CHALMERS
UNIVERSITY OF TECHNOLOGY

Large-scale structure formation in bimetric gravity

Master of Science Thesis in Fundamental Physics
Malin Nicole Renneby

Department of Fundamental Physics
Chalmers University of Technology
Gothenburg, Sweden, 2014

Large-scale structure formation in bimetric gravity

MALIN NICOLE RENNEBY

MASTER OF SCIENCE THESIS IN FUNDAMENTAL PHYSICS
FUFX03

THE OSKAR KLEIN CENTRE FOR COSMOPARTICLE PHYSICS
STOCKHOLM UNIVERSITY

DEPARTMENT OF FUNDAMENTAL PHYSICS
CHALMERS UNIVERSITY OF TECHNOLOGY

SUPERVISORS:
DR. EDVARD MÖRTSELL
JONAS ENANDER

EXAMINER:
PROF. DR. GABRIELE FERRETTI

SEPTEMBER 16, 2014

Large-scale structure formation in bimetric gravity

Author: Malin Nicole Renneby

FUFX03 - Master of Science Thesis in Fundamental Physics

Supervisors:

Dr. Edvard Mörtsell and Jonas Enander
Cosmology, Particle Astrophysics and String Theory Group
Oskar Klein Centre for Cosmoparticle Physics
Department of Physics
Stockholm University
Stockholm, Sweden

Examiner:

Prof. Dr. Gabriele Ferretti
Division of Elementary Particle Physics
Department of Fundamental Physics
Chalmers University of Technology
Gothenburg, Sweden

Contact: renneby@student.chalmers.se

Department of Fundamental Physics
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
+ 46 (0) 31 772 1000

Printed by Chalmers Reproservice
Gothenburg, Sweden 2014

Abstract

One of the most prominent questions in modern cosmology is the origin of the accelerated expansion of the universe. A solution might lie in modifying gravity in the infrared by adding a small mass to its mediating particle. In recent years, dRGT massive gravity and its dynamical extension Hassan-Rosen bimetric massive gravity have been shown to be classically consistent theories. This prompts for a phenomenological investigation of their predictions in cases already examined in general relativity, such as in spherically-symmetric geometries and on cosmological scales.

The thesis, conducted within the Cosmology, Particle Astrophysics and String Theory group at Stockholm University, elaborates on aspects related to the evolution of large-scale structures through analysis of the bimetric equations of motion for linear perturbations. We review the foundations of relativistic perturbation theory in general relativity. A particular emphasis is placed on superhorizon signatures with the integrated Sachs-Wolfe effect as a candidate mechanism. Moreover, we present the theoretical framework of bimetric massive gravity with applications to cosmology, both on the background level and concerning linear perturbations. The single-coupled theory with FLRW ansätze for the background metrics is investigated in the self-accelerating minimal bimetric β_1- and the infinite-branch $\beta_1\beta_4-$models, where the latter has attracted significant interest recently, with a normalization density condition provided by the Planck 2013 survey. We solve the system of equations of motion for the linear scalar perturbation fields in the $F_g = F_f = 0-$gauge following the notation of Solomon et al. and find an additional alternative by examining the Noether identities for the second-order perturbed action based on the method demonstrated by Lagos et al. To conclude, we plot the evolution of the linearized bimetric gravitational potentials from the early universe till today and discuss the relation to the predictions of ΛCDM.

Acknowledgements

Firstly, I would like to express my sincerest gratitude towards my supervisors Edvard Mörtsell and Jonas Enander for making this work possible. Many thanks to Angnis Schmidt-May and Fawad Hassan for many fruitful discussions as well as the development of this beautiful theory. Un certain regard to Angnis for her exceptional, accessible PhD thesis which has helped to clarify this analysis tremendously. Moreover, I would like to convey my appreciation to Adam R. Solomon for being a splendid host during the First Annual Bigravity World Congress at DAMTP, Centre for Mathematical Sciences, University of Cambridge in June 2014, as well as for inciting plenty of rewarding discussions with outstanding contributions from Frank Könnig, who actually solved the system, and Yashar Akrami, who has provided many prominent strategic insights. To continue, I would like to extend my thanks to Stefan Hilbert and Andreas Müller for endowing me with the opportunity to present some results from this treatise at the Excellence Cluster Universe in Garching, Germany, in August 2014, which was very beneficial and delightful. As for administrative matters, I am most grateful for the assistance from Lars Brink, Bo Sundborg, Lars Bergström, Joakim Edsjö and my examiner Gabriele Ferretti without whom I could never have carried out this thesis in Stockholm. In addition I would like to thank Florian Kühnel for being a superb office mate, offering astute advice and many laughs, as well as all members of the CoPS-group for hosting me, introducing me to the paramount scientific environment and enlightening my days. To conclude, I recognize with my most profound heartfelt thankfulness the massive support from my family and my friends which has bestowed upon me a beacon of hope, facing the impossible, the improbable, presenting a road ahead.

Contents

Chapter 1 Introduction

1.1 Motivation

Numerous questions remain unanswered in modern theoretical physics, the experimentally observed accelerated expansion of the universe one of them, [1] [2], usually attributed to an elusive *dark energy*. This prompts a theoretical description, and motivates a study of modified gravity theories. The accelerating universe could be described by a non-zero cosmological constant in Einstein's equations, but this constant is far below the value one would expect from studying the energy density arising from the non-zero vacuum expectation value predicted by quantum field theory. This acute situation is known as *the cosmological constant problem*. A natural attempt to resolve this issue could be through the introduction of a massive spin 2-mode for gravity. Such theories have attracted renewed interest in recent years with the proof of the classical consistency of *massive gravity* and its generalized extension *bimetric gravity*, [3] [4] [5], which are theories with two metrics. The Cosmology, Particle astrophysics and String theory (CoPS) group at Stockholm University, part of the Oskar Klein Center for Cosmoparticle Physics, has been leading in this paradigm shift. In order to test the validity of this theory one has compared zero:th order solutions with cosmological observations, [6]. Moreover, a study of first-order linear perturbations for general homogeneous and isotropic backgrounds have been conducted, [7], with an analysis of explicit solutions for de Sitter and quasi-de Sitter approximations. As linear perturbations represent the advent of structure formation, their investigation is crucial to phenomenologically test the predictions of bimetric gravity in the light of observational data. Optimally, one could obtain a class of theories providing viable cosmological scenarios, which would be subjected to further examinations.

At Stockholm University, I have been supervised by Dr. Edvard Mörtsell and Jonas Enander. Examiner at the Department of Fundamental Physics at Chalmers University of Technology is Prof. Dr. Gabriele Ferretti.

1.2 Aim

In this project, the goal is to find one or several numerical solution(s) to the equations of motion for linear perturbations in bimetric gravity, see [7], using gauge invariant variables. Previously, solutions have been found for de-Sitter and quasi-de Sitter backgrounds analytically, and now the quest is to comprehend the general case. Moreover, the theoretical foundations of the theory are studied with help from recent reviews, [8] [9], together with the excellent PhD thesis [10]. Additionally, insights from recent progress in the field, [11], are analyzed and incorporated in the final result.

1.3 Objectives

Serving as a guide throughout the project, the work is trying to find answers to the following questions of scientific interest:

(i) Which approximations could successfully be made?

(ii) Which initial conditions are appropriate to allow for a general solution that is physically viable?

(iii) How can we test the mathematical properties of the solutions (stability)?

(iv) Which symmetries do the solutions exhibit?

(v) In which gauge is the physical interpretation most evident?

(vi) How do the solutions relate to experimental data from telescopes?

1.4 Scope and notation

Bimetric gravity is a nascent field and hence there remains a lot to explore. This project is limited to the study of first-order linear perturbations for independent wave modes with no anisotropic stress factors. Such a restriction simplifies the equations of motion, reducing the description of the problem in terms of partial differential equations to differential algebraic equations and yield a class of solutions which presumably is close to observational data, which is true for general relativity. Primarily, superhorizon length scales are investigated as subhorizon solutions were previously examined in [11]. Currently, the subject of suspected instabilities in the perturbation equations, highlighted for instance in [12], remains an unresolved, moot point and hence our excursion into this territory will be circumscribed. We will not discuss bimetric black hole solutions but urge the interested reader to ponder results from for instance [13] and [14]. Effects in the context of gravitational lensing are highlighted in [15]. Concerning different frameworks, we will present the work using the covariant formulation of the theory. An alternative, the vielbein formalism, has facilitated the venture into more general multimetric theories [16] [17]. Still, its equivalence with the covariant formulation involves interpretative subtleties [16] [18]. Together with the amount of research conducted using the prior framework, this serves as an impetus to work covariantly but we encourage the interested reader to study [9] for a comprehensive review. Moreover, the purpose of this thesis is restricted to investigate properties of bimetric gravity and we will not perform a benchmark against other modified theories of gravity. An overview of bimetric gravity in this context can be found in [19]. Regarding notation, we use the mostly plus $-, +, \dots, +$ metric convention with $c = 1$. Repeated indices are summed over unless otherwise explicitly specified.

1.5 Method

The work is principally conducted through simulations in `Mathematica` complemented by calculations by pen and paper. Numerical methods and physical aspects of bimetric theory and its connection to general relativity are investigated through literature studies preceding the coding and have continue throughout the project. Several discussions with supervisors and group members as well as active participation in relevant seminars and meetings help in clarifying some of the more challenging theoretical and computational aspects. Non-plot figures are drawn in `Adobe Illustrator CC` unless an external source is credited.

Chapter 2 Large-scale structure formation

The universe consists of more than empty space, including nebulae, interstellar dust and a tiny blue-green world. To allow for these things formally, we examine small deviations from the unperturbed homogeneous and isotropic background solution and see how these develop over time, resulting in *structure formation*. Observationally when we refer to *large-scale structures*, we are chiefly interested in counting the number of galaxies in the sky and analyzing how they are clustered together at different times (see an artistic sketch in figure 2.1). An individual galaxy evolves in a nonlinear fashion and the same is true for a group of galaxies, but on larger scales one could predict a *linear* evolution given the equations of motion for the perturbations in a chosen model.

Physically, we seek to determine the density contrast $\delta(t) = \delta\rho(t)/\bar{\rho}(t)$, where $\bar{\rho}(t)$ denotes the unperturbed density and $\delta\rho(t)$ its variation, between galaxy clusters and void. In Einstein's theory of general relativity (GR), [20], we model these inhomogeneities by perturbing the stress-energy-tensor which corresponds to perturbing the metric on the other side of the equal sign. If we only consider perturbations at the linear order, which predominately govern the formation of large-scale structures and which we model as small perturbations of the background values, it is possible to relate the perturbed quantities through

$$\mathcal{L}\left(\bar{g}_{ab}\right)\delta g_{ab} = \delta T_{ab}\,, \tag{2.1}$$

where the perturbations are denoted by δ and $\mathcal{L}$ is a second-order differential operator depending on the unperturbed background metric $\bar{g}_{ab}$. If one has a maximal symmetric background, which is the case in homogeneous+isotropic cosmology, one may decompose the space and time dependencies separately for the perturbations. In our analysis we use this to perform a Fourier transform with respect to the spatial $\boldsymbol{x}$. We are then able to write down this equation (2.1) for each mode which is labelled by a wavenumber $\boldsymbol{k}$. At the linear level there will be no mixing between different modes in the equations of motion. Assuming that this $\boldsymbol{k}$ is spatially isotropic, our system of equations for the different perturbations is reduced from a system of partial differential equations to one of ordinary differential equations or one of differential-algebraic equations, i.e. a system where some variables appear in algebraic constraints.

The origin of these inhomogeneities remains speculative at present, although quantum fluctuations of an primordial inflation field following the Big Bang is a popular theory which agrees with our current observations. Still, it is possible to make predictions without knowing the explicit details of the progenitors to the observable inhomogeneities as we have a good understanding on how to model the evolution using the framework of general relativity. Despite this promise some uncertainties remain as we are not fully aware of the nature of *dark matter* and *dark energy*. If our goal is to probe the viability of modified theories of gravity, we must redo and adapt the steps which we will lay out in this first chapter on large-scale

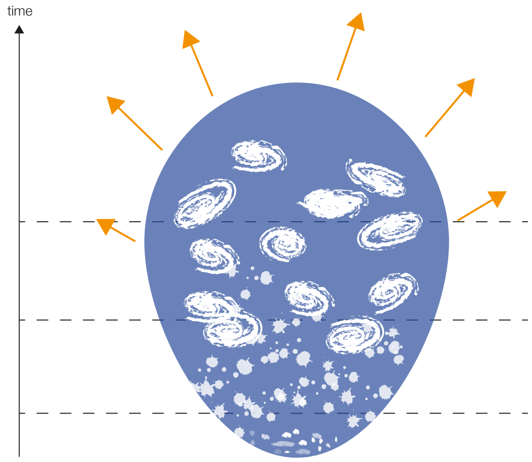


Figure 2.1: Illustration of the development of large-scale structures in an expanding universe here depicted as a balloon, characterized by the density contrast $\delta(t)$, from primordial inhomogeneities represented by the small dots at the bottom of the balloon. Note that many more galaxies than those drawn here are required to form large-scale structures.

structure formation in GR.

We will begin by introducing concepts such as redshift and the Hubble length to be able to define what we mean with large structures in more detail. Later, we proceed to a review of cosmological solutions and then into the mathematical modeling of structure. A corresponding treatment within the framework of bimetric theory will be presented in chapter 4.

2.1 Preliminaries

To refresh the readers' memories and facilitate further discussions on the topic, we will present a brief overview of Friedmann's equations for a homogeneous and isotropic universe. This is the foundation for the Λ*CDM model*, the standard model of cosmology which is our best fit to the observed experimental data. ΛCDM mean "cold dark matter with a nonzero cosmological constant Λ", whose background solution is governed by the Friedmann equations where we have added a cosmological constant term plus have split the matter density component into ordinary and dark matter.

Firstly, we will discuss homogeneity and isotropy before we move on to the derivation. Expressed in topological terms the following definition applies: A spacetime (manifold $\mathcal{M}$ with metric g) is *spatially homogeneous* if there exists a group of isometries whose orbits are three-dimensional spacelike surfaces. The orbit of a point $p \in \mathcal{M}$ are all points obtained by acting on p with the isometry group. A surface is classified as spacelike if all tangent vectors are spacelike and the spacetime metric will induce a homogeneous metric (pullback) on each of these surfaces. This means that through each point in a spatially homogeneous spacetime there exists a three-dimensional surface with a homogeneous metric. Proceeding to *isotropy*, i.e. that the universe looks the same in all directions, it is important to emphasize that only a certain class of observers can see it in that way. Suppose an observer at p has a 4-velocity with a non-zero component along the surface of spatial homogeneity through p then this selects a favored spatial direction, incompatible with isotropy. Hence, only observes whose 4-velocities are normal to the surfaces of spatial homogeneity can perceive the isotropy. Such observers are denoted *comoving*.

As a starting point we start with the Einstein-Hilbert action with a matter coupling and a cosmological constant term Λ in dimension $d = 4$,

$$S_{\mathrm{EH}+m} = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{1}{M_{\mathrm{PL}}^2}(R - 2\Lambda) + \mathcal{L}_m\right], \tag{2.2}$$

where M_{PL} is the Planck mass, $M_{\mathrm{PL}}^2 = 1/(8\pi G)$ in units $c = \hbar = 1$, whose variation yields Einstein's field equations

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = \frac{1}{M_{PL}^2}T_{\mu\nu}, \tag{2.3}$$

with $T_{\mu\nu}$ as the stress-energy tensor obtained from varying the matter Lagrangian piece. In spherical polar coordinates, the Friedmann-Lemaître-Robertson-Walker (FLRW) line element becomes

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + a(t)^2\left[\frac{\mathrm{d}r^2}{1 - kr^2} + r^2\,\mathrm{d}\Omega\right], \tag{2.4}$$

where $\mathrm{d}\Omega$ is the differential solid angle, $a(t)$ the cosmological scale factor and k[1] associated to the geometry of spacetime, corresponding to

$$k = \begin{cases} +1, & \text{spherical,} \\ 0, & \text{flat, Euclidean,} \\ -1, & \text{hyperbolical.} \end{cases} \tag{2.5}$$

Unless explicitly written out, we will generally assume $k = 0$.

Imagine an observer who is positioned at the center of an FLRW coordinate system at $r = 0$ and who receives radiated light at a time $t = t_0$ emitted from a source some radial distance $r = r_s$ at time $t = t_s$. The two events are connected by a null geodesic and, taking the distance to be purely radial, the light ray obeys

$$\mathrm{d}s^2 = 0 \rightarrow \mathrm{d}t = \pm a(t)\frac{\mathrm{d}r}{(1 - kr^2)^{1/2}}. \tag{2.6}$$

It is possible to establish a relation between r_s and t_s through integration

$$\int_{t_s}^{t_0} \frac{\mathrm{d}t}{a(t)} = \int_0^{r_s} \frac{\mathrm{d}r}{(1 - kr^2)^{1/2}}. \tag{2.7}$$

If we examine the differential of this equation, keeping in mind that radial coordinate of comoving sources r_1 is time-independent, one obtains that the interval between emitted subsequent light signals δt_s is related to the interval of arrival of the same signals δt_0 through

[1] Since there are few available fitting letters, we will also denote different wave modes by k, but it will be evident from the context which definition applies.

$$\frac{\delta t_s}{a(t_s)} = \frac{\delta t_0}{a(t_0)}. \tag{2.8}$$

For subsequent wave crests, one deduces from this relation that the observed frequency $\nu_0 = 1/\delta t_0$ is connected to the emitted frequency $\nu_s = 1/\delta t_s$ through the ratio $\nu_0/\nu_1 = a(t_s)/a(t_0)$. For increasing $a(t)$, this corresponds to a *redshift*, i.e. a decrease in frequency with a corresponding increase in wavelength by a factor commonly denoted:

$$1 + z = \frac{a(t_0)}{a(t_s)}. \tag{2.9}$$

If we choose to set $a(t_0) = 1$ today, the scale factor a is given in terms of the redshift as

$$a = \frac{1}{1+z}. \tag{2.10}$$

Astronomers have a preference for expressing results using redshift and we will try to follow this convention for our plots in this report. For sources in proximity to the observer, one can expand $a(t)$ in a power series

$$a(t) \approx a(t_0)\left[1 + (t - t_0) H_0 + \ldots\right], \tag{2.11}$$

where H_0 is known as the *Hubble constant*,

$$H_0 = \frac{\dot{a}(t_0)}{a(t_0)}, \tag{2.12}$$

which measures the expansion rate of the universe today. We will use this factor to normalize our equations in the chapters that follow and this has also been the case historically, and hence H_0 is measured somewhat obtrusively in units $H_0 = 100\,h\,\mathrm{km\,s^{-1}Mpc^{-1}}$ where the dimensional scaling factor h incorporated the uncertainties in the measurement. Currently, the best estimation of is $h \approx 0.678 \pm 0.077$ [21]. A parsec (abbreviated pc), is a unit of distance derived from the theoretical annual parallax of one arc-second (measured as the inverse of the parallax). A parallax is the apparent measured distance in the position of a celestial object as seen by one observer on the Earth and another hypothetical one on the Sun. Roughly, a parsec is 3.26 lightyears. Concerning large-scale structures, they are of the order of hundreds of megaparsecs. At an arbitrary time, the expansion rate is denoted $H(t) = \dot{a}(t)/a(t)$. We will now use this to introduce a central concept in this work, the *horizon*. Before we proceed we have to provide a brief interlude to introduce a more fitting temporal parametrization *conformal time*. Consider a photon which travel radially outward from the source in FLRW geometry with $k = 0$[2],

$$u^\mu u_\mu = 0 \rightarrow -(u^t)^2 + a(t)^2 (u^i)^2 = 0, \tag{2.13}$$

with 4-velocity u^μ. For a purely radial motion

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{u^r}{u^t} = \pm\frac{1}{a}, \tag{2.14}$$

[2] This line of reasoning can easily be translated into geometries where $k \neq 0$ by changing the parametrization of the spatial part of the metric. However, we will not present any case in this report where k is different from zero, which allows us to make this shortcut through the derivation, avoiding extra notation.

