mod. Phys. D **17** (2008) 367 [gr-qc/0409089].

[33] D. G. Boulware and S. Deser, *Can gravitation have a finite range?*, Phys. Rev. D **6** (1972) 3368.

[34] A. H. Chamseddine, A. Salam and J. A. Strathdee, *Strong Gravity and Supersymmetry*, Nucl. Phys. B **136** (1978) 248.

[35] P. Creminelli, A. Nicolis, M. Papucci and E. Trincherini, *Ghosts in massive gravity*, JHEP **0509** (2005) 003 [hep-th/0505147].

[36] S. Groot Nibbelink, M. Peloso and M. Sexton, *Nonlinear Properties of Vielbein Massive Gravity*, Eur. Phys. J. C **51** (2007) 741 [hep-th/0610169]

[37] C. de Rham and G. Gabadadze, *Generalization of the Fierz-Pauli Action*, Phys. Rev. D **82** (2010) 044020 [arXiv:1007.0443 [hep-th]].

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

[39] S. F. Hassan and R. A. Rosen, *Resolving the Ghost Problem in non-Linear Massive Gravity*, Phys. Rev. Lett. **108** (2012) 041101 [arXiv:1106.3344 [hep-th]].

[40] S. F. Hassan, R. A. Rosen and A. Schmidt-May, *Ghost-free Massive Gravity with a General Reference Metric*, JHEP **1202** (2012) 026 [arXiv:1109.3230 [hep-th]].

[41] S. F. Hassan and R. A. Rosen, *Confirmation of the Secondary Constraint and Absence of Ghost in Massive Gravity and Bimetric Gravity*, JHEP **1204** (2012) 123 [arXiv:1111.2070 [hep-th]].

[42] K. Hinterbichler and R. A. Rosen, *Interacting Spin-2 Fields*, JHEP **1207** (2012) 047 [arXiv:1203.5783 [hep-th]].

[43] S. F. Hassan, A. Schmidt-May and M. von Strauss, *On Consistent Theories of Massive Spin-2 Fields Coupled to Gravity*, JHEP **1305** (2013) 086 [arXiv:1208.1515 [hep-th]].

[44] K. Aoki and K. i. Maeda, *Dark matter in ghost-free bigravity theory: From a galaxy scale to the universe*, arXiv:1409.0202 [gr-qc].

[45] Y. Akrami, T. S. Koivisto and A. R. Solomon, *The nature of spacetime in bigravity: two metrics or none?*, arXiv:1404.0006 [gr-qc].

[46] C. de Rham, L. Heisenberg and R. H. Ribeiro, *On couplings to matter in massive (bi-)gravity*, arXiv:1408.1678 [hep-th].

[47] Y. Yamashita, A. De Felice and T. Tanaka, *Appearance of Boulware-Deser ghost in bigravity with doubly coupled matter*, arXiv:1408.0487 [hep-th].

[48] C. de Rham, L. Heisenberg and R. H. Ribeiro, *Ghosts & Matter Couplings in Massive (bi-&multi-)Gravity*, arXiv:1409.3834 [hep-th].

[49] S. F. Hassan, M. Kocic and A. Schmidt-May, *Absence of ghost in a new bimetric-matter coupling*, arXiv:1409.1909 [hep-th].

[50] A. Schmidt-May, *Mass eigenstates in bimetric theory with matter coupling*, arXiv:1409.3146 [gr-qc].

[51] J. Noller, *On Consistent Kinetic and Derivative Interactions for Gravitons*, arXiv:1409.7692 [hep-th].

[52] J. Noller and S. Melville, *The coupling to matter in Massive, Bi- and Multi-Gravity*, arXiv:1408.5131 [hep-th].

[53] L. Heisenberg, *Quantum corrections in massive bigravity and new effective composite metrics*, arXiv:1410.4239 [hep-th].