Choosing the direction outward from the source, $-$, one arrives at

$$r_i = \int_{t_i}^{t_0} \frac{\mathrm{d}t}{a}, \tag{2.15}$$

since $r_0 = 0$. This propels us to define conformal time as

$$\eta(t) \stackrel{\text{def}}{=} \int \frac{\mathrm{d}t}{a}, \text{ i.e. } \mathrm{d}\eta = a\,\mathrm{d}t, \tag{2.16}$$

so that $r_i = \eta_0 - \eta_i$. Now, imagine if we were to move the point t_i back in time to the Big Bang, equation (2.16) will represent the maximum distance wherein signals traveling at the speed of light may reach an observer at a later time t_0. This is known as the *comoving particle horizon*. Moreover, we can rewrite this integral

$$\int_{t_i}^{t} \frac{\mathrm{d}t}{a} = \int_{a_i}^{a} \frac{\mathrm{d}a}{\dot{a}a} = \int_{a_i}^{a} (aH)^{-1}\,\mathrm{d}\ln a, \tag{2.17}$$

where the term $(aH)^{-1}$ is the *comoving Hubble radius*. In normal, standard cosmology, one obtains that the integral is proportional to the comoving Hubble radius, and both are confusingly called "the horizon". Still, when we move beyond standard cosmology into the theoretical regime of inflation, this relation will not longer hold, instead $r_i \gg (aH)^{-1}$, implying that particles which were initially causally connected to one another become separated. For the perturbations which we will examine, this means that wave modes which were connected, below the horizon, $(aH)^{-1}$, can be pushed outside the horizon since $(aH)^{-1}$ changes with time. We will learn more about this in section 2.2.

From the background cosmology in ΛCDM we will now calculate the Friedmann equations by inserting the metric ansatz (2.4) into (2.3) with no assumption on k, calculating the Christoffel symbols through

$$\Gamma^{\mu}_{\rho\sigma} = \frac{1}{2} g^{\mu\nu} \left(\partial_\rho g_{\sigma\nu} + \partial_\sigma g_{\rho\nu} - \partial_\nu g_{\rho\sigma} \right), \tag{2.18}$$

which produce the components of the Ricci tensor

$$R_{\mu\nu} = \partial_\rho \Gamma^{\rho}_{\mu\nu} - \partial_\nu \Gamma^{\rho}_{\mu\rho} + \Gamma^{\rho}_{\sigma\rho}\Gamma^{\sigma}_{\mu\nu} - \Gamma^{\rho}_{\sigma\nu}\Gamma^{\sigma}_{\mu\rho}. \tag{2.19}$$

Firstly, we determine which components of the Ricci tensor that could be set to zero based on the homogeneity and isotropy of the metric. For a comoving observer, the Ricci tensor has in total 10 components, of which three are time-space (zero due to the isotropy), 1 time-time and 6 space-space. The space-space components are further restricted to the diagonal by isotropy; they must be the same for the x-, y- and z-directions and could only depend on time due to homogeneity. We denote this time-dependent function L. Hence we are only interested in the R_{tt} and R_{ii} components, which are given for a co-moving observer in FLRW as

$$\begin{aligned} R_{tt} &= -3\frac{\ddot{a}}{a}, \\ R_{ii} &= a\ddot{a} + 2\dot{a}^2 + 2K. \end{aligned} \tag{2.20}$$

This implies that R_{ii} can be expressed in terms of L through

$$L = \frac{R_{ii}}{a^2}. \tag{2.21}$$

The scalar curvature is given by taking the trace of $R_{\mu\nu}$, which yields $R = -R_{tt} + 3L$. Hence, the Einstein tensor is given by

$$G_{\mu\nu} \overset{\text{def}}{=} R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{1}{2}\begin{pmatrix} R_{tt}+3L & & & \\ & R_{tt}-L & & \\ & & R_{tt}-L & \\ & & & R_{tt}-L \end{pmatrix}. \tag{2.22}$$

On the matter side of the equal sign in (2.3) we introduce a stress-energy tensor in the perfect fluid-form, namely

$$T_{\mu\nu} = (\bar{\rho} + \bar{p})u_\mu u_\nu + \bar{p}g_{\mu\nu}, \tag{2.23}$$

where $\bar{\rho}$ is the (unperturbed) energy density of the fluid and $\bar{p}$ the (unperturbed) spatial pressure. Here, we do neither have anisotropic stress nor momentum flows and hence $T_{\mu\nu}$ only has none-trivial elements on the diagonal. For the 4-velocity of the fluid, u^μ, we have in cosmic time that $g^{\mu\nu}u_\mu u_\nu = -1$. The cosmological constant term can be incorporated as a density $\rho_\Lambda \overset{\text{def}}{=} \Lambda/M_{\text{PL}}^2$ appearing on the diagonal $\rho_\Lambda g_{\mu\nu}$. Einstein's equations then give

$$\begin{aligned} \frac{1}{M_{\text{PL}}^2}\bar{\rho} &= \frac{1}{2}\left(R_{tt} + 3L\right), \\ \frac{1}{M_{\text{PL}}^2}\bar{p} &= \frac{1}{2}\left(R_{tt} - L\right). \end{aligned} \tag{2.24}$$

Dividing the first of these with a third, we obtain

$$\frac{1}{3}\frac{1}{M_{\text{PL}}^2}\bar{\rho} = \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = H^2 + \frac{k}{a^2}, \tag{2.25}$$

and if we take $-1/6$ times the $\bar{\rho}$ equation together with $-1/3$ times of the $\bar{p}$ equation, we arrive at

$$-\frac{1}{3}\frac{1}{2M_{\text{PL}}^2}(\bar{\rho} + 3\bar{p}) = \frac{\ddot{a}}{a}. \tag{2.26}$$

These two are called the *Friedmann equations.* Moving on, the continuity equation is given by the second Bianchi identity applied on the Einstein tensor. For the stress-energy tensor, this implies

$$\nabla_\mu T^{\mu\nu} = \partial_\mu T^{\mu\nu} + \Gamma^\mu_{\mu\rho}T^{\rho\nu} + \Gamma^\nu_{\mu\rho}T^{\mu\rho} = 0, \tag{2.27}$$

which after a few calculations[3] results in the continuity equation

$$\dot{\bar{\rho}} + 3H(\bar{\rho} + \bar{p}) = 0. \tag{2.28}$$

For fluids without pressure, such as pressure-less dust, this is reduced to

$$\dot{\bar{\rho}} = -3H\bar{\rho}. \tag{2.29}$$

In cosmology, one usually assumes that the ratio between $\bar{\rho}$ and $\bar{p}$ is constant, yielding the equation of state

[3]In this work, these have been carried out through the construction of a dedicated `Mathematica` script.

$$\omega = \frac{p}{\bar{\rho}}. \tag{2.30}$$

If we place this in the continuity equation (2.30), one obtains

$$\dot{\bar{\rho}} = -3H(1+\omega)\bar{\rho} = -3(1+\omega)\bar{\rho}\frac{\dot{a}}{a}, \tag{2.31}$$

whose solution for the energy density is

$$\bar{\rho} = \bar{\rho}_0 a^{-3(1+\omega)}, \tag{2.32}$$

where $\bar{\rho}_0$ is the energy density at $t = 0$ (today). There are a few interesting cases concerning this equation:

(i) **Dark matter or baryons:** A fitting model is that of *non-relativistic matter*, i.e. when $|\bar{p}| \ll \bar{\rho}$. Here, one is allowed to use Newtonian physics to describe small perturbations from the background solution, given small velocities and subhorizon scales. The redshifts are proportional to a^{-3}, as $\bar{\rho} \approx \rho_0 a^{-3}$.

(ii) **Radiation:** For photons, or neutrinos as well in the early universe, the energy density is $\bar{\rho} = 3\bar{p}$, which gives $\omega = 1/3$ and then resulting redshift goes as a^{-4}.

(iii) **Cosmological constant:** From the relation $\rho_\Lambda g_{\mu\nu}$, reminding us of the minus sign in the time-component, we see that $\omega = -1$ and hence we have a constant energy density.

To conclude, it is advantageous to introduce a *critical density*, i.e. at which density the universe becomes flat ($k = 0$). According to the first Friedmann equation, (2.25), this occurs at

$$\bar{\rho}_{\text{crit}} = 3H^2 M_{\text{PL}}^2. \tag{2.33}$$

Proceeding, one can now define a convenient density parameter for each constituent χ in the universe,

$$\Omega_\chi \stackrel{\text{def}}{=} \frac{\rho_\chi}{\bar{\rho}_{\text{crit}}}. \tag{2.34}$$

For the geometry, one can define a coefficient as

$$\Omega_k \stackrel{\text{def}}{=} -\frac{k}{a^2 H^2}. \tag{2.35}$$

Caution must be taken since this coefficient has nothing to do with density and is only introduced to simplify the expressions. Equation (2.25) tells us that the different components in the universe today satisfy

$$\Omega_\Lambda + \Omega_m + \Omega_\gamma + \Omega_k = \left(\frac{H}{H_0}\right)^2\bigg|_{a=1} = 1, \tag{2.36}$$

with γ denoting radiation. The best present measurements from the 2013 Planck survey for Ω_Λ and Ω_m, decomposed into baryonic matter and a dark matter components, are illustrated in figure 2.2. As for the residual two, $\Omega_\gamma \sim 10^{-5}$ and $\Omega_k < 10^{-3}$. Hence, they are not

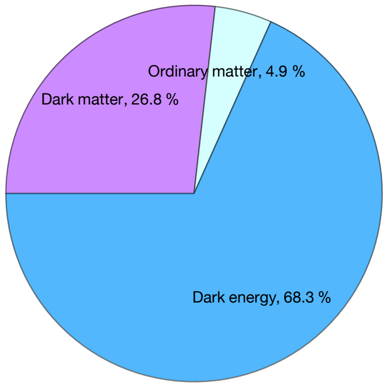


Figure 2.2: Percentages of the constituents of the universe today as measured with and published by Planck in 2013 [21].

important in the contemporary universe, but the radiation component played an important part earlier on, as one can surmise from $\bar{\rho}_\gamma \propto a^{-4}$. At an arbitrary previous time, we include redshift factors in (2.36) and rewrite them as fractions of the critical density as

$$\bar{\rho}(a) = \rho_\Lambda + \bar{\rho}_m a^{-3} + \bar{\rho}_\gamma a^{-4} = \bar{\rho}_{\text{crit}} \left(\Omega_\Lambda + \Omega_m a^{-3} + \Omega_\gamma a^{-4} \right), \tag{2.37}$$

which leads to

$$H^2 = H_0^2 \left(\Omega_\Lambda + \Omega_m a^{-3} + \Omega_\gamma a^{-4} + \Omega_k a^{-2} \right). \tag{2.38}$$

Here, it is common to define an energy function

$$\mathcal{E}(a)^2 \stackrel{\text{def}}{=} \Omega_\Lambda + \Omega_m a^{-3} + \Omega_\gamma a^{-4} + \Omega_k a^{-2}, \tag{2.39}$$

which contains the evolving part of the expansion factor. In redshift language, this means that the following relation

$$H = H_0 \mathcal{E}(z), \tag{2.40}$$

emerges. We will use that $\mathcal{E}(z=0) = \mathcal{E}(a=1) = 1$ in our normalization schemes during the analysis of bimetric gravity in section 4.2. To continue, we will start investigating how one treats small perturbations of this ideal background solution in general relativity.

2.2 Cosmic perturbations

To examine small deviations from the background solution, we restrict ourselves to linear, first-order perturbations as we mentioned in the introduction. We assume that they are small and hence we are able to neglect terms of higher order. This section is inspired by the treatments in [22], [23], [24], [25], [26], [27] and [28]. After having decomposed the perturbation fields, $\delta\phi$, in terms of their Fourier components, i.e.

$$\delta\phi(t, \boldsymbol{r}) = \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \delta\phi_{\boldsymbol{k}}(t), \tag{2.41}$$

one can examine the individual modes with wavenumber k using approximations on different scales[4] and different times as

(i) **Subhorizon, $k \gg aH$:** These are modes with wavelengths within the comoving Hubble radius. As they are remote from scales where curvature becomes important, we can use *Newtonian perturbation theory* to examine them and the ordinary equations of non-relativistic fluid dynamics. All cosmological relevant perturbations, those that we may observe, originate from this regime. However, whereas k remains unaltered through the eras, the comoving Hubble radius changes. Specifically, for theories of inflation, the Hubble radius shrinks during the rapid expansion, which pushes the subhorizon modes into the superhorizon region.

(ii) **Superhorizon, $k \ll aH$:** We are dealing with perturbations whose wavelengths start to approach the curvature scale of the universe set by H. Hence, we cannot neglect relativistic effects and have to use *relativistic perturbation theory*[5].

(iii) **Intermediate region, $k \sim aH$:** For perturbations with wavelengths close to the horizon, we must consider all terms and not make any approximations.

Effectively, all regimes can be described by relativistic perturbation theory where the subhorizon regime is accessed by taking the limit $aH \to 0$ and the superhorizon through $k \to 0$. As we remarked upon briefly, the comoving Hubble radius is something which changes overtime. In an inflationary model, where the expansion is governed by a primordial scalar field $\varphi(t)$, the fluctuations $\delta\varphi_{\boldsymbol{k}} = \delta\varphi_{\boldsymbol{k}}(t,\, \boldsymbol{x})$ can be described through a nonzero quantum-mechanical variance,

$$\left\langle |\delta\varphi_{\boldsymbol{k}}|^2 \right\rangle \overset{\text{def}}{=} \left\langle 0 \left| |\delta\varphi_{\boldsymbol{k}}|^2 \right| 0 \right\rangle, \tag{2.42}$$

around a zero-average, arising from the uncertainty principle. In practice, the examination of the field at this level is often pseudo-classical, i.e. the field is quantized but the gravitational background remains classical. Of particular interest is the variance at the horizon crossing as the universe expands. At this point, the quantum fluctuations are driven into the classical regime, and the quantum expectation values are matched to the ensemble average of a classical stochastic field. Here, it is convenient to switch parametrization of the perturbations from fluctuations of the inflaton field to fluctuations in the *comoving curvature*, $\mathcal{R}_{\boldsymbol{k}}$. This is the perturbation of the intrinsic 3-curvature on hypersurfaces of constant time $\Sigma(t) = t$, evaluated in the *comoving gauge*, see equation (2.96). It turns out that $\delta\dot{\mathcal{R}}_{\boldsymbol{k}}$, obeying the relation in (2.98), is conserved in the superhorizon domain in general relativity for perturbations with negligible non-adiabatic pressure, which means that we can relate the theoretical conditions at the horizon exit which is characterized by high energies to those after horizon re-entry at low energies. Perturbations whose equation of state can be written as $p = p(\rho)$ are called *adiabatic perturbations*. Such perturbations' state at some point in the perturbed spacetime $(t,\, \boldsymbol{x})$ is the same as in the background unperturbed universe at a slightly different time

[4] The method to treat the perturbations also depends on which kind of matter one analyzes; the Newtonian approximation is valid for matter characterized by nonrelativistic pressures and velocities whereas relativistic effects must be taken into account to describe photons and neutrinos.

[5] More accurately, we are doing *cosmic relativistic perturbation theory*, since we are describing small deviations from the underlying isotropic and homogeneous background solution, and using its properties.

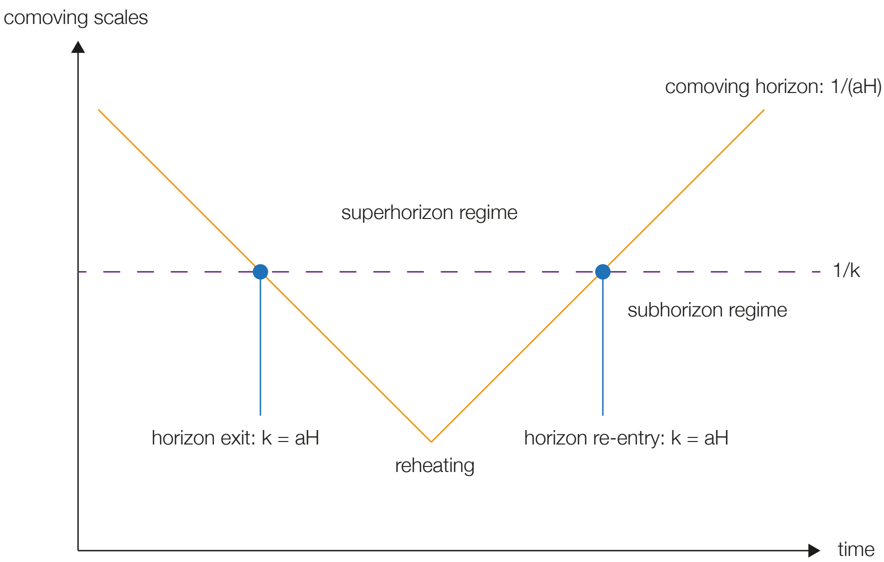


Figure 2.3: Cosmologically relevant perturbations were generated inside the horizon, below the line $(aH)^{-1}$. They are modeled as quantum fluctuations of some scalar inflationary field just before the horizon exit. As the universe expanded they were forced into the superhorizon regime above the line, where the conservation of the comoving curvature perturbation $\dot{\mathcal{R}} \approx 0$ froze them in, and at some late time epoch they will once again be subhorizon and start to evolve.

$t + \delta t(\boldsymbol{x})$. This implies that parts in the perturbed universe are before/behind the evolution of the background universe. From the equation of state,

$$p(\rho) = \bar{p}(\bar{\rho}) + \frac{\mathrm{d}p}{\mathrm{d}\rho}(\bar{\rho})\delta\rho \Rightarrow \delta p = \frac{\mathrm{d}p}{\mathrm{d}\rho}(\bar{\rho})\delta\rho \Rightarrow \dot{\bar{p}} = \frac{\mathrm{d}p}{\mathrm{d}\rho}\dot{\bar{\rho}}, \tag{2.43}$$

from which we infer the relation

$$\frac{\dot{\bar{p}}}{\dot{\bar{\rho}}} = \frac{\delta p}{\delta \rho} = c_s^2, \tag{2.44}$$

where c_s is the sound-speed defined through

$$c_s^2 \stackrel{\text{def}}{=} \left(\frac{\partial p}{\partial \rho}\right)_S, \tag{2.45}$$

with the entropy S kept fixed (no entropy is produced during the expansion). Since single-field inflation yields adiabatic perturbations, we will focus on these in the subsequent analysis. Moreover, since our equation of state $p = \omega\rho$, it means that $c_s^2 = \omega = \text{const.}$ for adiabatic perturbations. Thus $c_s^2 = \omega = 0$ designates the matter era and $c_s^2 = \omega = \frac{1}{3}$ the radiation era. Between the re-entry and today, it is possible to compute the evolution of the perturbations. Hence, one may for some quantities bridge the gap between today and the primordial universe.

2.2.1 Subhorizon perturbations

In this report, we are chiefly interested in superhorizon perturbations, and we will not digress much into the domain of Newtonian perturbation theory. The main idea is to treat matter as a fluid and use the ordinary fluid mechanics equations

$$\begin{aligned} \frac{\partial \bar{\rho}}{\partial t} + \nabla \cdot (\bar{\rho}\boldsymbol{u}) &= 0, \\ \frac{\partial \bar{\boldsymbol{u}}}{\partial t} + (\bar{\boldsymbol{u}} \cdot \nabla)\,\bar{\boldsymbol{u}} + \frac{1}{\bar{\rho}}\nabla\bar{p} + \nabla\bar{\Phi} &= 0, \\ \Delta\bar{\Phi} - 4\pi G\bar{\rho} &= 0, \end{aligned} \tag{2.46}$$

where $\bar{\boldsymbol{u}}$ is the 3-velocity of the fluid and $(\partial_t + \bar{\boldsymbol{u}} \cdot \nabla)\bar{\boldsymbol{u}}$ is the convective time derivative, which is an adapted derivative that follows a fluid element as it moves. Then, we add linear perturbations to these variables, $\bar{\rho} \to \bar{\rho} + \delta\rho$, $\bar{p} \to \bar{p} + \delta p$, $\bar{\Phi} \to \bar{\Phi} + \delta\Phi$, $\bar{\boldsymbol{u}} \to \bar{\boldsymbol{u}} + \delta\boldsymbol{u}$ and replace the quantities in the system to obtain the corresponding equations involving only the perturbations. In an expanding background universe, there is a small subtlety involved when translating between the position in terms of physical coordinates $\boldsymbol{r}$ and in comoving coordinates $\boldsymbol{x}$. They are related to one another through

$$\boldsymbol{r} = \frac{a(t)}{a_0}\boldsymbol{x} \Rightarrow \boldsymbol{u} = \boldsymbol{v} + H\boldsymbol{r}, \tag{2.47}$$

where the last relation relates the velocities in the two systems, with the proper velocity in comoving coordinates as $\boldsymbol{v} = a\dot{\boldsymbol{x}}$. If we normalize our equations we can replace $a_0 = 1$. Moreover, we would like to locate an expression for the gradient in comoving coordinates which acknowledges the spatial expansion. At a given time t, the two gradients are related as

$$\nabla_{\boldsymbol{r}} = a^{-1}\nabla_{\boldsymbol{x}}, \tag{2.48}$$

which is easily deduced from (2.47). Henceforth, we will abbreviate $\nabla_{\boldsymbol{x}}$ as $\nabla \stackrel{\text{def}}{=} \nabla_{\boldsymbol{x}}$. Now, if we plug in perturbations expressed in these coordinates into the system and perform a Fourier transform according to (2.41), a short calculation yields

$$\dot{\delta}_{\boldsymbol{k}} - ia^{-1}\boldsymbol{k} \cdot \boldsymbol{v}_{\boldsymbol{k}} = 0, \tag{2.49}$$

$$\dot{\boldsymbol{v}}_{\boldsymbol{k}} + H\boldsymbol{v}_{\boldsymbol{k}} - ia^{-1}\boldsymbol{k}\left[\frac{\delta p}{\bar{\rho}} + \delta\Phi_{\boldsymbol{k}}\right] = 0, \tag{2.50}$$

$$-\boldsymbol{k}^2\delta\Phi_{\boldsymbol{k}} = 4\pi G a^2 \bar{\rho}\delta_{\boldsymbol{k}}, \tag{2.51}$$

where we have used the contrast density $\delta_{\boldsymbol{k}}$ instead of $\delta\rho_{\boldsymbol{k}}$ and the continuity equation $\dot{\bar{\rho}} = -3H\bar{\rho}$ from the zero:th order components. One may combine these by plugging in $\partial_t \cdot$ (2.49) and its non-differentiated version into (2.50) with (2.51) to produce *Jeans' stability equation*,

$$\ddot{\delta}_{\boldsymbol{k}} + 2H\dot{\delta}_{\boldsymbol{k}} + \left[\left(\frac{k}{a}\right)^2 \frac{\delta p}{\bar{\rho}} - 4\pi G\bar{\rho}\right]\delta_{\boldsymbol{k}} = 0. \tag{2.52}$$

For adiabatic perturbations, one may use (2.44) to recast this expression as

$$\ddot{\delta}_{\boldsymbol{k}} + 2H\dot{\delta}_{\boldsymbol{k}} + \left[\left(\frac{k}{a}\right)^2 c_s^2 - 4\pi G\bar{\rho}\right]\delta_{\boldsymbol{k}} = 0. \tag{2.53}$$

The wavenumber for which the total bracket equals zero is known as the *Jeans' wavenumber* k_J which has a corresponding *Jeans' wavelength* λ_J,

$$k_J \stackrel{\text{def}}{=} \frac{a}{c_s}\sqrt{4\pi G\bar{\rho}}, \;\; \lambda_J \stackrel{\text{def}}{=} \frac{2\pi}{k_J}. \tag{2.54}$$

For $k \gg k_J$, we can approximate Jeans' equation by omitting the Poisson term, which yields oscillating solutions which are dampened by $2H\dot{\delta}_{\boldsymbol{k}}$. Hence, there will be no growing solution and no structure formation sub-Jeans. Above this limit, however, for scales significantly larger than the Jeans' length, where we have the relations $k_J \gg k \gg H$, we can instead neglect the sound-speed term. As a consequence, the solution will involve fluctuations subjected to a power-law growth, which incite structure formation.

Granted, these last few equations are only valid for scalar velocities. Following Helmholtz' theorem, we can as usual decompose $\boldsymbol{v}$ into a transverse, curl-less component and a longitudinal, divergence-less component. In the full system, the longitudinal component only appear in the time-derivative term of (2.50). Its solution is obviously proportional to a^{-1}, i.e. these perturbations decay and are not considered pertinent in the structure formation context. This ends our overview of Newtonian perturbation theory.

2.2.2 General relativistic treatment

In the fully general case, suitable for the superhorizon and intermediate region and for relativistic matter, we start by adding a small perturbation to the background metric,

$$\bar{g}_{\mu\nu} \to g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu}, \tag{2.55}$$

where $\bar{g}_{\mu\nu}$ is the unperturbed isotropic and homogeneous background FLRW metric. Owing to these spatial symmetries of the background metric, it is possible to decompose the perturbations into fields of scalars, vectors or tensors type based on how they transform, *SVT* *decomposition.* These components will not mix at the linear level, facilitating the examination. The most general ansatz of the perturbed line element can be written as [22],

$$\mathrm{d}s^2_{\bar{g}+\delta g} = -N(t)^2\,(1+E)\,\mathrm{d}t^2 + 2N(t)a(t)F_i\,\mathrm{d}t\,\mathrm{d}x^i + a(t)^2\,[\delta_{ij}+\delta g_{ij}]\,\mathrm{d}x^i\,\mathrm{d}x^j, \tag{2.56}$$

with $N(t) = 1,\, a(t)$ for cosmic and conformal time respectively and where the perturbation fields depend on both space and time. As previously, we split the 3-vector $\boldsymbol{F}$ into a divergenceless and a curl-free part,

$$F_i = \partial_i F + G_i,\;\; \partial^i G_i = 0, \tag{2.57}$$

whose first (curl-free) constituent is the scalar part. Equivalently, any rank-2 tensor δg_{ij} can be decomposed[6] as

$$\delta g_{ij} = A\delta_{ij} + \partial_i\partial_j B + 2\partial_{(i}C_{j)} + D_{ij}, \tag{2.58}$$

where the round parentheses announce a symmetrization over the indices i, j. Here, the fields C_i and D_{ij} contain the vector and tensor parts respectively. Both of these are divergenceless, $D_{ij} = D_{ji}$ and the tensor perturbation is traceless $D_{ii} = 0$. This implies that G_i, C_i and D_{ij} carry two degrees of freedom each. In total, the ten degrees of freedom of the metric perturbation $\delta g_{\mu\nu}$ are decomposed as

$$\delta g_{\mu\nu} = \{E,\, F_i,\, \delta g_{ij}\} = \{E,\, (F,\, G_i)\,,\, (A,\, B,\, C_i,\, D_{ij})\}\,, \tag{2.59}$$

where the degrees of freedom are distributed as $1, (1,\, 2), (1, 1, 2, 2)$. One can decompose the perturbed stress-energy tensor, $\bar{T}_{\mu\nu} \to T_{\mu\nu} = \bar{T}_{\mu\nu} + \delta T_{\mu\nu}$, in a similar approach. The perturbed stress-energy tensor part is given adding small perturbations to (2.23) and keeping terms of linear order as

$$\delta T^{\mu}{}_{\nu} = (\delta\rho + \delta P)\bar{u}^{\mu}\bar{u}_{\nu} + (\bar{\rho} + \bar{P})(\delta u^{\mu}\bar{u}_{\nu} + \bar{u}^{\mu}\delta u_{\nu}) - \delta P\delta^{\mu}{}_{\nu} - \Sigma^{\mu}{}_{\nu}, \tag{2.60}$$

whose last term incorporates anisotropic inertia, denoted Σ, which measures the deviation from the perfect fluid form. It is possible to define it such that $u^{\mu}\Sigma_{\mu\nu} = 0$, which implies that such terms only appear in the spatial part of the stress-energy tensor, where the may be decomposed into scalar Σ_{S}, vectorial Σ_{V} and tensorial components Σ_{T}. Moreover, one could redefine the pressure by including the trace $\Sigma^{i}{}_{i}$, meaning that one can demand that $\Sigma^{i}{}_{i} = 0$. The normalization condition on the 4-velocity yields at linear order that

$$\delta g_{\mu\nu}\bar{u}^{\mu}\bar{u}^{\nu} + 2\bar{u}_{\mu}\delta u^{\mu} = 0, \tag{2.61}$$

which leads to the conclusion that

$$\delta u_0 \propto E, \tag{2.62}$$

whilst the spatial part $\delta\boldsymbol{u}$ is an independent, dynamical variable. As in (2.57), $\delta\boldsymbol{u}$ can be split into a scalar and a vectorial part,

$$\delta u_i = \partial_i \delta u + \delta u_i^{\mathrm{V}},\;\; \partial^i \delta u_i^{\mathrm{V}} = 0, \tag{2.63}$$

[6] We mainly adhere to the convention established in [22]. Some authors, [23], define the fields with an extra factor of two in front for later convenience in equations (2.80) and (2.81). In addition, others [24], may write the scalar component B in such a manner that the trace $\mathrm{Tr}\,\delta g_{ij}$ features a term $2B$.

where u_i will receive contributions from (2.57) as we lower the indices in (2.63) with the perturbed metric tensor. The scalar perturbed stress-energy components for a perfect fluid are acquired by inserting this into (2.23) and are

$$\begin{aligned} T^0_{\ 0} &= -\bar{\rho}(1+\delta), \\ T^i_{\ 0} &= -\left(\bar{\rho}+\bar{p}\right) u^i, \\ T^0_{\ i} &= \left(\bar{\rho}+\bar{p}\right)\left(u_i+\partial_i F\right), \\ T^i_{\ j} &= \left(\bar{p}+\delta p\right)\delta^i_{\ j}+\Sigma^i_{\ j}, \end{aligned} \tag{2.64}$$

where u^i are the components of $\boldsymbol{u} = \mathrm{d}\boldsymbol{x}/\,\mathrm{d}t$ and $\Sigma^i_{\ j}$ is anisotropic stress, $\Sigma^i_{\ i} = 0$. If we specify that we are dealing with pressure-less dust, $\bar{p} = \delta p = \Sigma^i_{\ j} = 0$. To obtain the tensor and vectorial parts, we simply change the velocities and the F-terms as well as the terms in $\Sigma^i_{\ j}$.

Before we proceed and insert our perturbed metric into Einstein's equations to acquire the equations of motion for the linear perturbations fields in (2.59), we must deal with the subtle transformation properties of these variables. Suppose that we simply shifted our spatial coordinates slightly, $x^i \to x^i + \xi^\mu(t,\,\boldsymbol{x})$ with a small transformation ξ and calculated the line element in (2.56). If the shift is small, it is possible to treat it as a perturbation. Expressed in terms of $\mathrm{d}x^i$ the perturbation becomes $\mathrm{d}x^i = \mathrm{d}\tilde{x}^i - \partial_t\xi^i\,\mathrm{d}t - \partial_j\xi^j\,\mathrm{d}x^i$. At linear order, this introduces uninvited extra components in the spatial parts. Yet, these do not represent the physical degrees of freedom of the theory, as they disappear when we perform an inverse transform. Such artificial fields are known as *gauge modes*. They arise since our ansätze of perturbation fields do not abide by the underlying fundamental symmetry of general relativity, namely *general covariance* also known as *diffeomorphism invariance*. To remind the reader; this means that the form of the physical laws of the theory is the same in all coordinate systems. Choosing a specific coordinate system, a special foliation of spacetime, can be thought of as picking a *gauge*, echoing the terminology of electromagnetism and Yang-Mills theories. The theory comes with a redundant description and we need to associate a metric transformation to $\delta g_{\mu\nu}$, a *gauge transformation*, which allows it to change in the same manner as the unperturbed metric $\bar{g}_{\mu\nu}$ subjected to the same transformation.

These gauge transformations can be expressed *passively* (coordinate system approach) or *actively* (pure diffeomorphism mapping approach) for the perturbations following [23]. For the passive case one considers a background manifold $\mathcal{M}$ (spacetime) where one has a chosen coordinate system x^ρ and arbitrary functions $Q\left(x^\rho\right)$ with a fixed dependence on the coordinates. On this manifold it is possible to introduce a supplemental coordinate system $\tilde{x}^\rho$ with associated functions $\tilde{Q}\left(\tilde{x}^\rho\right)$, which also have a fixed coordinate dependence. Should we introduce a small perturbation δQ of the function Q in the coordinate system x^ρ at a point $p \in \mathcal{M}$, it can be expressed as

$$\delta Q(p) = Q\left(x^\rho\right) - \bar{Q}\left(x^\rho\right), \tag{2.65}$$

where the bar denotes the unperturbed function. In the second coordinate system, the perturbation could be written as

$$\widetilde{\delta Q}(p) = \tilde{Q}\left(x^\rho\right) - \bar{\tilde{Q}}\left(x^\rho\right). \tag{2.66}$$

A gauge transformation in this framework corresponds to switching from

$$\delta Q(p) \to \widetilde{\delta Q}(p), \text{ while } x^\rho \to \tilde{x}^\rho \text{ on } \mathcal{M}. \tag{2.67}$$

Viewed actively, one can instead envisage two manifolds; one physical manifold $\mathcal{M}$ and one background spacetime $\mathcal{N}$ with associated coordinates x^ρ_{b} rigidly fixed where "b" means background. One then introduces a diffeomorphism $\mathcal{D}: \mathcal{N} \to \mathcal{M}$, which induces a coordinate system on $\mathcal{M}$ through $\mathcal{D}: x^\rho_{\mathrm{b}} \to x^\rho$. For a chosen diffeomorphism, one can express a perturbation δQ of a function Q on $\mathcal{M}$ as

$$\delta Q(p) = Q(p) - \bar{Q}\left(\mathcal{D}^{-1}(p)\right), \text{ for } p \in \mathcal{M}. \tag{2.68}$$

Here $\bar{Q}$ is a fixed function defined on the background spacetime. In this formalism a gauge transformation corresponds to

$$\delta Q(p) \to \widetilde{\delta Q}(p), \text{ generated by a change of } \mathcal{D} \to \widetilde{\mathcal{D}}, \tag{2.69}$$

between the manifolds $\mathcal{N}$ and $\mathcal{M}$. With such a change in correspondence, an associated change of the induced coordinates on $\mathcal{M}$, $x^\rho \to \tilde{x}^\rho$ follows. The two different fashions can be interpreted as

(i) **Passive formalism:** One connects the gauge transformations with the choice of coordinate systems on $\mathcal{M}$, wherein the perturbations are expressed.

(ii) **Active formalism:** The amplitude of the perturbations depends on the correspondence between $\mathcal{N}$ and $\mathcal{M}$.

In both cases, an infinitesimal coordinate transformation with a 4-vector ξ,

$$x^\rho \to \tilde{x}^\rho = x^\rho + \xi^\rho(x), \tag{2.70}$$

brings about a change in δQ:

$$\Delta Q = \widetilde{\delta Q} - \delta Q = \mathscr{L}_\xi Q, \tag{2.71}$$

where $\mathscr{L}_\xi$ denotes the Lie derivative[7] of Q along ξ. If we set the function Q to be the metric $g_{\mu\nu}$, the transformation of the perturbation becomes

$$\delta g_{\mu\nu} \to \widetilde{\delta g}_{\mu\nu} = \delta g_{\mu\nu} + \mathscr{L}_\xi \bar{g}_{\mu\nu}, \tag{2.72}$$

with

$$\mathscr{L}_\xi \bar{g}_{\mu\nu} = -\xi^\rho \partial_\rho \bar{g}_{\mu\nu} - \partial_\mu \xi^\rho \bar{g}_{\rho\nu} - \partial_\nu \xi^\rho \bar{g}_{\mu\rho}. \tag{2.73}$$

This is not hard to prove if we use the invariance of the line element in the two coordinates systems and write out

$$\mathrm{d}s^2 = \tilde{g}_{\mu\nu}(\tilde{x})\, \mathrm{d}\tilde{x}^\mu\, \mathrm{d}\tilde{x}^\nu = g_{\rho\sigma}(x)\, \mathrm{d}x^\rho\, \mathrm{d}x^\sigma, \tag{2.74}$$

and then defines the 1-forms $\mathrm{d}x^\rho = \partial x^\rho / \partial \tilde{x}^\mu\, \mathrm{d}\tilde{x}^\mu$ and $\mathrm{d}x^\sigma$ in a similar fashion. Such a correspondence means that the other metric can be written as

[7] The interested reader can refer to chapter 2 of [29] for an excellent formal introduction to the Lie derivative as well as the context in the differential geometric foundations of general relativity.

$$\tilde{g}_{\mu\nu}(\tilde{x}) = \frac{\partial x^\rho}{\partial \tilde{x}^\mu}\frac{\partial x^\sigma}{\partial \tilde{x}^\nu} g_{\rho\sigma}(x) \,. \tag{2.75}$$

From this we infer that the coordinate change $x^\rho \to \tilde{x}^\rho$ yields a metric transformation $g_{\mu\nu}(x) \to \tilde{g}_{\mu\nu}(\tilde{x})$ with $\tilde{g}_{\mu\nu}(\tilde{x})$ according to (2.75). On the other hand, viewed actively, by setting the diffeomorphism as $\mathcal{D} = \phi$, the same transformation can be equally interpreted,

$$g_{\mu\nu}(x) \to \frac{\partial \phi^\alpha}{\partial x^\mu}\frac{\partial \phi^\beta}{\partial x^\nu} g_{\alpha\beta}(\phi(x)). \tag{2.76}$$

In this report, we will frequently discuss topics from a diffeomorphism perspective and this transformation rule in (2.76) will be our reference.