[54] C. M. Will, *The Confrontation between General Relativity and Experiment*, Living Rev. Rel. **17** (2014) 4 [arXiv:1403.7377 [gr-qc]].

[55] C. de Rham and S. Renaux-Petel, *Massive Gravity on de Sitter and Unique Candidate for Partially Massless Gravity*, JCAP **1301** (2013) 035 [arXiv:1206.3482 [hep-th]].

[56] S. F. Hassan, A. Schmidt-May and M. von Strauss, *On Partially Massless Bimetric Gravity*, Phys. Lett. B **726,** (2013) 834 [arXiv:1208.1797 [hep-th]].

[57] S. F. Hassan, A. Schmidt-May and M. von Strauss, *Higher Derivative Gravity and Conformal Gravity From Bimetric and Partially Massless Bimetric Theory*, arXiv:1303.6940 [hep-th].

[58] S. F. Hassan, A. Schmidt-May and M. von Strauss, *Particular Solutions in Bimetric Theory and Their Implications*, arXiv:1407.2772 [hep-th].

[59] C. de Rham, K. Hinterbichler, R. A. Rosen and A. J. Tolley, *Evidence for and obstructions to nonlinear partially massless gravity*, Phys. Rev. D **88** (2013) 2, 024003 [arXiv:1302.0025 [hep-th]].

[60] K. Hinterbichler, *Manifest Duality Invariance for the Partially Massless Graviton*, arXiv:1409.3565 [hep-th].

[61] S. Garcia-Saenz and R. A. Rosen, *A non-linear extension of the spin-2 partially massless symmetry*, arXiv:1410.8734 [hep-th].

[62] P. A. R. Ade *et al.* [Planck Collaboration], *Planck 2013 results. XVI. Cosmological parameters*, Astron. Astrophys. (2014) [arXiv:1303.5076 [astro-ph.CO]].

[63] E. Aubourg, S. Bailey, J. E. Bautista, F. Beutler, V. Bhardwaj, D. Bizyaev, M. Blanton and M. Blomqvist *et al.*, *Cosmological implications of baryon acoustic oscillation (BAO) measurements*, arXiv:1411.1074 [astro-ph.CO].

[64] M. Betoule *et al.* [SDSS Collaboration], *Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples*, [arXiv:1401.4064 [astro-ph.CO]].

[65] A. Conley *et al.* [SNLS Collaboration], *Supernova Constraints and Systematic Uncertainties from the First 3 Years of the Supernova Legacy Survey*, Astrophys. J. Suppl. **192** (2011) 1 [arXiv:1104.1443 [astro-ph.CO]].

[66] N. Suzuki *et al.*, *The Hubble Space Telescope Cluster Supernova Survey. V. Improving the Dark-energy Constraints above z¿1 and Building an Early-type-hosted Supernova Sample*, Astrophys. J., **746** (2012) [arXiv:1105.3470 [astro.ph.CO]].

[67] G. D'Amico, C. de Rham, S. Dubovsky, G. Gabadadze, D. Pirtskhalava and A. J. Tolley, *Massive Cosmologies*, Phys. Rev. D **84** (2011) 124046 [arXiv:1108.5231 [hep-th]].

[68] A. E. Gümrükçüoğlu, C. Lin and S. Mukohyama, *Open FRW universes and self-acceleration from nonlinear massive gravity*, JCAP **1111** (2011) 030 [arXiv:1109.3845 [hep-th]].

[69] B. Vakili and N. Khosravi, *Classical and quantum massive cosmology for the open FRW universe*, Phys. Rev. D **85** (2012) 083529 [arXiv:1204.1456 [gr-qc]].

[70] A. E. Gümrükçüoğlu, C. Lin and S. Mukohyama, *Cosmological perturbations of self-accelerating universe in nonlinear massive gravity*, JCAP **1203** (2012) 006 [arXiv:1111.4107 [hep-th]].