Expanding the lefthand side of (2.75) including terms linear in the infinitesimal coordinate transformation, yields

$$\tilde{g}_{\mu\nu}(\tilde{x}) = \bar{g}_{\mu\nu}(\tilde{x}) + \delta\tilde{g}_{\mu\nu} = \bar{g}_{\mu\nu}(x+\xi) + \delta\tilde{g}_{\mu\nu} = \bar{g}_{\mu\nu}(x) + \xi^\rho\partial_\rho\bar{g}_{\mu\nu}(x) + \delta\tilde{g}_{\mu\nu}. \tag{2.77}$$

Now, to rewrite the righthand side of (2.75), we seek an expression for the matrix $\partial x^\rho/\partial\tilde{x}^\mu$. Its inverse is easier to locate and follows from $\partial\tilde{x}^\rho/\partial x^\mu = \delta_\mu{}^\rho + \partial_\mu\xi^\rho$. Using the fact that we are dealing with a small perturbation added to the identity matrix which is its own inverse and neglecting ξ terms of higher order, we deduce that the inverse matrix must be $\partial x^\rho/\partial\tilde{x}^\mu = \delta_\mu{}^\rho - \partial_\mu\xi^\rho$. Hence the righthand side becomes

$$\left(\delta_\mu{}^\rho - \partial_\mu\xi^\rho\right)\left(\delta_\nu{}^\sigma - \partial_\nu\xi^\sigma\right)\left(\bar{g}_{\rho\sigma}(x) + \delta g_{\rho\sigma}\right) = \bar{g}_{\mu\nu}(x) + \delta g_{\mu\nu} - \partial_\mu\xi^\rho\bar{g}_{\rho\nu} - \partial_\nu\xi^\sigma\bar{g}_{\mu\sigma}, \tag{2.78}$$

keeping only terms linear in $\delta g_{\mu\nu}(x)$ and ξ. Combining this with the previous information from (2.77) give (2.73) [25] [30]. From this conclusion we surmise that $\mathscr{L}_\xi\bar{g}_{\mu\nu}$ will provide the sought transformation rules for $\Delta\delta g_{\mu\nu} \stackrel{\text{def}}{=} \mathscr{L}_\xi\bar{g}_{\mu\nu}$. Splitting the spatial part of the shift into a scalar and a vector part, $\xi^i = \partial^i\xi^S + \xi^V$, $\partial_i\xi^V = 0$, and analyzing (2.73) with our set of variables in (2.59), we find the following transformations for the scalar perturbations fields,

$$\Delta E = \frac{2}{N}\partial_t\left(\frac{\xi_0}{N}\right),\ \Delta A = \frac{2H}{N^2}\xi_0,\ \Delta B = -\frac{2}{a^2}\xi^S,\ \Delta F = -\frac{1}{Na}\left(\xi_0 + \dot{\xi}^S - 2H\xi^S\right), \tag{2.79}$$

where we have lowered the indices with the unperturbed metric, $\xi_0 = \bar{g}_{00}\xi^0$ etc. The tensor perturbations have the same symmetry as the background metric by default and will hence not be affected. The vector perturbations transform as $\Delta G_i = -1/Na\left(\dot{\xi}^V + 2H\xi^V\right)$ and $\Delta C_i = -1/a^2\xi^V$. To proceed and obtain a physical solution, we can either introduce combinations of these variables that are invariant under these transformations, *gauge-invariant variables*, or pick a specific gauge where no information is lost and the redundancy removed. By visual inspection of the form of the scalar transformation rules in (2.79) and by choosing to work in conformal time, we find two gauge-invariant combinations

$$\Phi \stackrel{\text{def}}{=} -A - H(2F - \dot{B}), \tag{2.80}$$

$$\Psi \stackrel{\text{def}}{=} E + H(2F - \dot{B}) + \partial_t(2F - \dot{B}), \tag{2.81}$$

which we can use to determine the system uniquely for the scalar perturbations. These are known as the *Bardeen variables.* For the stress-energy tensor, the same procedure applies, $\Delta\delta T^{\mu}{}_{\nu} = \mathscr{L}_{\xi}\bar{T}^{\mu}{}_{\nu}$.

If we calculate[8] the equations of motion for the linear scalar perturbations in (2.59) following (2.1) for Einstein's field equations with the line element in (2.56) with help from the connection in (2.18) and the Ricci tensor, followed by a Fourier transform we arrive at

- $\delta g^{(\mathrm{s})}_{0-0}$:

$$\frac{3H}{N^2}\left(HE - \dot{A}\right) - k^2\left[\frac{A}{a^2} + \frac{H}{N}\left(\frac{2F}{a} - \frac{\dot{B}}{N}\right)\right] = \frac{1}{M_g^2}\delta T^{0}{}_{0}\,, \tag{2.82}$$

- $\delta g^{(\mathrm{s})}_{0-i}$:

$$-\,ik\frac{1}{N^2}\left(HE - \dot{A}\right) = \frac{1}{M_g^2}\delta T^{0}{}_{i}\,, \tag{2.83}$$

- $\delta g^{(\mathrm{s})}_{i-i}$ (spatial trace components, no sum implied):

$$\begin{aligned}&\frac{1}{N^2}\left[\left(2\dot{H} + 3H^2 - 2\frac{\dot{N}}{N}H\right)E + H\dot{E} - \ddot{A} - 3H\dot{A} + \frac{\dot{N}}{N}\dot{A}\right] + \\ &-k^2\left[\frac{A+E}{a^2} + \frac{H}{N}\left(\frac{4F}{a} - \frac{3\dot{B}}{N}\right) + \frac{2\dot{F}}{aN} - \frac{1}{N^2}\left(\ddot{B} - \frac{\dot{N}}{N}\dot{B}\right)\right] = \frac{1}{M_g^2}\delta T^{i}{}_{i}\,,\end{aligned} \tag{2.84}$$

- $\delta g^{(\mathrm{s})}_{i-j}$ (off-diagonal components):

$$\frac{k^2}{2}\left[\frac{A+E}{a^2} + \frac{H}{N}\left(\frac{4F}{a} - \frac{3\dot{B}}{N}\right) + \frac{2\dot{F}}{aN} - \frac{1}{N^2}\left(\ddot{B} - \frac{\dot{N}}{N}\dot{B}\right)\right] = \frac{1}{M_g^2}\delta T^{i}{}_{j}\,. \tag{2.85}$$

For pressure-less dust, neither anisotropic nor isotropic pressure are present and the terms $\delta T^{i}{}_{j} = \delta T^{i}{}_{i} = 0$ vanish. This allows us to evaluate the second line of equation (2.84) to zero using equation (2.85). Here, we have assumed that the magnitude k is isotropic. To continue, we can either rewrite the system in terms of gauge-invariant variables or picking a suitable gauge. We will work this out in the most common *Newton's gauge.* In this setup, where one takes advantage of the gauge freedom to set ξ^S and ξ_0 such that $B = 0$ and $F = 0$ respectively, all off-diagonal components in the perturbed metric tensor vanish and the Bardeen variables reduce to an agreeable form; $\Phi = -A$ and $\Psi = E$. To be more in line with current conventions we will shift these variables $\Phi \to 2\Phi$, $\Psi \to 2\Psi$ in the line element (2.56) which for these fields in Newton's gauge in conformal time is written

$$\mathrm{d}s^2_{\bar{g}+\delta g} = a(t)^2\left[-\left(1+2\Psi\right)\mathrm{d}t^2 + \left(1-2\Phi\right)\delta_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j\right]. \tag{2.86}$$

In terms of physics, interpretations of results tend to be facilitated in this gauge. The choice $F = 0$ guarantees that hypersurfaces of constant time are orthogonal to the worldlines of observers at rest and from $B = 0$ we deduce that induced geometry on these hypersurfaces is isotropic. Returning to the equations of motion for the perturbations with (2.86) in mind, we immediately discern that (2.85) tells us that $\Phi = \Psi$ for no anisotropic stress, and the system in conformal time with the stress-energy components as in (2.64) is reduced to

[8] This was performed using a `Mathematica` script.

- $\delta g^{(\mathrm{s})}_{0-0}$:

$$\frac{3H}{a^2}\left(H\Phi+\dot{\Phi}\right)-k^2\frac{\Phi}{a^2}=-\frac{\delta\bar{\rho}}{2M_g^2}, \tag{2.87}$$

- $\delta g^{(\mathrm{s})}_{0-i}$:

$$-ik\frac{1}{a^2}\left(H\Phi+\dot{\Phi}\right)=-\frac{\bar{\rho}u^i}{2M_g^2}, \tag{2.88}$$

- $\delta g^{(\mathrm{s})}_{i-i}$ (spatial trace):

$$\left(2\dot{H}+H^2\right)\Phi+\ddot{\Phi}+3H\dot{\Phi}=0, \tag{2.89}$$

where we discover that the evolution of all matter perturbations is governed by the last equation (2.89). When we encounter the same problem in bimetric theory, this is the equation of interest which we want to find, together with (2.91). By combining (2.88) and (2.87), we obtain the corresponding Poisson equation,

$$k^2\Phi=\frac{a^2\bar{\rho}}{2M_g^2}\left(\delta+3H\frac{iu^i}{k}\right). \tag{2.90}$$

which one can compare with equation (2.51) in Newtonian perturbation theory, which was calculated in the comoving gauge where the velocity disappears. The corresponding equation for second-order differential equation for δ can be calculated with help from the Bianchi constraints, (2.27), described in the context of bimetric gravity in section 4.1 and explicitly presented in equations (4.43)-(4.44).

For adiabatic perturbations, where the sound speed is related to the perturbed pressure and density as in (2.44), the combination of (2.87) and (2.89), including its pressure term from (2.64), yields a close-form expression for the gravitational potential as

$$\ddot{\Phi}+3(1+\omega)H\dot{\Phi}-\omega k^2\Phi=0, \tag{2.91}$$

where the residual Φ-terms have been eliminated through the Friedmann equation (2.26). The superhorizon and the subhorizon approximations are easy to distinguish by taking either $k\to 0$ or $H\to 0$ in (2.91). In the superhorizon limit, $\ddot{\Phi}+3(1+\omega)H\dot{\Phi}=0$, and hence one plausible solution is $\Phi=\text{const}$. Solutions during the matter era or radiation era can be accessed by setting $\omega=0$ and $\omega=1/3$ respectively.

In order to calculate the comoving curvature perturbation, we require the induced metric on surfaces of constant time, which is the spatial part of equation (2.56). Since the quantity is a perturbation of scalar variable, the vector and tensor components drop out. With our perturbation fields in (2.59), this is

$$\gamma_{ij}\stackrel{\text{def}}{=}a^2\left[(1+A)\delta_{ij}+\partial_i\partial_j B\right]. \tag{2.92}$$

The intrinsic curvature, which is a concept that we will return to in section 3.3, on these surfaces is given by the three-dimensional Ricci scalar[9]. In component form, it is calculated through

$${}^{(3)}R=\gamma^{ij}\partial_k{}^{(3)}\Gamma^k_{ij}-\gamma^{ij}\partial_j{}^{(3)}\Gamma^k_{ik}+\gamma^{ij}{}^{(3)}\Gamma^k_{ij}{}^{(3)}\Gamma^l_{kl}-\gamma^{ij}{}^{(3)}\Gamma^l_{ik}{}^{(3)}\Gamma^k_{jl}, \tag{2.93}$$

[9]Its relation to the four-dimensional Ricci scalar is detailed in equation (3.24).

with ${}^{(3)}\Gamma$ as the connection corresponding to the induced metric, given like in (2.18)

$$ {}^{(3)}\Gamma^i_{jk} = \frac{1}{2}\gamma^{il}\left(\partial_j\gamma_{kl} + \partial_k\gamma_{jl} - \partial_l\gamma_{jk}\right), \tag{2.94} $$

where γ^{il} is the inverse of (2.92). Fortunately for us, the spatial derivatives of the perturbation variables in (2.94) are all of linear order which implies that we can write $\gamma^{il} = a^{-2}\delta^{il}$ and omit the perturbative parts. The scale factors cancel one another and one obtains an expression purely in terms of the perturbation variables. Concerning (2.93), only the components

$$ {}^{(3)}R \approx \gamma^{ij}\partial_k {}^{(3)}\Gamma^k_{ij} - \gamma^{ij}\partial_j {}^{(3)}\Gamma^k_{ik}, \tag{2.95} $$

survive, which after some lines of calculations produce

$$ a^{2\,(3)}R = -2\nabla^2\left[A - \frac{1}{2}\nabla^2 B\right], \tag{2.96} $$

where the content in the brackets is the *comoving curvature perturbation.* As the name hints at, it has here been calculated in the *comoving gauge* whose coordinates follow the the flow of the matter fluid. Hence, in order to obtain a gauge-invariant expression, we need to supply extra terms to (2.96), which yields the gauge-invariant comoving curvature perturbation $\mathcal{R}$. In Newton's gauge, [25], one can use (2.88) to write $\mathcal{R}$ as

$$ \mathcal{R} = -\Phi - \frac{4M_{\mathrm{PL}}^2 H(\dot{\Phi} + H\Phi)}{a^2(\bar{\rho} + \bar{p})}, \tag{2.97} $$

and show that it scales as

$$ \frac{\mathrm{d}\ln\mathcal{R}}{\mathrm{d}\ln a} \sim \left(\frac{k}{H}\right)^2, \tag{2.98} $$

for adiabatic perturbations, meaning that $\dot{R} \approx 0$ for superhorizon regimes. In the same limit [26], (2.97) takes the following form

$$ \mathcal{R} = -\frac{5 + 3\omega}{3(1+\omega)}\Phi, \tag{2.99} $$

using equation (2.82). As we discovered when we scrutinized equation (2.91), $\Phi = \text{const.}$ for superhorizon scales meaning that one can relate its magnitude during the radiation epoch, Φ_{RD}, to the matter era Φ_{MD} through, $\omega = 0$ for matter and $\omega = 1/3$ for radiation in (2.99). This yields $\Phi_{\mathrm{MD}} = 9/10\Phi_{\mathrm{RD}}$, which we will use in section 2.3.1.

2.3 Cosmological observables

In order to examine large-scale structures (LSS), we must first have a good grasp on how the background universe evolves during the eras. A nice introduction to these topics is presented in [22]. There are three primarily probes, which complement one another, for us to determine properties of the background, namely

(i) **SN Ia data:** Allows us to estimate the expansion history, i.e. $H(z)$, through measuring distances to supernovae explosions at different epochs.

(ii) **Baryonic acoustic oscillations (BAO) + cosmic microwave background (CMB):** Relics from the early universe which indicate angular scales and hence provide a statistical standard ruler.

Given a bright interstellar object such as a supernova, we can measure the distance d_L to it if we know its luminosity L (i.e. the total energy it emits) and notes the apparent luminosity l which we measure in our observatories on Earth and in its orbit,

$$l = \frac{L}{4\pi d_L^2}, \tag{2.100}$$

which is simply the fraction of the energy radiated through a sphere of radius d_L surrounding the object. By Liouville's theorem in statistical physics, the occupational number of photons or massive particles traveling through vacuum remains the same, provided no interaction occurs, i.e.

$$\frac{\mathrm{d}n}{\mathrm{d}t} = \frac{\partial n}{\partial t} + \frac{\partial n}{\partial \boldsymbol{x}}\frac{\partial \boldsymbol{x}}{\partial t} + \frac{\partial n}{\partial \boldsymbol{p}}\frac{\partial \boldsymbol{p}}{\partial t} = 0, \tag{2.101}$$

for particles following characterized by $n(\boldsymbol{x}, \boldsymbol{p}, t)$ traveling along a trajectory $[\boldsymbol{x}(t), \boldsymbol{p}(t)]$ in phase space. This means that the amount of photons or other particles which reaches us traveling through interstellar space is the same as the amount that was emitted towards us in the specific sold angle given no interaction with the foreground. Since the universe is mostly empty, this implies that l is roughly unaffected and if one has knowledge of the foreground one can compensate for the loss in apparent luminosity.

However, for distance $z > 0.1$, effects associated with the expansion of the universe cannot be neglected, forcing us to modify (2.100), taking account the energy shift owing to redshift, time delay between the arrival of individual photons and the changing radius of the sphere surrounding the interstellar object given by the metric. In all, these considerations yield,

$$d_L \to d_L(z) = a\,(t_0)\,r_1(1+z), \tag{2.102}$$

where a is the scale factor and r_1 the coordinate distance to Earth viewed from the interstellar object. This suffices to estimate the expansion rate, since the distance in a flat universe is calculated as

$$d_L(z) = (1+z)\int_0^z \frac{\mathrm{d}z'}{H(z')}, \tag{2.103}$$

where extra terms have to be added to incorporate curvature. In modified gravity, this function might be altered, which means that these observations can constraint such theories. Naturally, one can only hope to reach an estimation since no measurement is perfect and we cannot travel to the object and measure L. Still, there are certain types of supernovae in the sky, the SN Ia, that serve as approximate cosmic candles, meaning that they emit approximately the same amount of energy. Theoretically, one presumes that these cosmic explosions are the products of the interplay in a binary star system where a white dwarf accretes the gas of its fellow star, until its mass approaches the Chandrasekhar limit, i.e. the maximum mass supported by the pressure from electronic degeneracy. At this point it becomes unstable triggering the explosion. Owing to the similarity of the mass between the objects, close to the Chandrasekhar limit, the luminosity of the different supernovae will be similar to one another and hence they can serve as distance indicators [22].

As mentioned, the other two complementary indicators are remnants of the early universe. A flow-chart of different epochs is presented in figure 2.4 with the presumed inflation

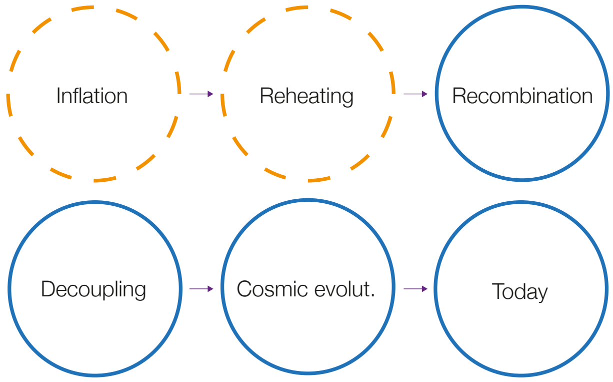


Figure 2.4: Overview of different time epochs in the evolution history of the universe. Reheating marks the end of the inflationary expansion and the recombination and decoupling events of photons and baryonic matter occurred $\sim$ 380 000 years after the Big Bang. The cosmic evolution phase is further characterized by an event known as *re-ionization*, roughly 200 million to 1 billion years after the Big Bang. This marked the advent of re-emerged ionized hydrogen, emitting photons, created in gravitational potential wells by dark matter structures. In the time frame between decoupling, perceived by the emission of the cosmic microwave background from the last-scattering surface, and re-ionization, there was only neutral hydrogen in the universe, which did not emit any light. This interlude is known as *the dark ages*.

related events in orange circles and the latter more established events in blue circles. The two phenomena originate from the recombination/decoupling epochs. With *recombination* we imply the time where the universe had cooled down sufficiently so that electrons and protons could form neutral hydrogen from the primordial plasma. A short time thereafter, photons decoupled from matter and this is referred to as *decoupling*. This event has been preserved as a snapshot in time, $\sim$ 380 000 years after the Big Bang corresponding to a redshift $z \sim 1100$, through the *cosmic background radiation (CMB)*, which is the light from the *last-scattering surface* between matter and photons. *Baryon acoustic oscillations (BAO)*, on the other hand, are the result of the interplay between pressure/gravitational forces in the transition from the primordial plasma [31]. Overdensities in this plasma gravitationally attracted more matter/radiation which in turn yielded an outward pressure force. As dark matter only interacts gravitationally, it remained in the potential well whereas the photons/baryonic matter were shifted away, producing a spherical sound wave in their wake outwards from the overdense region. After the decoupling of photons and baryons, the former constituent dissipates, which diminishes the pressure and leaves a shell of baryonic matter at a fixed distance, the sound horizon [32], from the center of the overdense region. Since the only remaining force is purely gravitational, matter will accrete at the center of the overdensity but also at the radius of the baryonic shell. As the universe expands and matures, galaxies will statistically gather at these points and by establishing the initial scale from the CMB, we can determine the expansion history. The present sound horizon has been measured to be $\sim$ 150 Mpc [32].

Concerning the large-scale structures themselves, their evolution is determined by the gravitational potential(s), Φ, which gives the density evolution δ. δ is measured by correlating

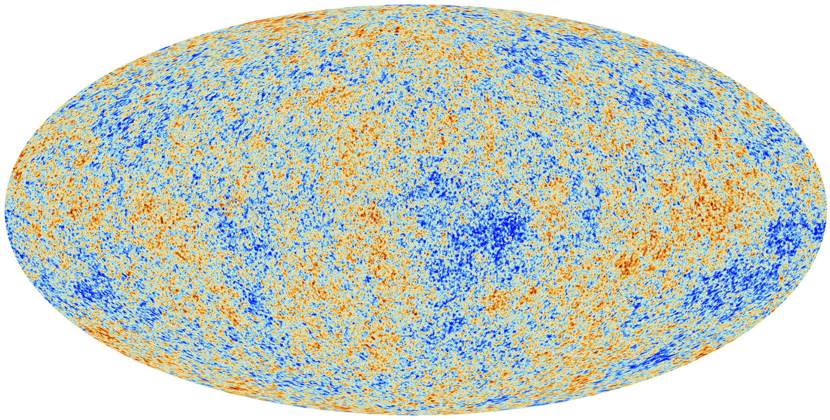


Figure 2.5: Primordial temperature fluctuations of the order 10^{-5} in the cosmic microwave background as measured by the Planck experiment in 2013 [35]. Courtesy of Planck/ESA.

the amount of galaxies in survey catalogues. From the observational perspective, this vital counting of galaxies is achieved for instance by the `Sloan Digital Sky Survey (SDSS)`. In addition, this allows us to use the BAO mechanism as a statistical ruler. Still, only visible luminous matter is measured, but on large scales it is possible to relate this proportionally to the dark matter density, $\delta_{\mathrm{DM}} = \mathrm{const} \cdot \delta_{\mathrm{ord.matter}}$. In the next section 2.3.1, we will detail another probe to investigate the evolution of the gravitational potentials in the LSS context.

One of the major difficulties in this endeavor is that many analyses of measured data and simulations incorporate predictions by general relativity and are hence not model-independent. It is hard to construct any observables at all without some prior assumptions and one seeks to narrow them down, which has for instance been carried out in [33]. A possible way forward is to use several observables, SN Ia + BAO/CMB + data from galaxy survey catalogues, to obtain a more model independent estimation of for instance the densities of the constituents of the universe [11]. In this report, we make a simplification by simply relying on data from Planck in section 4.2, meaning that there may be an inherent bias in the results. However, the results are qualitatively the same as in previous more extensive studies, [12], and parameters differ with a few decimal points. Future analyses should benefit from a larger selection of experimental data, where important contributions will probably come from upcoming experiments such as ESA's high definition large-scale structure survey `Euclid` mission [34].

2.3.1 Integrated Sachs-Wolfe effect

A specific probe to investigate the gravitational impact of large-scale structures is the *integrated Sachs-Wolfe effect* (ISW) [36] [37]. This phenomenon corresponds to a shift in the energy of CMB photons passing through the intergalactic medium, while encountering time-evolving potentials originating from large-scale structures. Firstly, we parametrize this problem in terms of the relative temperature fluctuations of the CMB, which define a brightness function on Earth, Θ, as

$$\Theta(\boldsymbol{x}, \boldsymbol{n}, \eta) \stackrel{\mathrm{def}}{=} \frac{\delta T_{\mathrm{CMB}}}{T_{\mathrm{CMB}}}(\boldsymbol{x}, -\boldsymbol{n}, \eta), \tag{2.104}$$

where $\boldsymbol{n}$ is the upward pointing normal vector on Earth's surface, η conformal time and $\boldsymbol{x}$ the physical position of the anisotropy in the sky. As we measure Θ over the whole sky, it is convenient to expand the function in spherical harmonics $Y_l^m(\boldsymbol{n})$

$$\Theta(\boldsymbol{x},\, \boldsymbol{n},\, \eta) = \sum_{l=1}^{\infty} \sum_{m=-l}^{l} a_{lm}(\boldsymbol{x},\, \eta) Y_l^m(\boldsymbol{n}), \tag{2.105}$$

Since Θ is real, the coefficients must satisfy $a^*_{lm} = a_{l\,-m}$. The average of these coefficients is zero, but they have a nonzero variance C_l, which is given by

$$\langle a^*_{lm} a_{l'm'} \rangle = \delta_{ll'} \delta_{mm'} C_l, \tag{2.106}$$

owing to the rotational invariance of the average of the product of two different δTs. Regrettably, it is not possible to measure the full C_l since we cannot average over position. What we actually observe is a variance averaged over m,

$$C_l^{\text{obs}} = \frac{1}{2l+1} \sum_{m=-l}^{l} a_{lm} a_{l\,-m} = \frac{1}{4\pi} \int \mathrm{d}^2\boldsymbol{n}\, \mathrm{d}^2\boldsymbol{n}' P_l \left(\boldsymbol{n} \cdot \boldsymbol{n}' \right) \Theta(\boldsymbol{n}) \Theta(\boldsymbol{n}'), \tag{2.107}$$

where the P_l are Legendre polynomials. The fractional difference between C_l and C_l^{obs} is known as the *cosmic variance* and the mean square of this difference is quantified

$$\left\langle \left(\frac{C_l - C_l^{\text{obs}}}{C_l} \right)^2 \right\rangle = 1 - 2 + \frac{1}{(2l+1)^2 C_l^2} \sum_m \sum_{m'} \langle a_{lm} a_{l\,-m} a_{lm'} a_{l\,-m'} \rangle = \frac{2}{2l+1}, \tag{2.108}$$

where the last equality is valid provided Θ follows a Gaussian distribution and where we have used relation (2.106) together with the observation that the coefficients are real. This equation provide us with a limit to which accuracy we can measure C_l. For small l, roughly $l \leqslant 5$, corresponding to the largest scales, the cosmic variance is large and hence we cannot determine C_l with precision. Next, we seek to find a direct expression for C_l directly in terms of Θ. We start by Fourier transforming Θ and express the brightness function evaluated at today, $\eta = 0$, in terms of multipole moment functions Θ_l [28],

$$\Theta(\boldsymbol{n}) = \sum_l i^l (2l+1) \int \mathrm{d}^3\boldsymbol{k} \Theta_l(\boldsymbol{k}) P_l \left(\frac{\boldsymbol{k} \cdot \boldsymbol{n}}{k} \right). \tag{2.109}$$

Then, we would like to decompose $\Theta_l(\boldsymbol{k})$ into a form where the contribution from the primordial inhomogeneities and the contribution from the structure evolution of the universe are visible i.e.,

$$\Theta_l(\boldsymbol{k}) = \Theta_l(k) \Phi_{(i)}(\boldsymbol{k}), \tag{2.110}$$

where $\Phi_{(i)}(\boldsymbol{k})$ is the primordial term and $\Theta_l(k)$ the evolution part, independent of direction due to isotropy and solely dependent on the magnitude k. This means that the variance can be recast into

$$\langle \Theta_l(\boldsymbol{k}) \Theta_l(\boldsymbol{k}') \rangle = \langle \Theta_l(k) \Theta_l^*(k') \rangle \left\langle \Phi_{(i)}(\boldsymbol{k}) \Phi^*_{(i)}(\boldsymbol{k}') \right\rangle. \tag{2.111}$$

Then, one proceeds by inverting equation (2.105) in Fourier space and inserting the result in (2.106), which together with the expansions above and the properties of the spherical harmonics allows us to write

$$C_l = 4\pi \int \frac{\mathrm{d}k}{k} \mathcal{P}_\Phi(k) \Theta_l^2(k), \tag{2.112}$$

with $\mathcal{P}_\Phi(k)$ independent of l and related to the variance of $\Phi_{(i)}(\boldsymbol{k})$ through

$$\left\langle \Phi_{(i)}(\boldsymbol{k})\Phi^*_{(i)}(\boldsymbol{k}')\right\rangle = \frac{1}{4\pi k^3}\mathcal{P}_\Phi(k)\delta(\boldsymbol{k}-\boldsymbol{k}'). \tag{2.113}$$

To summarize, we have related the field that transcribes the evolution of the universe, which is influenced by structure growth potentials, to the observable C_l. Now, our goal is to describe the prediction for $\Theta_l(k)$ in general relativity and in modified theories of gravity and quantify the difference. To achieve this, we need knowledge of C_l at the last-scattering surface and today. In general relativity, $T^4 \propto \rho_\gamma$ [25], which implies that, at the recombination era when the universe was dominated by radiation, the relative temperature perturbation Θ was related to the density perturbation through

$$\Theta_{\text{lss}} = \frac{1}{4}\delta_\gamma. \tag{2.114}$$

where Θ_{lss} is the perturbation at the last scattering surface. Today, the picture is slightly different, and to derive the evolution of the perturbations influenced by gravitational potentials we start from the geodesic equation

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} = 0, \tag{2.115}$$

which expressed for a photon's 4-momentum $P^\mu \overset{\text{def}}{=} \mathrm{d}x^\mu/\,\mathrm{d}\lambda$ is

$$\frac{\mathrm{d}P^\mu}{\mathrm{d}\lambda} + \Gamma^\mu_{\alpha\beta}P^\alpha P^\beta = 0. \tag{2.116}$$

Since the photons are massless, $P^2 = g_{\mu\nu}P^\mu P^\nu = 0$, and in the Newtonian gauge with a line element (2.86) this means,

$$-a^2(1+2\Psi)(P^0)^2 + a^2p^2 = 0 \Rightarrow P^0 = p(1-\Psi), \tag{2.117}$$

if one neglects terms above linear order and where $a^2p^2 = g_{ij}P^iP^j$. Next, we seek an expression for the spatial components P^i which can be written as $P^i = \alpha n^i$, where $\boldsymbol{n}$ is a unit vector in the direction of the 3-momentum and α a constant. To find α we write

$$a^2p^2 = a^2(1-2\Phi)\alpha^2 \Rightarrow \alpha = p(1+\Phi). \tag{2.118}$$

For the zero:th component in equation (2.116), this implies

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(p(1-\Psi)\right) = -\Gamma^0_{\alpha\beta}P^\alpha P^\beta \cdot \frac{1+\Psi}{p}, \tag{2.119}$$

where we have used $\mathrm{d}t/\,\mathrm{d}\lambda = P^0$. Expanding this and multiplying each side with $(1+\Psi)/p$ while keeping only linear terms yield

$$\frac{1}{p}\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\partial\Psi}{\partial t} - \Gamma^0_{\alpha\beta}P^\alpha P^\beta \cdot \frac{1+2\Psi}{p^2}, \tag{2.120}$$

with non-zero perturbed Christoffel symbols where the $\cdot$ signifies a partial time derivative,

$$\Gamma^0_{00} = H + \dot{\Psi},\ \Gamma^0_{0i} = \partial_i\Psi,\ \Gamma^0_{ii} = H - \dot{\Phi} - 2H(\Psi+\Phi). \tag{2.121}$$

After a brief, straightforward calculation, one deduces that (2.120) is reduced to

$$\frac{1}{p}\frac{\mathrm{d}p}{\mathrm{d}t} = -2H + \frac{\partial \Phi}{\partial t} - n^i \frac{\partial \Psi}{\partial x^i}, \tag{2.122}$$

at linear order. For a homogenous, isotropic universe

$$\frac{1}{p}\frac{\mathrm{d}p}{\mathrm{d}t} = -2H, \tag{2.123}$$

in conformal time. Since we can recast

$$n^i \frac{\partial \Psi}{\partial x^i} = \frac{\mathrm{d}\Psi}{\mathrm{d}t} - \frac{\partial \Psi}{\partial t}, \tag{2.124}$$

in terms of the total derivative given by $\mathrm{d}/\mathrm{d}t$, the perturbative part of (2.122) is

$$-\frac{\mathrm{d}\Psi}{\mathrm{d}t} + \dot{\Psi} + \dot{\Phi}. \tag{2.125}$$

With this to aid us, we are able to surmise an expression for the temperature perturbation today at t_0, namely

$$\Theta_0(\boldsymbol{n}) = \left(\Theta_{\mathrm{lss}} + \Psi + \boldsymbol{n}\cdot\boldsymbol{v}_b\right)|_{t_\star} + \int_{t_\star}^{t_0} \mathrm{d}t \left(\dot{\Psi} + \dot{\Phi}\right), \tag{2.126}$$

where the first parenthesis originates from the last scattering at $t_\star$ with t written in conformal time and where $\boldsymbol{n}\cdot\boldsymbol{v}_b$ is the Doppler shift for the photons with respect to an observer comoving with the baryons at the last scattering surface. Note that this expression has to be modified with an extra term to fit observations here on Earth due to the Doppler shift from the Earth's motion around the Sun. Here we have limited ourselves to an approximation of instantaneous recombination. In reality, recombination occurred over a time scale $\Delta z \sim 10$. The total derivative, $\mathrm{d}\Psi/\mathrm{d}t$, has been integrated from recombination till today, but the present term is only gauge-dependent, associated with how we view the background cosmology, and hence not an observable. Photons from the last scattering surface are redshifted due to the climb out of the potential well present at the time and this effect is incorporated in the Ψ-term, which together with Θ forms the *Sachs-Wolfe term*.

The last integral is the contribution from the *integrated Sachs-Wolfe effect (ISW)*, Θ_{ISW}. Owing to its relation to gravitational potentials, the effect is predominantly observed on large scales [37]. For cosmological models with $\Omega_m = 1$, the gravitational potentials are constant through the process of linear structure formation. However, in a universe with dark energy, the potentials are supposed to decay due to the accelerated expansion. As a consequence, CMB photons which travel through overdense regions, such as superclusters of galaxies, will be subjected to a positive ISW effect and those traversing voids will experience a negative ISW effect. It has been proposed as an indicator for modified gravity [38]. In a model with pressureless dust in general relativity, there will be no anisotropic stress present and $\Phi = \Psi$, which means that we can simplify the integral in (2.126). When we explore bimetric gravity we will have two propagating modes which yield two potentials which might yield extra anisotropic stress, implying $\Phi \neq \Psi$ for pressureless dust, and hence we can theoretically predict the deviation from general relativity and then compare this with experimental data. In general relativity in Fourier space, with no anisotropic stress, Θ_{ISW} can be expressed as

$$\Theta_{\mathrm{ISW}}(\boldsymbol{n}) = 2\int_{t_\star}^{t_0} \mathrm{d}t \int \mathrm{d}^3\boldsymbol{k}\, \dot{\Phi}(t,\, \boldsymbol{k}) e^{i\boldsymbol{k}\cdot\boldsymbol{n}(t_0 - t)}. \tag{2.127}$$

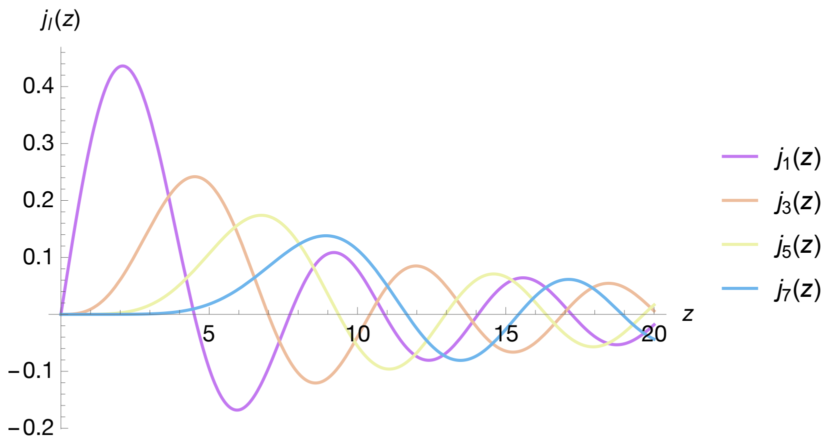


Figure 2.6: Illustration of some of the first spherical Bessel functions for small l.

For the large, superhorizon scales which we consider, it is possible to decompose $\Phi(t, \boldsymbol{k}) = b(t) \cdot \Phi_{\mathrm{MD}}(\boldsymbol{k})$, where $b(t)$ is a decreasing function and $\Phi_{\mathrm{MD}}(\boldsymbol{k})$ the value of Φ during the matter-dominated era, Φ_{MD}, when it was constant. From the expression of the comoving curvature perturbation at superhorizon scales, (2.99), we can relate the value of Φ during the radiation dominated era $\Phi_{(i)} = \Phi_{\mathrm{RD}}$ as $\Phi_{\mathrm{MD}} = 9/10 \cdot \Phi_{\mathrm{RD}}$ and consequentially

$$\Theta_{\mathrm{ISW}}(\boldsymbol{n}) = \frac{9}{5} \int_{t_\star}^{t_0} \mathrm{d}t \frac{\mathrm{d}b(t)}{\mathrm{d}t} \int \mathrm{d}^3\boldsymbol{k}\, \dot{\Phi}_{\mathrm{RD}}(\boldsymbol{k}) e^{i\boldsymbol{k}\cdot\boldsymbol{n}(t_0-t)}, \tag{2.128}$$

where one can use a Legendre expansion of the exponential [22],

$$e^{i\boldsymbol{k}\cdot\boldsymbol{n}(t_0-t)} = \sum_l (2l+1) i^l P_l\left(\frac{\boldsymbol{k}\cdot\boldsymbol{n}}{k}\right) j_l(k(t-t_0)), \tag{2.129}$$

with the functions j_l as spherical Bessel functions, related to the ordinary Bessel functions J_l through $j_l(z) \overset{\mathrm{def}}{=} (\pi/2z)^{1/2} J_{l+1/2}(z)$, to produce

$$\Theta_{\mathrm{ISW}}(\boldsymbol{n}) = \frac{9}{5} \int_{t_\star}^{t_0} \mathrm{d}t \frac{\mathrm{d}b(t)}{\mathrm{d}t} \int \mathrm{d}^3\boldsymbol{k}\, \dot{\Phi}_{\mathrm{RD}}(\boldsymbol{k}) \sum_l (2l+1) i^l P_l\left(\frac{\boldsymbol{k}\cdot\boldsymbol{n}}{k}\right) j_l(k(t-t_0)). \tag{2.130}$$

Making a comparison with the expansion of Θ in terms multipole moment functions in (2.109), this allows us to conclude that the multipole moment functions of Θ_{ISW} are given by

$$\Theta_{\mathrm{ISW},\,l}(\boldsymbol{k}) = \frac{9}{5} \int_{t_\star}^{t_0} \mathrm{d}t \frac{\mathrm{d}b(t)}{\mathrm{d}t} j_l(k(t-t_0)), \tag{2.131}$$

which enables us to calculate the observable C_l in (2.112) for the ISW effect,

$$C_{\mathrm{ISW},\,l} = 4\pi \cdot \left(\frac{9}{5}\right)^2 \int \frac{\mathrm{d}k}{k} \mathcal{P}_\Phi(k) \int_{t_\star}^{t_0} \int_{t_\star}^{t_0} \mathrm{d}t\, \mathrm{d}t' \frac{\mathrm{d}b(t)}{\mathrm{d}t} j_l(k(t-t_0)) \frac{\mathrm{d}b(t')}{\mathrm{d}t'} j_l(k(t'-t_0)). \tag{2.132}$$

By plotting some of the first few spherical Bessel functions in figure 2.6, we see that these terms in the integrals will be most significant for small l, i.e. at large scales, and hence the observable effect $C_{\mathrm{ISW},\,l}$ will be most evident at large scales. However, as we discovered in (2.108), we cannot go to arbitrarily small l. The ISW effect approximately starts to manifest itself for $l \leqslant 10$ [22], leaving us with a narrow window $5 < l \leqslant 10$ for experimental investigation. Still, we are primarily interested in it as a superficial probe for modified gravity and hence we are searching for abnormal signatures, deviating a lot from the predictions made by general relativity.