[71] A. De Felice, A. E. Gümrükçüoğlu and S. Mukohyama, *Massive gravity: nonlinear instability of the homogeneous and isotropic universe*, Phys. Rev. Lett. **109** (2012) 171101 [arXiv:1206.2080 [hep-th]].

[72] M. Fasiello and A. J. Tolley, *Cosmological perturbations in Massive Gravity and the Higuchi bound*, JCAP **1211** (2012) 035 [arXiv:1206.3852 [hep-th]].

[73] A. De Felice, A. E. Gümrükçüoğlu, C. Lin and S. Mukohyama, *Nonlinear stability of cosmological solutions in massive gravity*, JCAP **1305** (2013) 035 [arXiv:1303.4154 [hep-th]].

[74] A. E. Gümrükçüoğlu, L. Heisenberg and S. Mukohyama, *Cosmological perturbations in massive gravity with doubly coupled matter*, arXiv:1409.7260 [hep-th].

[75] M. S. Volkov, *Cosmological solutions with massive gravitons in the bigravity theory*, JHEP **1201** (2012) 035 [arXiv:1110.6153 [hep-th]].

[76] D. Comelli, M. Crisostomi, F. Nesti and L. Pilo, *FRW Cosmology in Ghost Free Massive Gravity*, JHEP **1203** (2012) 067 [Erratum-ibid. **1206** (2012) 020] [arXiv:1111.1983 [hep-th]].

[77] Y. Akrami, T. S. Koivisto and M. Sandstad, *Accelerated expansion from ghost-free bigravity: a statistical analysis with improved generality*, JHEP **1303** (2013) 099 [arXiv:1209.0457 [astro-ph.CO]].

[78] F. Koennig, A. Patil and L. Amendola, *Viable cosmological solutions in massive bimetric gravity*, JCAP **1403** (2014) 029 [arXiv:1312.3208 [astro-ph.CO]].

[79] Y. Akrami, T. S. Koivisto, D. F. Mota and M. Sandstad, *Bimetric gravity doubly coupled to matter: theory and cosmological implications*, JCAP **1310** (2013) 046 [arXiv:1306.0004 [hep-th]].

[80] K. Aoki and K. i. Maeda, *Cosmology in ghost-free bigravity theory with twin matter fluids: The origin of "dark matter"*, Phys. Rev. D **89** (2014) 064051 [arXiv:1312.7040 [gr-qc]].

[81] D. S. Gorbunov and V. A. Rubakov, *Introduction to the theory of the early universe: Cosmological perturbations and inflationary theory*, Hackensack, USA: World Scientific (2011)

[82] D. Comelli, M. Crisostomi and L. Pilo, *Perturbations in Massive Gravity Cosmology*, JHEP **1206** (2012) 085 [arXiv:1202.1986 [hep-th]].

[83] D. Comelli, M. Crisostomi and L. Pilo, *FRW Cosmological Perturbations in Massive Bigravity*, Phys. Rev. D **90** (2014) 084003 [arXiv:1403.5679 [hep-th]].

[84] F. Könnig and L. Amendola, *Instability in a minimal bimetric gravity model*, Phys. Rev. D **90** (2014) 044030 [arXiv:1402.1988 [astro-ph.CO]].

[85] A. R. Solomon, Y. Akrami and T. S. Koivisto, *Linear growth of structure in massive bigravity*, JCAP **1410** (2014) 066 [arXiv:1404.4061 [astro-ph.CO]].

[86] F. Könnig, Y. Akrami, L. Amendola, M. Motta and A. R. Solomon, *Stable and unstable cosmological models in bimetric massive gravity*, arXiv:1407.4331 [astro-ph.CO].

[87] N. Khosravi, H. R. Sepangi and S. Shahidi, *Massive cosmological scalar perturbations*, Phys. Rev. D **86** (2012) 043517 [arXiv:1202.2767 [gr-qc]].