Chapter 3 Bimetric theory

3.1 Foundations

In this chapter we provide a brief review of the theoretical foundations and motivation for bimetric gravity. Primarily, it is based upon [10] and the recent reviews [8] [9].

Among the four forces, gravity is the one which is least well measured, with an uncertainty in the value of G of order 10^{-4} [39] and its presumable quantization is still a remote dream. In a modern description, general relativity can be viewed as an *effective field theory* of gravity, i.e. an approximation of the underlying unknown physical theory, valid up to the energy scales comparable with the Planck mass.

If we consider spacetime symmetries we identify particles as representations of the underlying group characterized by their spin and mass values. Over the years a number of particles have been found and consistent theories have been developed for several, theorized additional ones[1].

Spin	Particle
0	Higgs boson (dilaton, inflaton, etc.)
$\frac{1}{2}$	fermions
1	gauge bosons
$\frac{3}{2}$	Rarita-Schwinger fields
2	graviton

As for spin-2 fields, general relativity, consistent in the classical regime, proposes a massless graviton. Still, as far as we know today, there is no explicit theoretical restriction that by default rules out other possibilities. Granted, such theories must of course be mathematically consistent and physically sound according to our current observational capabilities. Concerning physical consistency one would like to avoid terms which could introduce negative kinetic energy, as such terms violate unitarity in a quantum description. At the classical level these bring instabilities. Such terms are known as *ghosts*. These are distinguished by being kinetic terms with a wrong sign in front of them in the Lagrangian of the theory. More precisely, they appear as terms with the opposite sign in front of them coupled to terms with the correct sign in front of them. On the other hand, one might also encounter mass terms with the wrong sign and these are known as *tachyonic terms*. We shall primary focus on how to build a ghost-free modified gravity theory in this work.

According to experimental surveys, the expansion of the universe has now entered an accelerated phase [1] [2]. The origin of this acceleration is usually attributed to an elusive *dark*

[1] The Rarita-Schwinger fields are theoretical fields which arise in supersymmetric applications and are governed by the Rarita-Schwinger equation, which is similar to the Dirac equation for spin 1/2 particles. For more information please refer to [40].

energy. Mathematically, this can be accommodated within Einstein's equations of general relativity as a non-zero *cosmological constant* Λ. Yet, the observationally deduced value which would fit the current expansion history of the order $10^{-47}\,(\mathrm{GeV})^4$, strongly disagrees with the number derived from calculations of the expectation value of the vacuum in quantum field theory, which could be up to 60-120 orders of magnitude larger. This apparent discrepancy is known as *the cosmological constant problem*, which is one of the lingering major, headache-causing issues in modern theoretical physics. A nice review on the subject can be found in [41]. Since the density which propels the accelerated expansion is relatively small, i.e. representing infrared energy levels, a good starting point would be to try to modify Einstein's theory in the infrared, which could be achieved by adding a small mass.

Before we proceed, it is a good idea to define what we mean by *linear* and *nonlinear*. Given a certain background, curved or flat, a *linear* interaction only include first order terms in perturbation theory whereas a *nonlinear* one also covers higher order corrections.

In the Standard Model of particle physics, massless fermion fields acquire mass through Yukawa couplings with the Higgs field. Through spontaneous symmetry breaking the fields acquire masses proportional to the vacuum expectation value of the Higgs field. When renormalizing the theory, these masses are protected by the chiral symmetry of the unbroken theory with massless fermions. Briefly, the symmetry prohibits certain problematic interaction terms, which would cause large corrections. This ensures that the loop corrections are of the same order as the fermion masses. These masses are considered to be technically natural owing to their stability under quantum corrections following *t' Hooft naturalness* argument [42] [43]. In general, the procedure can be extended to any theory with a symmetry, be it local or global, which (re-)appears as the masses involved go to zero. Hence, from a quantum field theory point of view, a theory of gravity with a massive field with a small mass where diffeomorphism invariance appear as the mass go to zero would be technically natural, [44] [45]. The cosmological constant, on the other hand, is not accompanied by a symmetry in the zero-limit which would prevent quantum corrections running amok. This means that a graviton with a small mass can be viewed as preferable compared to a cosmological constant of similar size.

To address the discrepancy between the observationally small Λ and its large quantum counterpart, one could imagine that the vacuum expectation value could be screened by a mass term involving a small graviton mass in Einstein's equations. If so, the vacuum energy's implication on cosmological expansion could be tiny in size. The force mediated by a massive graviton would have a Yukawa profile, i.e.

$$F \sim \frac{1}{r}\mathrm{e}^{-mr} \tag{3.1}$$

where m is the mass of the graviton and r the distance. This term emerges from the ordinary case with a massless graviton at length scales $r \sim m^{-1}$. In Einstein's equations one manifestation of this idea is to promote Newton's constant G to a differential operator $G\left(L^2\Box\right)$, which leads to the following relation

$$G^{-1}\left(L^2\Box\right)G_{\mu\nu} = 8\pi T_{\mu\nu}. \tag{3.2}$$

Being a function of the covariant d'Alembertian, Newton's constant acts as a high-pass filter where the scale of the filter is determined by L. A large excitation on the matter side with a long wavelength $\gg L$, i.e. the cosmological constant, is then *degravitated* by passing through the filter, [46]. It can be argued that this not realizable with an ordinary massless graviton (the propagating degrees of freedom does not add up) but with a massive spin-2 particle with mass $m \sim L^{-1}$ or a resonance of an infinite set of such particles with masses $m \sim L^{-1}$ [47].

Therefore by setting m approximately equal to the Hubble constant one might explain the smallness of the observed value. Unfortunately, so far attempts to find a working straightforward screening mechanism have been futile.

The tread which we will follow throughout this chapter is counting the number of degrees of freedom which the propagating fields carry. If they do not add up to the expected value for a (massive/massless) particle of a given spin, they indicate the presence of ghosts which carry the redundant degrees of freedom. Hence, it is essential to find theoretical constraints which eliminate these problematic degrees of freedom. In the spin-2 case, we will see how it is performed in general relativity and in the consistency argument for bimetric theory.

3.2 Origins

A theory with massive spin-2 particles was already proposed by Fierz and Pauli as early as 1939 [48], following the same spirit as the Proca Lagrangian, the massive version of Maxwell's electromagnetism. They managed to write down a unique ghost-free mass term for a linearized fluctuation of a spin-2 field. In d dimensions in flat space, the Fierz-Pauli action for a single free spin-2 field with mass m takes the following form

$$S = \int \mathrm{d}^d x \left[-\frac{1}{2}\partial_\rho \left(h_{\mu\nu}\partial^\rho h^{\mu\nu} - h\partial^\rho h \right) + \partial_\mu \left(h_{\nu\rho}\partial^\nu h^{\mu\rho} - h^{\mu\nu}\partial_\nu h \right) - \frac{m^2}{2}\left(h_{\mu\nu}h^{\mu\nu} - h^2 \right) \right], \tag{3.3}$$

where the field is propagated by a symmetric tensor $h_{\mu\nu}$ [8]. Here, the trace of $h_{\mu\nu}$ is denoted as h. The kinetic term is simply a linearized version of the Einstein-Hilbert term, and contains all contractions of two powers of $h_{\mu\nu}$ with two derivatives. The coefficients of the involved parts are fixed so that the action with $m = 0$ is invariant under

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_{(\mu}\xi_{\nu)}, \tag{3.4}$$

for a spacetime dependent parameter $\xi(x)$. This transformation is a linearized version of general relativity's diffeomorphism invariance. To create a mass term one requires an object with no free indices and with only one tensor $h_{\mu\nu}$ available, there are two options $h_{\mu\nu}h^{\mu\nu}$ and h^2. Hence, the prototype mass term in the Lagrangian emerges as

$$\mathcal{L}_{\text{mass}} \propto -m^2 \left(h_{\mu\nu}h^{\mu\nu} - \alpha h^2 \right), \tag{3.5}$$

where α is a dimensionless parameter to be determined [9]. In order to find $\alpha = -1$ as one observes in (3.3), it is convenient to introduce a concept known as *Stückelberg fields*. These are ubiquitous tools in analyzing propagating degrees of freedom in massive gravity. Their purpose is to restore the underlying symmetry of a theory, in this particular case linearized diffeomorphism invariance, at the expense of introducing supplemental fields. Let these fields be called υ_μ and let them transform as

$$\upsilon_\mu \to \upsilon_\mu - \frac{1}{2}\xi_\mu, \tag{3.6}$$

simultaneously as $h_{\mu\nu}$ transforms as in (3.4). Then, the mass term

$$-m^2\left(\left(h_{\mu\nu}+2\partial_{(\mu}\upsilon_{\nu)}\right)^2-\alpha\left(h^2+2\partial_\rho\upsilon^\rho\right)^2\right), \tag{3.7}$$

is left invariant and hence the action is invariant under linear diffeomorphisms. In this new mass term one obtains a purely kinetic term for the Stückelberg fields, namely

$$\mathcal{L}_m^{\text{S, kin}}\propto m^2\left((\partial_\mu\upsilon_\nu)^2-\alpha\,(\partial_\rho\upsilon^\rho)^2\right). \tag{3.8}$$

Now, in order to investigate which choice(s) of α result in a ghost-free theory we split the fields υ^μ into a transversal and a longitudinal part, where the transversal part obeys $\partial^\mu\upsilon_\mu^\perp=0$,

$$\upsilon_\mu=\upsilon_\mu^\perp+\partial_\mu\upsilon^{\text{L}}. \tag{3.9}$$

After performing the split, we examine the kinetic term for the purely longitudinal component,

$$(1-\alpha)\,\partial_\mu\partial_\nu\upsilon^{\text{L}}\partial^\mu\partial^\nu\upsilon^{\text{L}}=(1-\alpha)\left(\Box\upsilon^{\text{L}}\right)^2, \tag{3.10}$$

where $\Box$ is the d'Alembertian operator, $\eta^{\mu\nu}\partial_\mu\partial_\nu$, in flat Minkowski space. If $\alpha=1$, no higher order derivatives in space or time appear of the longitudinal component. This is promising, since higher order derivatives would indicate the presence of a ghost as inferred from Ostrogradsky's theorem [49]. In fact, two degrees of freedom are contained in the longitudinal part and each comes with a kinetic term with the opposite sign with respect to the other one. The problem can be illustrated by rewriting the d'Alembertian as a sum of two propagators

$$\frac{1}{\Box^2}=\lim_{m\to 0}\frac{1}{2m^2}\left(\frac{1}{1-m^2}-\frac{1}{1+m^2}\right), \tag{3.11}$$

where the last term clearly looks odd and will couple in the wrong way to external sources. As the mass can be put arbitrarily small, this means that the issue will be present for all $\alpha\neq 1$, which motivates its value. Hence, the $\alpha=1$ brings about the unique consistent mass term and we reach the sought form in (3.3) by choosing the *unitary* gauge for the Stückelberg fields, i.e. the specific gauge where these take on zero values. Later, we will see how this mass term can be generalized to the nonlinear regime of contemporary massive gravity and bimetric gravity.

Arriving at this conclusion, we would like to count the degrees of freedom of the theory, which will become important when we examine the consistency of massive and bimetric gravity since they must reproduce this result. Here, we will work in four dimensions but the result can be extended to higher dimensions in a straightforward manner. For the linear, free Fierz-Pauli action this counting can be done in several ways. Firstly, one can perform a Legendre transformation and do a Hamiltonian analysis or take certain linear combinations of the terms and derivatives of them to make the different modes visible [8]. Another approach, [9], is to work with Stückelberg fields, where one performs an alternative split of the fields to the one in (3.9) with,

$$\upsilon_\mu^\perp=\frac{1}{m}A_\mu,\ \ \partial_\mu\upsilon^{\text{L}}=\frac{1}{m^2}\partial_\mu\pi, \tag{3.12}$$

where the mass factors have been included for convenience and where A_μ will prove to represent the helicity-1 mode and π the helicity-0 mode respectively later on. If we rewrite the Fierz-Pauli action in terms of $h_{\mu\nu}$ and the two Stückelberg fields A_μ and π, replacing $h_{\mu\nu}$ and h in the same manner as in (3.7), it takes the following form [9],

$$\begin{aligned}\mathcal{L} = -\frac{1}{4}h^{\mu\nu}\mathcal{E}^{\rho\sigma}_{\mu\nu}h_{\rho\sigma} - \frac{1}{2}h^{\mu\nu}\left(\partial_\mu\partial_\nu\pi - \mathrm{Tr}(\partial_\mu\partial_\nu\pi)\eta_{\mu\nu}\right) - \frac{1}{8}F_{\mu\nu}F^{\mu\nu} - \\ \frac{1}{8}m^2\left(h_{\mu\nu}h^{\mu\nu} - h^2\right) - \frac{1}{2}m\left(h^{\mu\nu} - h\eta^{\mu\nu}\right)\partial_{(\mu}A_{\nu)},\end{aligned} \tag{3.13}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and where $\mathcal{E}$ is a kinetic operator[2] which when acted on $h_{\mu\nu}$ gives the kinetic term in (3.3) where an additional $h^{\mu\nu}$ has been used to contract its two free indices and the term has been divided by two. At this point we observe that we have obtained one purely kinetic term for $h_{\mu\nu}$ and one for A_μ, but the one for π is still contracted with $h_{\mu\nu}$. To remedy this, one can diagonalize the mixing through the shift $h_{\mu\nu} \to \tilde{h}_{\mu\nu} + \pi\eta_{\mu\nu}$, which leads to the following kinetic terms

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4}\tilde{h}^{\mu\nu}\mathcal{E}^{\rho\sigma}_{\mu\nu}\tilde{h}_{\rho\sigma} - \frac{3}{4}\left(\partial\pi\right)^2 - \frac{1}{8}F_{\mu\nu}F^{\mu\nu}. \tag{3.14}$$

Here, we can recognize the kinetic terms for a helicity-2 mode $\tilde{h}_{\mu\nu}$ which carries two degrees of freedom as in general relativity, a helicity-1 mode A_μ which propagates two degrees of freedom and finally a helicity-0 mode which brings an additional one. This means that Fierz-Pauli massive gravity carries five degrees of freedom in four dimensions. An alternative route is to note the propagating degrees of freedom through analysis of the representations of the Poincaré group for massive particles in a specific dimension. A nice introduction to the Poincaré group can be found in [50].

A few decades later, in the 1970s, van Dam and Veltman, [51], and Zacharov, [52], independently discovered that the Fierz-Pauli action with interactions coupled to sources did not reduce to the massless theory in the zero-mass limit, due to the helicity-0 mode which we identified previously which is not present in the massless theory. This is known as the *vDVZ discontinuity*. The helicity-1 mode does not pose such difficulties as it does not couple to external sources and hence it does not matter that it is nonexistent for the massless theory. A loop-hole to reconcile these results was found a few years afterwards in 1972 though the so-called *Vainshtein mechanism*, [53]. A nice introduction to the topic is provided in [54]. For a theory of massive gravity, at low energies, nonlinear effects start to dominate when the mass is small as the theory become strongly coupled. More explicitly, the linear approximation ceased to be valid inside a certain radius of massive sources, called the *Vainshtein radius*. The vDVZ discontinuity is derived using a linear approximation of the theory and special non-linear interactions could thus suppress the additional degrees of freedom, i.e. the helicity-0 mode, which are propagated by massive gravity in the zero-mass limit. To find the detailed form of such interactions, a nonlinear completion of Fierz-Pauli massive gravity had to be established. However, this endeavor must be pursued with great care, since theoretical inconsistencies are generally introduced. In the same year as the Vaihnstein mechanism was formulated, Bouleware and Deser nailed down such a pathology, notoriously known as the *Bouleware-Deser ghost*, [55] [56] which proved to be excruciatingly hard to avoid and thus extinguished the interest in these theories until the beginning of the 21st century. Basically,

[2] It is known as the Lichnerowicz operator [9].

the Bouleware-Deser ghost is yet a consequence of Ostrogradsky's theorem as the extra non-linear terms in a generalized Fierz-Pauli action generically yield cubic and quartic differential operators in the equations of motion which give rise to an additional degree of freedom, a ghost. In four dimensions, the theory suddenly propagates six degrees of freedom instead of the expected five where the last one is the ghost.

In 2005, Creminelli et. al. reformulated the analysis from an effective field theory point-of-view [57], based on an earlier work by Arkani-Hamed et al. [58], which seemingly verified the inevitability of ghosts in massive gravity. However, a sign mistake in their paper changed the conclusion completely which was discovered in 2010 by de Rham, Gabadadze and Tolley and allowed them to write down the first consistent nonlinear theory of massive gravity the same year [59] [60]. This theory, known as *massive gravity*, requires two metrics where one is dynamical and one is a fixed reference metric. Still, their consistency proof was only valid in a certain decoupling limit. In 2011, Hassan and Rosen reformulated the theory in such a way [61] that the full consistency could be confirmed [3] [62]. In addition, this setup allowed for a simple generalization to promote the reference metric to a dynamical one [5] , leading to a theory of two dynamical metrics on equal footing, *bimetric massive gravity*, which is the central topic of this thesis. In the next section, we will explore the basics of the proof formulation and gain insight about the dynamical structure of spacetime theories.

3.3 ADM formalism

An introductory course in general relativity typically involves deriving Einstein's equations by varying the Einstein-Hilbert action. Through this one studies the Lagrangian, or more specifically the Lagrangian density, of the theory. However, it is not apparent which terms encode the true dynamical degrees of freedom and which that merely provide constraints. To elaborate, this has to do with the diffeomorphism invariance of the theory and this freedom to choose coordinates introduces gauge modes which are not real degrees of freedom. In order to illuminate this, one studies the Hamiltonian of the theory. This requires an explicit split between space and time and hence reformulation of the covariant theory. In this treatment, we are principally inspired by chapter 12 in [24] and appendix E of [63]. The application to massive gravity follows the approach in [10] closely.

We start by introducing a scalar field $t = t(x^a)$, where x^a are coordinates on the spacetime manifold, which is defined in a such way that spacetime is foliated into non-intersecting spacelike hypersurfaces $\Sigma(t)$ of constant t. An illustration is provided in figure 3.1. The normal n^a of these hypersurfaces is proportional to $\partial_a t$. One may then introduce a spatial coordinate system x^α on these surfaces. To connect points on different surfaces a congruence of curves parametrized by t is installed so that events on different hypersurfaces have the same x^α provided they lie on the same curve. Combined with the t-coordinate, one obtains a four-dimensional coordinate system $x^b = (t, x^\alpha)$ with tangent vector to the curves $t^a = \partial x^a/\partial t$ satisfying $t^a \partial_a t = 1$. To continue, the construct $e^a_\alpha = (\partial x^a/\partial y^\alpha)$ projects onto the surface $\Sigma(t)$. For a unit normal to the hypersurfaces one defines $n_a = -N\partial_a t$ where N is a scalar function to ensure proper normalization known as the *lapse*. Naturally, $n_a e^a_\alpha = 0$. With these tools it is possible to decompose the curves' tangent vector along the normal and the tangent plane to the hypersurfaces, i.e. $t^a = Nn^a + N^\alpha e^a_\alpha$. An infinitesimal coordinate transformation $x^a = x^a\,(t,\, x^\alpha)$ could then be formulated as

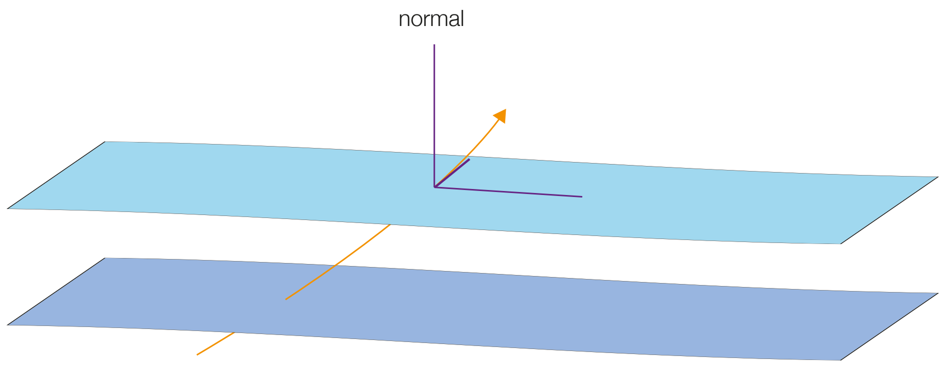


Figure 3.1: Spacetime foliated in terms of hypersurfaces $\Sigma(t)$ intersected by curves parametrized by t. The surface unit normal is given by $n_a = -N\partial_a t$ and a tangent vector to the curve can be decomposed as $t^a = Nn^a + N^\alpha e^a_\alpha$, where $N^\alpha e^a_\alpha$ lies in the tangent plane to the surface, which is three-dimensional.

$$\mathrm{d}x^a = t^a\,\mathrm{d}t + e^a_\alpha\,\mathrm{d}y^\alpha = (N\,\mathrm{d}t)n^a + (N^\alpha\,\mathrm{d}t + \mathrm{d}y^\alpha)\,e^a_\alpha, \tag{3.15}$$

with a line element

$$\mathrm{d}s^2 = \mathrm{d}x^a\,\mathrm{d}x_a = g_{ab}\,\mathrm{d}x^a\,\mathrm{d}x^b = -N^2\,\mathrm{d}t^2 + h_{\alpha\beta}\left(N^\alpha\,\mathrm{d}t + \mathrm{d}x^\alpha\right)\left(N^\beta\,\mathrm{d}t + \mathrm{d}x^\beta\right), \tag{3.16}$$

where the induced metric on the hypersurface is

$$h_{\alpha\beta} = g_{ab}e^a_\alpha e^b_\beta = g_{\alpha\beta}. \tag{3.17}$$

This implies that it is possible to write the time-time, time-space and space-space components in terms of the lapse, shift and induced metric variables as $g_{00} = N^\rho N_\rho - N^2$, $g_{0\alpha} = N_\alpha$ and $g_{\alpha\beta} = h_{\alpha\beta}$, respectively. In matrix form this is equivalent to

$$g_{ab} = \begin{pmatrix} -N^2 + N_\alpha h^{\alpha\beta}N_\beta & N_\beta \\ N_\alpha & h_{\alpha\beta} \end{pmatrix}, \tag{3.18}$$

with $h^{\alpha\beta}$ the inverse of the induced metric. From this we readily deduce the inverse

$$g^{ab} = \frac{1}{N^2}\begin{pmatrix} -1 & N^\beta \\ N^\alpha & N^2h^{\alpha\beta} - N^\alpha N^\beta \end{pmatrix}. \tag{3.19}$$

One recognizes that $\sqrt{-g} = N\sqrt{h}$ where $h = \det h_{\alpha\beta}$. After having sliced up spacetime in this fashion, one would like to define variables which carry information about the surfaces

$\Sigma(t)$. We already have the induced metric, which tells us about the intrinsic properties of the surfaces (and lets us define a covariant derivative on them). Yet, one could also define an *extrinsic curvature*, which contains information about surrounding higher-dimensional spacetime's structure. This is given by

$$K_{cd} = -h^a{}_c h^b{}_d \nabla_a n_b = -h^a{}_c \nabla_a n_d. \tag{3.20}$$

Here $h^a{}_c = \delta^a{}_c + n^a n_c$, the natural projection tensor onto $\Sigma(t)$. The last equality follows from differentiating the normalization condition $n^a n_a = -1$, i.e. $n^a \nabla_b n_a = 0$. This convey information about how the normal is changing in the surrounding space, projected onto the surface $\Sigma(t)$. Equivalently, one can give the extrinsic curvature in terms of the shift and the lapse as

$$-K_{cd} = h^a{}_c \nabla_a n_d = \nabla_c n_d + n^a n_c \nabla_a n_d = \nabla_c n_d + n_c \left(n^a \nabla_a n_d\right) = \nabla_c n_d + n_c a_d, \tag{3.21}$$

where we denote the acceleration analogous to the normal vector as a_d. In practice, this means that the covariant derivative of the normal vector can be decomposed into a part tangent to the hypersurface encoded within the extrinsic curvature and a part normal to it given by the acceleration part. By using this and the decomposition of the induced metric, one observes that the covariant spatial components can be calculated from $K_{\alpha\beta} = -\nabla_\beta n_\alpha = -N\Gamma^0_{\alpha\beta}$. Evaluating this, one concludes

$$K_{ij} = \frac{1}{2N}\left(\nabla_i N_j + \nabla_j N_i - \partial_0 h_{ij}\right), \tag{3.22}$$

where ∇_i is the three-dimensional covariant derivative which acts on vectors tangent to the hypersurface. Should we express this in a coordinate system where the components of the shift take on a zero value, the extrinsic curvature is simply given by

$$K_{\mu\nu} = -\frac{1}{2N}\partial_t h_{\mu\nu}. \tag{3.23}$$

From this it follows that the extrinsic curvature is manifestly spatial. Its non-zero components convey information about the time-derivative of $h_{\alpha\beta}$. Together these variables are known as *ADM variables* after Arnowitt-Deser-Misner [64] who first formulated it in 1962. It is possible to show, [24], that the Ricci scalar can be expressed as

$$R = {}^{(3)}R + K_{ab}K^{ab} + K^2 + \text{surface term in the action}, \tag{3.24}$$

using the Gauss-Codazzi equations, relating the Riemann tensor in the higher dimensional space to that on the hypersurfaces. Together with the previous determinant relation, the Einstein-Hilbert action can be written as

$$S = \int \mathrm{d}^d x N\sqrt{h}\left({}^{(3)}R + K_{ab}K^{ab} + K^2\right), \tag{3.25}$$

without the surface term. We note that the lapse and the shift do not appear with any time derivatives, i.e. they do not encode the dynamics of the theory. This means that their canonical momenta vanish identically and that the variations of the action with respect to these variables, yielding zero, will act as constraint equations. By performing a subsequent Legendre transformation on the dynamical spatial variables, h_{ij}, using the extrinsic curvature in (3.22), one acquires its non-zero canonical momenta. The process is simplified if one notes that the variables only appear in terms with K_{ij},

$$\pi^{ij} = \frac{\delta L_{\mathrm{ADM}}}{\delta \dot{h}_{ij}} = \frac{\partial K_{ij}}{\partial \dot{h}_{ij}} \frac{\delta L_{\mathrm{ADM}}}{\delta K_{ij}} = -\sqrt{h}\left(K^{ij} - K h^{ij}\right). \tag{3.26}$$

With this definition it is possible to find an expression for the Hamiltonian

$$H_{\mathrm{ADM}} = \left(\int_{\Sigma} \pi^{ij} \dot{h}_{ij} \, \mathrm{d}^{d-1}x\right) - L_{\mathrm{ADM}} = \int_{\Sigma} \mathcal{H}_{\mathrm{ADM}} \, \mathrm{d}^{d-1}x. \tag{3.27}$$

Evaluating the expression as a Hamiltonian density using (3.25), (3.22) and (3.26), one discerns

$$\begin{aligned} \mathcal{H}_{\mathrm{ADM}} &= \sqrt{h}\left(K^{ij} - K h^{ij}\right)\left(2NK_{ij} - \nabla_j N_i - \nabla_i N_j\right) - N\sqrt{h}\left({}^{(3)}R + K_{ij}K^{ij} + K^2\right) \\ &= N\left(-{}^{(3)}R + K_{ij}K^{ij} - K^2\right)\sqrt{h} - 2\sqrt{h}\left(K^{ij} - Kh^{ij}\right)\nabla_j N_i \\ &= \sqrt{h}\left[N\left(-{}^{(3)}R + K_{ij}K^{ij} - K^2\right) + 2N_i \nabla_j\left(K^{ij} - Kh^{ij}\right)\right], \end{aligned} \tag{3.28}$$

where we have omitted a total derivative on the last line. This implies that the Hamiltonian density can be expressed as the lapse and the shift multiplied by functions of the canonical momenta and the spatial variables, i.e.

$$\mathcal{H}_{\mathrm{ADM}} = N\mathcal{R}^0(h, \pi) + N_i \mathcal{R}^i(h, \pi), \tag{3.29}$$

where one inverts (3.26) in order to recast the extrinsic curvature as a function of π and h, namely

$$\sqrt{h}K^{ij} = -\mathrm{const}\left(\pi^{ij} - \frac{1}{d-2}\pi h^{ij}\right). \tag{3.30}$$

With this result in mind, we now turn to some general results from analytical mechanics regarding constrained Hamiltonian systems. Please refer to [65] for an extensive review. As stated before, the lapse and shift do not appear with any dynamical terms and all their terms are linear. Hence, they are Lagrange multipliers. On-shell, i.e. when the equations of motion are satisfied on the constraint surface, $\mathcal{R}^0$ and $\mathcal{R}^i$ are identically zero. These provide d *primary constraints* in phase space, constraints valid for all times. Moreover, we can use gauge invariance to remove additional degrees of freedom. If we evaluate the Poisson brackets between the independent linear combinations of the primary constraints on the constraint surface, we will discover that they yield zero in $2d - d$ cases. These are known as *first-class constraints* and represent a gauge redundancy, in practice d opportunities to choose N and N^i, in the Hamiltonian formalism. Another way to see it is to return to the Hamiltonian expression, again noting that no time derivatives of the lapse and shift appear which means that their time evolution is unconstrained and that they are solely determined through a gauge choice. These are responsible for the d diffeomorphism symmetries of the action. Counting degrees of freedom, $d(d-1)/2$ comes from h_{ij} and its canonical momentum π^{ij} bring an additional $d(d-1)/2$ possibilities. From this we subtract the d primary constraints and the d first class constraints, leaving us with $d(d-3)$ remaining dynamical degrees of freedom in phase space. Physically, this corresponds to the helicity states, and their canonical momenta, of a massless spin-2 particle. In four dimensions, one concludes that gravitational waves have two polarizations.

We would like to perform a similar analysis in modified theories of gravity to check that they propagate the correct number of degrees of freedom, corresponding to the force-mediating particle of the theory.

3.4 Ghost-free massive gravity

If we would like to introduce a modified theory of gravity with a massive particle, we must specify a mass term in the action. For a rank-2 tensor g_{ab} this should be a scalar density, namely some nontrivial scalar function $V(g)$ (potential) multiplied by the scalar density $\sqrt{-g}$. This scalar function $V(g)$ should not involve any derivatives. Since the metric comes with two loose indices, we require another rank-2 tensor in order to contract them. The first hand option would be to use the inverse of g_{ab} to achieve it, yet this would only add a trivial cosmological constant term to the action. Regrettably, this implies that one cannot formulate a covariant nonlinear interaction term suited for a spin-2 field with a sole tensor field available. To amend this, one could supply an additional rank-2 tensor field, denoted f_{ab}, to construct a nonlinear interaction term together with g_{ab}.

Subjected to an ADM decomposition, the corresponding shift and lapse functions of such a theory would still come without time derivatives. Nevertheless, there is nothing a priori which guarantees that they appear linearly in the interaction term. If they do not, their equations of motion would involve N and N^i themselves meaning that N and N^i would be constrained rather than the metric. Additionally the d constraints arising from gauge redundancy are lost. Hence we are left with $d(d-1)/2$ propagating degrees of freedom plus those from their canonical momenta. However, for a massive spin-2 particle this represents only an extra two degrees of freedom, one propagating and one from its canonical momentum. This is the Bouleware-Deser ghost and it has been shown that a clever choice of the interaction potential accompanied by a secondary condition furnish two additional constraints which eliminate this pathological part.

In order to build a potential term which automatically bring a supplemental constraint, we first split the new tensor field f_{ab} into ADM variables,

$$f_{ab} = \begin{pmatrix} -\tilde{N} + \tilde{N}_l \tilde{h}^{lk} \tilde{N}_k & \tilde{N}_j \\ \tilde{N}_i & \tilde{h}_{ij} \end{pmatrix}. \tag{3.31}$$

Next, we recast $\sqrt{-g} = N\sqrt{h}$ as previously, and hence obtains an interaction term of the form

$$\sqrt{-g} \cdot V\left(g^{-1} f\right) = N\sqrt{h} \cdot V\left(h_{ij},\, N,\, N^i;\, \tilde{h}_{ij},\, \tilde{N},\, \tilde{N}^i\right). \tag{3.32}$$

To enable this term to give the sought constraint, it suffices to require that the Lagrangian density should be linear in N and that N should be absent in the equations of motion, as it will render the Lagrangian linear in N^i automatically [4]. More precisely, the constraint could arise due to a combination of the equations for N and N^i should such a combination be independent of these variables. Concerning the action, this means that one ought to permit a field-dependent redefinition of the shift $N^i \to n^i$ which would make the Lagrangian linear in the lapse. Before we observed that N^i appeared linearly in the kinetic terms in GR and it will also be the case here. As a consequence, the redefinition of the shift itself must be linear in N; otherwise this property would be lost. In addition, since the new shift is assumed to be present in the constraint it must be determined through its own equation of motion, i.e. it has to be independent of N. From this one deduces that the variation of the action $\mathcal{S}$ with respect to the redefined shift can be written as

$$\frac{\delta S}{\delta n^i} = \frac{\delta N^j}{\delta n^i}\frac{\delta S}{\delta N^j} = 0. \tag{3.33}$$

As the redefinition is linear in N, so must also be the case for the (nonzero) Jacobian of the redefinition, $\delta N^j/\delta n^i$, and this leaves $\delta S/\delta N^j = 0$. By examining (3.32), one immediately recognizes that the potential term already comes with a factor of N in front of it, so to acquire a Lagrangian linear in N it must resemble

$$V\left(g^{-1}f\right) = \frac{1}{N}V_1 + V_2, \tag{3.34}$$

where V_1 and V_2 are scalar functions of the remaining variables. Now, we proceed to find candidate expressions which could provide us with terms of this form. It is inferred from the ADM decomposition of g_{ab}, (3.18) (3.19), that our sole chance to obtain a term linear in $1/N$ is through the square root of the inverse of g_{ab}. This motivates a potential which is a function of a matrix $S = \sqrt{g^{-1}f}$, defined through

$$\sqrt{g^{-1}f}\cdot\sqrt{g^{-1}f} = g^{-1}f. \tag{3.35}$$

Yet, performing an ADM decomposition of such a square-root matrix proves to be a formidable challenge at first glance. Fortunately, we can use the redefinition of the shift to recast the expression into a simpler form to make it linear in $1/N$. The procedure is as following; start by investigating the matrix $g^{-1}f$ and redefine the shift as

$$N^i = c_1^i\left(h_{ij},\, n^i\right) + Nc_2^i\left(h_{ij},\, n^i\right), \tag{3.36}$$

where the functions c_1 and c_2 are independent of the lapse. Then, write

$$N^2 g^{-1} f = \mathbb{E}_0 + N\mathbb{E}_1 + N^2\mathbb{E}_2, \tag{3.37}$$

with matrices $\mathbb{E}_i$ independent of N. Here, replace the shift with its new components, (3.36), and compare coefficients using (3.19) and (3.31). To obtain a potential term linear in N in the Lagrangian, we require

$$N\sqrt{g^{-1}f} = \mathbb{A} + N\mathbb{B}. \tag{3.38}$$

In order for this to be consistent with (3.37), the matrices $\mathbb{A}$ and $\mathbb{B}$ must be given by $\mathbb{A}^2 = \mathbb{E}_0$, $\mathbb{B}^2 = \mathbb{E}_2$ and $\mathbb{A}\mathbb{B} + \mathbb{B}\mathbb{A} = \mathbb{E}_1$. The prescribed redefinition of the shift can be read out of the last relation. Still, the redefinition is not unique and one has some liberty to choose variables suitable for a specific calculation. One choice is presented in appendix C. The gist of the argument to obtain a palpable ghost-free potential lies in the structure of the matrices $\mathbb{A}$ and $\mathbb{B}$. To begin with, the calculation shows that $\mathbb{A}$ must be of rank one, which means that it can be written as an outer product of two vectors,

$$\mathbb{A} = uv^{\mathrm{T}}. \tag{3.39}$$

With this fact in mind, it is actually possible to construct a potential term $V(S)$ with powers of S, as terms with multiple powers of $1/N \cdot \mathbb{A}$ will vanish. Let us see how this can be deduced.

Since $\mathbb{A}$ has rank one, it means that it will act as projection operator to a one-dimensional subspace. Owing to this property, antisymmetric products of $\mathbb{A}$:s will be linear in $\mathbb{A}$. In d dimensions, inspect a term with multiple products of S:s with the following ansatz,

$$S^n = \left(\frac{1}{N}\mathbb{A} + \mathbb{B}\right)^n = \sum_{m=0}^{n} \binom{n}{m} \left(\frac{1}{N}\right)^m \mathbb{A}^l \mathbb{B}^{n-m}, \tag{3.40}$$

with binomial coefficients $\binom{n}{m} = n!/(m!(n-m)!)$. In component form this can be reformulated as

$$S^n = \sum_{m=0}^{n} \binom{n}{m} \left(\frac{1}{N}\right)^m \mathbb{A}^{\nu_1}{}_{\mu_1} \dots \mathbb{A}^{\nu_m}{}_{\mu_m} \mathbb{B}^{\nu_{m+1}}{}_{\mu_{m+1}} \dots \mathbb{B}^{\nu_n}{}_{\mu_n}. \tag{3.41}$$

Next, form a potential term with antisymmetric products of $\mathbb{A}$ as

$$V(S) = \sum_{n=0}^{d} b_n \epsilon^{\mu_1\mu_2 \dots \mu_n\lambda_{n+1} \dots \lambda_d} \epsilon_{\nu_1\nu_2 \dots \nu_n\lambda_{n+1} \dots \lambda_d} S^{\nu_1}{}_{\mu_1} \dots S^{\nu_n}{}_{\mu_n}, \tag{3.42}$$

where b_n are arbitrary coefficients and the ϵ:s are completely antisymmetric Levi-Civita tensors in d dimensions. Now, replace the product of S:s in the generic expression for the potential and arrive at

$$V(S) = \sum_{n=0}^{d} b_n \sum_{m=0}^{n} \binom{n}{m} \left(\frac{1}{N}\right)^m V_n(\mathbb{A}, \mathbb{B}), \tag{3.43}$$

where we have collected

$$\begin{aligned} V_n(\mathbb{A}, \mathbb{B}) &= \epsilon^{\mu_1\mu_2 \dots \mu_n\lambda_{n+1} \dots \lambda_d} \epsilon_{\nu_1\nu_2 \dots \nu_n\lambda_{n+1} \dots \lambda_d} \cdot \mathbb{A}^{\nu_1}{}_{\mu_1} \dots \mathbb{A}^{\nu_m}{}_{\mu_m} \mathbb{B}^{\nu_{m+1}}{}_{\mu_{m+1}} \dots \mathbb{B}^{\nu_n}{}_{\mu_n} \\ &\overset{(3.39)}{=} \epsilon^{\mu_1\mu_2 \dots \mu_n\lambda_{n+1} \dots \lambda_d} \epsilon_{\nu_1\nu_2 \dots \nu_n\lambda_{n+1} \dots \lambda_d} \cdot u^{\nu_1} w_{\mu_1} \dots u^{\nu_m} w_{\mu_m} \mathbb{B}^{\nu_{m+1}}{}_{\mu_{m+1}} \dots \mathbb{B}^{\nu_n}{}_{\mu_n}. \end{aligned} \tag{3.44}$$