[88] A. De Felice, A. E. Gümrükçüoğlu, S. Mukohyama, N. Tanahashi and T. Tanaka, *Viable cosmology in bimetric theory*, JCAP **1406** (2014) 037 [arXiv:1404.0008 [hep-th]].

[89] M. Lagos and P. G. Ferreira, *Cosmological perturbations in massive bigravity*, arXiv:1410.0207 [gr-qc].

[90] C. de Rham, G. Gabadadze and A. J. Tolley, *Helicity Decomposition of Ghost-free Massive Gravity*, JHEP **1111** (2011) 093 [arXiv:1108.4521 [hep-th]].

[91] L. Alberte, *Massive Gravity on Curved Background*, Int. J. Mod. Phys. D **21** (2012) 1250058 [arXiv:1110.3818 [hep-th]].

[92] L. F. Abbott and S. Deser, *Stability of Gravity with a Cosmological Constant*, Nucl. Phys. B **195** (1982) 76.

[93] S. Deser and R. I. Nepomechie, *Gauge Invariance Versus Masslessness in De Sitter Space*, Annals Phys. **154** (1984) 396.

[94] S. Deser and R. I. Nepomechie, *Anomalous Propagation of Gauge Fields in Conformally Flat Spaces*, Phys. Lett. B **132** (1983) 321.

[95] A. Higuchi, *Forbidden Mass Range for Spin-2 Field Theory in De Sitter Space-time*, Nucl. Phys. B **282** (1987) 397.

[96] A. Higuchi, *Massive Symmetric Tensor Field in Space-times With a Positive Cosmological Constant*, Nucl. Phys. B **325** (1989) 745.

[97] A. Higuchi, *Symmetric Tensor Spherical Harmonics on the N Sphere and Their Application to the De Sitter Group SO(N,1)*, J. Math. Phys. **28** (1987) 1553

[98] S. Deser and A. Waldron, *Stability of massive cosmological gravitons*, Phys. Lett. B **508** (2001) 347 [hep-th/0103255].

[99] S. Deser and A. Waldron, *Gauge invariances and phases of massive higher spins in (A)dS*, Phys. Rev. Lett. **87** (2001) 031601 [hep-th/0102166].

[100] S. Deser and A. Waldron, *Partial masslessness of higher spins in (A)dS*, Nucl. Phys. B **607** (2001) 577 [hep-th/0103198].

[101] Y. Sakakihara, J. Soda and T. Takahashi, *On Cosmic No-hair in Bimetric Gravity and the Higuchi Bound*, PTEP **2013** (2013) 033E02 [arXiv:1211.5976 [hep-th]].

[102] M. Fasiello and A. J. Tolley, *Cosmological Stability Bound in Massive Gravity and Bigravity*, JCAP **1312** (2013) 002 [arXiv:1308.1647 [hep-th]].

[103] D. Comelli, M. Crisostomi, K. Koyama, L. Pilo and G. Tasinato, *Cosmology of bigravity with doubly coupled matter*, arXiv:1501.00864 [hep-th].

[104] G. Cusin, R. Durrer, P. Guarato and M. Motta, *Gravitational waves in bigravity cosmology*, arXiv:1412.5979 [astro-ph.CO].

[105] R. K. Sachs and A. M. Wolfe, *Perturbations of a cosmological model and angular variations of the microwave background*, Astrophys. J. **147** (1967) 73 [Gen. Rel. Grav. **39** (2007) 1929].

[106] A. J. Nishizawa, *The integrated Sachs-Wolfe effect and the Rees-Sciama effect*, PTEP **2014** (2014) 6, 06B110 [arXiv:1404.5102 [astro-ph.CO]].

[107] J. M. Bardeen, J. R. Bond, N. Kaiser and A. S. Szalay, *The Statistics of Peaks of Gaussian Random Fields*, Astrophys. J. **304** (1986) 15.

[108] R. G. Crittenden and N. Turok, *Looking for Lambda with the Rees-Sciama effect*, Phys. Rev. Lett. **76** (1996) 575 [astro-ph/9510072].

[109] S. Ferraro, B. D. Sherwin and D. N. Spergel, *A WISE measurement of the ISW effect*, arXiv:1401.1193 [astro-ph.CO].

[110] E. L. Wright, P. R. M. Eisenhardt, A. Mainzer, M. E. Ressler, R. M. Cutri, T. Jarrett, J. D. Kirkpatrick and D. Padgett *et al.*, *The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance*, Astron. J. **140** (2010) 1868 [arXiv:1008.0031 [astro-ph.IM]].

[111] L. Yan, E. Donoso, C.-W. Tsai, D. Stern, R. J. Assef, P. Eisenhardt, A. W. Blain and R. Cutri *et al.*, *Characterizing the Mid-IR Extragalactic Sky with WISE and SDSS*, Astron. J. **145** (2013) 55 [arXiv:1209.2065 [astro-ph.CO]].

[112] C. L. Bennett *et al.* [WMAP Collaboration], *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results*, Astrophys. J. Suppl. **208** (2013) 20 [arXiv:1212.5225 [astro-ph.CO]].

[113] D. Comelli, M. Crisostomi, F. Nesti and L. Pilo, *Spherically Symmetric Solutions in Ghost-Free Massive Gravity*, Phys. Rev. D **85**, 024044 (2012) [arXiv:1110.4967 [hep-th]].

[114] M. S. Volkov, *Hairy black holes in the ghost-free bigravity theory*, Phys. Rev. D **85** (2012) 124043 [arXiv:1202.6682 [hep-th]].

[115] E. Babichev and C. Deffayet, *An introduction to the Vainshtein mechanism*, Class. Quant. Grav. **30** (2013) 184001 [arXiv:1304.7240 [gr-qc]].

[116] E. Babichev and M. Crisostomi, *Restoring general relativity in massive bigravity theory*, Phys. Rev. D **88** (2013) 8, 084002 [arXiv:1307.3640].

[117] E. Babichev and A. Fabbri, *Rotating black holes in massive gravity*, Phys. Rev. D **90** (2014) 084019 [arXiv:1406.6096 [gr-qc]].

[118] E. Babichev and A. Fabbri, *A class of charged black hole solutions in massive (bi)gravity*, JHEP **1407** (2014) 016 [arXiv:1405.0581 [gr-qc]].

[119] R. Brito, V. Cardoso and P. Pani, *Massive spin-2 fields on black hole spacetimes: Instability of the Schwarzschild and Kerr solutions and bounds on the graviton mass*, Phys. Rev. D **88** (2013) 2, 023514 [arXiv:1304.6725 [gr-qc]].

[120] E. Babichev and A. Fabbri, *Instability of black holes in massive gravity*, Class. Quant. Grav. **30** (2013) 152001 [arXiv:1304.5992 [gr-qc]].

[121] E. Babichev and A. Fabbri, *Stability analysis of black holes in massive gravity: a unified treatment*, Phys. Rev. D **89** (2014) 081502 [arXiv:1401.6871 [gr-qc]].

[122] R. Brito, V. Cardoso and P. Pani, *Black holes with massive graviton hair*, Phys. Rev. D **88** (2013) 064006 [arXiv:1309.0818 [gr-qc]].

[123] M. S. Volkov, *Hairy black holes in theories with massive gravitons*, arXiv:1405.1742 [hep-th].

[124] M. Bartelmann and P. Schneider, *Weak gravitational lensing*, Phys. Rept. **340** (2001) 291 [astro-ph/9912508].

[125] S. Reynaud and M. T. Jaekel, *Testing the Newton law at long distances*, Int. J. Mod. Phys. A **20** (2005) 2294 [gr-qc/0501038].

[126] B. Bertotti, L. Iess and P. Tortora, *A test of general relativity using radio links with the Cassini spacecraft*, Nature **425** (2003) 374.