Here, one observes that all indices of the symmetric product of the vectors v^{ν_i} are contracted with indices of the antisymmetric tensor $\epsilon_{\nu_1\nu_2 \dots \nu_n\lambda_{n+1} \dots \lambda_d}$. This yields zero if we have more than one v^{ν_i} in the product and thus the potential only contain non-zero terms with at most one $\mathbb{A}$. Hence, one could argue that following the redefinition of the shift the most general appearance of the potential should be (3.42). One can simplify this expression considerably by writing the product of S:s in terms of the elementary symmetric polynomials of the eigenvalues of the matrix S. A short introduction to this topic is furnished in appendix A. We define the n:th elementary symmetric polynomial, $e_n(S)$, for a matrix of dimension $d \times d$ following [16] as

$$e_n(S) \overset{\text{def}}{=} \frac{1}{d!}\binom{d}{n} \epsilon^{\mu_1 \dots \mu_n\lambda_{n+1} \dots \lambda_d} \epsilon_{\mu_1 \dots \mu_n\lambda_{n+1} \dots \lambda_d} S^{\nu_1}{}_{\mu_1} \dots S^{\nu_n}{}_{\mu_n}, \tag{3.45}$$

with $e_0(S) = 1$ and where $n \leqslant d$. The determinant of S is easily discerned as the d:th term,

$$\det S = \frac{1}{d!} \epsilon^{\mu_1\mu_2 \dots \mu_d} \epsilon_{\nu_1\nu_2 \dots \nu_d} S^{\nu_1}{}_{\mu_1} \dots S^{\nu_d}{}_{\mu_d}. \tag{3.46}$$

In addition, these polynomials can be acquired through a recursive formula

$$e_n(S) = \frac{(-1)^{n+1}}{n} \sum_{k=0}^{n-1} (-1)^k \operatorname{Tr}\left(S^{n-k}\right) e_k(S). \tag{3.47}$$

Explicitly, the first five terms in this expansion, relevant for four spacetime dimensions, are

$$\begin{aligned} e_0(\mathbb{X}) &= 1, & e_1(\mathbb{X}) &= [\mathbb{X}], \\ e_2(\mathbb{X}) &= \frac{1}{2}\left([\mathbb{X}]^2 - [\mathbb{X}^2]\right), & e_3(\mathbb{X}) &= \frac{1}{6}\left([\mathbb{X}]^3 - 3[\mathbb{X}^2][\mathbb{X}] + 2[\mathbb{X}^3]\right) \\ e_4(\mathbb{X}) &= \det(\mathbb{X}), \end{aligned} \tag{3.48}$$

with $\mathbb{X} = S$, which is a standard convention, and with square-brackets denoting the trace. Note the similarity between $e_2(\mathbb{X})$ and the Fierz-Pauli mass term in equation (3.3). The potential is then given by

$$\sqrt{-g}V(S) = \sqrt{-g}\sum_{n=0}^{d} \frac{n!(d-n)!}{2} b_n e_n(S) = \sqrt{-g}\sum_{n=0}^{d} \beta_n e_n(S), \tag{3.49}$$

where we have baked in the coefficients in β_n. A potential of this form, will make the Lagrangian linear in N as well as N^i, as previously mentioned, providing us with an additional constraint. However, an additional constraint is required in order to obliterate the degree of freedom corresponding to the canonical momentum of the pathological ghost-mode. This can be obtained by postulating that the already acquired constraint to be time invariant. Since the constraint does not have an explicit time dependence, its evolution is generated by its Poisson bracket with the Hamiltonian H, [66],

$$\dot{C}(x) = \{C(x),\, H\}, \tag{3.50}$$

where $C(x)$ denotes the constraint. Here, the Poisson bracket between two operators is defined as

$$\{A,\, B\} = \int \mathrm{d}^{d-1}x \left(\frac{\delta A}{\delta h_{ij}(x)}\frac{\delta B}{\delta \pi^{ij}(x)} - \frac{\delta A}{\delta \pi_{ij}(x)}\frac{\delta B}{\delta h^{ij}(x)}\right), \tag{3.51}$$

where the integration is performed over the constraint surface in phase space. More explicitly, equation (3.50) can be written as

$$\{C(x),\, H\} = \int \mathrm{d}^{d-1}x \left[\{C(x),\, \mathcal{H}_0(y)\} + N(y)\{C(x),\, C(y)\}\right], \tag{3.52}$$

where the constraint-less part of the Hamiltonian density is denoted by $\mathcal{H}_0$ and where the constraint is incorporated in $N \cdot C(x)$. That the first bracket yields zero follows from the previous discussion and that the second bracket vanishes was proved in [62]. We will not discuss it further here as the proof is slightly convoluted. This concludes our explorative tour of the basic premises of massive gravity where we have motivated the validity of an action of type,

$$S_{\mathrm{HR}} = \frac{M_g^2}{2}\int \mathrm{d}^4x\sqrt{-g}R(g) - 2m^2M_g^2\sqrt{-g}\sum_{n=0}^{4}\beta_n e_n(S), \tag{3.53}$$

where one has yet to introduce a source term. This is often denoted as the *dRGT action*. In their original papers, [59] [60], the authors followed a slightly different approach than the one presented here following [4] [62]. Later, the consistency proof has been reformulated in other formalisms, such as in the language of Stückelberg fields in [67]. Note that we have note yet arrived at bimetric gravity where both metrics are dynamical, which we will examine in the next section.

3.5 The bimetric action

Finally, we can take a first look at the bimetric action, also known as the *Hassan-Rosen action*, in four dimensions:

$$S_{\mathrm{HR}} = \frac{M_g^2}{2}\int \mathrm{d}^4x\sqrt{-g}R(g) + \frac{M_f^2}{2}\int \mathrm{d}^4x\sqrt{-f}R(f) - 2m^2M_g^2\sqrt{-g}\sum_{n=0}^{4}\beta_n e_n(S) \tag{3.54}$$

where we have imbued the dRGT action in (3.53) with a dynamical f metric term.

One of the first things one would like to investigate is potential symmetries associated with the general action and also those which emerge for special subclasses defined by parameter choices. Since e_0 simply yields the identity matrix, this means that the β_0-term is nothing but a cosmological constant term for the ordinary g-metric. Equally, the β_4-term gives a cosmological constant term for the f-metric as e_4 is the determinant. These two remarks point at a distinguishable feature of the bimetric action owing to a particular relation between the elementary symmetric polynomials, namely

$$\sqrt{g}\cdot e_n(S) = \sqrt{f}\cdot e_{d-n}(S), \text{ with } n = 0, \ldots, d, \tag{3.55}$$

which ensures that the action is invariant under the exchanges

$$g_{\mu\nu} \leftrightarrow f_{\mu\nu},\ M_g \leftrightarrow M_f,\ \beta_n \to \beta_{d-n} \tag{3.56}$$

which form a discrete symmetry, [5]. In the previous section, we worked with a theory where g bore the dynamics while f was fixed. The symmetry which we have just described entails that we similarly could promote f to a dynamical field with its own Einstein-Hilbert term and keep g fixed and reach the same result vice versa.

In the context of general covariance, the Einstein-Hilbert term of massive gravity shares the same symmetry as general relativity. However, the dRGT action as a whole loses this symmetry since the fixed reference metric appears in the mass term. As for bimetric gravity, its two Einstein-Hilbert terms are individually invariant under separate diffeomorphisms parametrized by ϕ and $\tilde{\phi}$,

$$g_{\mu\nu}(x) \to \frac{\partial\phi^\alpha}{\partial x^\mu}\frac{\partial\phi^\beta}{\partial x^\nu}g_{\alpha\beta}\left(\phi(x)\right),\ \ f_{\mu\nu}(x) \to \frac{\partial\tilde{\phi}^\alpha}{\partial x^\mu}\frac{\partial\tilde{\phi}^\beta}{\partial x^\nu}f_{\alpha\beta}(\tilde{\phi}(x)). \tag{3.57}$$

However, these symmetries are not completely lost when considering the mass term, as it simply reduces the overall symmetry to a subgroup of diagonal diffeomorphisms where $\bar{\phi} \stackrel{\text{def}}{=} \phi = \tilde{\phi}$, [16],

$$g_{\mu\nu}(x) \to \frac{\partial\bar{\phi}^\alpha}{\partial x^\mu}\frac{\partial\bar{\phi}^\beta}{\partial x^\nu}g_{\alpha\beta}(\bar{\phi}(x)),\ \ f_{\mu\nu}(x) \to \frac{\partial\bar{\phi}^\alpha}{\partial x^\mu}\frac{\partial\bar{\phi}^\beta}{\partial x^\nu}f_{\alpha\beta}(\bar{\phi}(x)). \tag{3.58}$$

This leads us to conclude that bimetric gravity is more symmetric than massive gravity and perhaps better suited to tackle the naturalness argument in future quantum ventures.

Still, to conclude that a theory of two dynamical metrics, massive bimetric gravity, is consistent, one must again perform an ADM analysis. A successful result was obtained in [5]. Briefly, this theory a priori has two Bouleware-Deser ghosts so one requires that the action is linear in the lapses N and $\tilde{N}$ for both metrics. As before, one redefines the shift variables, N^i and $\tilde{N}^i$, and this also brings about a Lagrangian which is linear in the shifts.

To conclude, the secondary constraint calculated in [62] is used to eliminate the degrees of freedom corresponding to the canonical momenta of the ghosts. This leads us to a bimetric theory which propagates $d(d-2)-1$ degrees of freedom in d dimensions. Physically, this could be interpreted as a massless and a massive spin-2 field. Yet, each mass eigenstate is not formed solely by one metric, but consists of contributions from the two. In the subsequent chapter when we will discuss bimetric cosmology, we will see how one could obtain explicit expressions for the two propagating modes through analysis of the equations of motion. The bimetric action is the most general action involving two dynamic metrics with interaction terms that do not contain any derivatives of them.

3.6 Matter couplings

The next natural step is to include matter terms in the action. In the so-called *single-coupled theory*, one adds a matter Lagrangian in the same way as in general relativity,

$$\sqrt{-g}\mathcal{L}^m(g,\,\Phi), \tag{3.59}$$

where Φ denotes different matter fields. Notice that these fields only couples directly to the g-metric, but there is an implicit relation to the f-metric through the interaction potential. One geometric interpretation is to refer to g as our ordinary spacetime metric similar to the one in GR where theory covers an additional coupling to a spin-2 field, f, which renders the gravitons massive. In the previous section we mentioned that the two propagating modes were built from a superposition of the two metrics and simply coupling matter exclusively to the combined massless mode has proved to be inconsistent [68]. The absence of ghost proof for the free bimetric theory in [5] and [62] can be extended in a straightforward manner to a single-coupled matter theory [68], and is explicitly presented in [69] in the ADM formalism. Moreover, the single-coupled theory does not provide any extra contributions that ruin the special form of the potential and resuscitate the Bouleware-Deser ghost for quantum corrections at one-loop order [70].

An option is to let different types of matter, for instance ordinary matter and dark matter, couple to different metrics. Another conceivable, but conceptually challenging, path is to couple the same type of matter to both metrics. Such theories are commonly referred to as *double-coupled*. Owing to the discrete symmetry between the two metrics for the free theory depicted in (3.56), such couplings would be advantageous from the symmetry aspect. Their phenomenology has been investigated in [71] and they pose intriguing questions on the nature of the relation between spacetime geometry and matter prevalence. Should one decide to couple matter to both metrics, which would then be the physical one? This acute situation could be compared to having two rulers for distance measurements with no knowledge of which one to use. The speed of light could be different in the two geometries, challenging the notion of causality. These fundamental issues dampened the interest in this class of theories, but they have undergone a revival during the last year. However, recently, it has been shown that such models generically bring pathologies both on the classical level [69] and for first-order quantum corrections [70] unless the couplings have specific forms. A similar analysis has also recently been conducted in the vielbein formalism [72] for this allowed double-coupling, which was explicitly shown to be ghost-free in [73]. The associated background cosmology have been analyzed a couple of days ago [74], but comparisons with experimental data have yet to be performed. It is possible to form an effective metric of g and f and obtain consistent results while coupling it to matter. Concerning the geometrical aspect, another possible way to resolve these dilemmas might be achieved in a departure from our Riemannian view of spacetime to a description founded in *Finsler geometry* [75]. Within this framework, the line element would be given by a *quasi-metric*, involving contributions from both g and f, which might share features with and serve as an interpretation an effective metric of theory. Much rests to be discovered and there are excellent conditions for auspicious research findings in the years to come.

Chapter 4 Bimetric phenomenology: Cosmology

With the basic cornerstones of bimetric gravity presented in chapter 3, we move now on to its physical, cosmological predictions. Firstly, we will show how to derive the equations of motions and proceed with a brief analysis of their implications on the cosmic background solutions. Then we will approach the purpose of this thesis, namely the investigation of large-scale structures in the bimetric framework.

4.1 Equations of motion

To derive the equations of motion, one proceeds in a straightforward manner by varying the action. To probe spacetime solutions, we set $d = 4$ and work with a single-coupled theory with matter coupled to the g-metric. Hence, our starting point is the action,

$$S = \int \mathrm{d}^4x \left[\frac{M_g^2}{2}\sqrt{-g}R(g) + \frac{M_f^2}{2}\sqrt{-f}R(f) - 2m^2M_g^2\sqrt{-g}\sum_{n=0}^{4}\beta_n e_n(S) + \sqrt{-g}\mathcal{L}^m(g,\, \Phi) \right], \tag{4.1}$$

with terms previously discussed in sections 3.5-3.6. As in general relativity, we postulate that the variation of the matter Lagrangian gives,

$$-\frac{1}{\sqrt{-g}}\frac{\delta\mathcal{L}^m(g,\, \Phi)}{\delta g_{\mu\nu}} = T^g_{\mu\nu}, \tag{4.2}$$

where $T^g_{\mu\nu}$ is the stress-energy tensor in the g-sector and with $T^f_{\mu\nu} = 0$. Note that in a double-coupled theory the matter variation is defined likewise. The derivation steps are exactly the same as when one varies the Einstein-Hilbert action with a non-zero stress-energy tensor to obtain Einstein's equations (see for instance [24]) with exception of the interaction terms, with the evident difference that one obtains two equations, one for g and one for f. Considering these interaction terms, there is a subtlety involved while evaluating the variation of the square-root matrix $S = \sqrt{g^{-1}f}$. Still, since this matrix only is present in traces in the elementary symmetric polynomials, there exists special caveats which we can use. To begin with, the variation of S^2 equals $\delta(S^2) = (\delta S)S + S(\delta S)$. If we plug this into the variation of the trace, one can reformulate it as $\delta(\operatorname{Tr} S) = \operatorname{Tr}(\delta S) = 1/2\operatorname{Tr}(S^{-1}\delta(S^2))$ since the trace is a linear transformation. Subsequently one proceeds to higher powers of S [61]. Together with the recursive relation of the elementary symmetric polynomials in (3.47), the variation of the interaction terms is given by tensors of the following form

$$
\begin{aligned}
{V_{\mu\nu}}^{\text{int},\,g}(S) &= \sum_{n=0}^{d-1}(-1)^n\beta_n g_{\mu\rho}\left(Y^{(n)}\right)^{\rho}{}_{\nu}(S)\\
{V_{\mu\nu}}^{\text{int},\,f}(S) &= \sum_{n=0}^{d-1}(-1)^n\beta_{d-n} f_{\mu\rho}\left(Y^{(n)}\right)^{\rho}{}_{\nu}(S^{-1})
\end{aligned}
\tag{4.3}
$$

where we have also incorporated the symmetry in (3.56) and whose matrix functions $Y^{(n)}$ are defined according to

$$
\left(Y^{(n)}\right)^{\mu}{}_{\nu}(S) = \sum_{k=0}^{n}(-1)^k e_k(S)\left(S^{n-k}\right)^{\mu}{}_{\nu}. \tag{4.4}
$$

In addition, we have used the fact that the tensors are symmetric in (4.3) [68]. We readily obtain our terms in four dimensions as the first four terms in the series. Explicitly, they are given by

$$
\begin{aligned}
Y^{(0)}(S) &= I\\
Y^{(1)}(S) &= S - [S]I\\
Y^{(2)}(S) &= S^2 - [S]S + \frac{1}{2}\left([S]^2 - [S^2]\right)I\\
Y^{(3)}(S) &= \frac{1}{6}\left(6S^3 + 3S[S]^2 - [S]^3 I - 3S[S^2] + [S](-6S^2 + 3[S^2]I) - 2[S^3]I\right),
\end{aligned}
\tag{4.5}
$$

where I is the unit matrix and the square-brackets denote the trace as before.

Later on when we start to examine the perturbations required for structure formation, we must perform an additional variation of the action, now involving perturbed metric quantities as $g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu}$ and $f_{\mu\nu} \to f_{\mu\nu} + \delta f_{\mu\nu}$. In that case, the variation of the square-root of a perturbed matrix emerges as a delicate affair [76]. Fortunately, our ansätze for the two unperturbed metrics will be isotropic and homogeneous, i.e. of FLRW-type, and then we can compute these terms through a series expansion of the square-root, provided that the perturbations are small. In terms of matrix components one acquires

$$
\left(\sqrt{M+\epsilon}\right)_{\mu\nu} = \sqrt{M_{\mu\nu}} + \frac{\epsilon_{\mu\nu}}{\sqrt{M_{\mu\mu}} + \sqrt{M_{\nu\nu}}} + \mathcal{O}\left(\epsilon^2\right), \tag{4.6}
$$

where M is the unperturbed, diagonal background matrix and ϵ the matrix with the small perturbations. One then replaces S with this expansion in (4.3) to reach the equations of motion for the perturbations while neglecting all terms of order $\mathcal{O}\left(\epsilon^2\right)$.

To simplify the analysis one clearly observes that the action in (4.1) can be subjected to a rescaling,

$$
f_{\mu\nu} \to \left(\frac{M_g}{M_f}\right)^2 f_{\mu\nu} \text{ and } \beta_n \to \left(\frac{M_f}{M_g}\right)^n \beta_n, \tag{4.7}
$$

which implies that the ratio between the two Planck masses in each sector M_f/M_g is a redundant variable which can be set to unity. Note that this combination would still appear in the equations of motion for the f-sector in a double-coupled theory. Henceforth, we will refer to this ratio as $M_\star$ following the convention in [6]. With this in mind, we arrive at the following equations of motion,

$$G_{\mu\nu}(g) + m^2 \sum_{n=0}^{3} (-1)^n \beta_n g_{\mu\rho} \left(Y^{(n)}\right)^{\rho}{}_{\nu}(S) = \frac{1}{M_g^2} T^g_{\mu\nu} \tag{4.8}$$

$$G_{\mu\nu}(f) + m^2 \sum_{n=0}^{3} (-1)^n \beta_{4-n} f_{\mu\rho} \left(Y^{(n)}\right)^{\rho}{}_{\nu}(S^{-1}) = 0, \tag{4.9}$$

where $G_{\mu\nu}(g) \stackrel{\text{def}}{=} R_{\mu\nu} - 1/2 g_{\mu\nu} R(g)$ is the Einstein tensor in each sector. During the successive treatment of the perturbations the Einstein tensors will transform as $G_{\mu\nu}(g) \to G_{\mu\nu}(g) + \delta G_{\mu\nu}(g)$ as the metrics transform as $g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu}$.

In general relativity, the Bianchi identity $\nabla^\mu G^g_{\mu\nu} = 0$, where ∇ is the covariant derivative with respect to the g metric, implies that the stress-energy tensor obeys $\nabla^\mu T^g_{\mu\nu} = 0$, i.e. it is covariantly conserved, as we saw in equation (2.27). Hence, the conservation of the sources is a consequence of the equations of motion. The result is mimicked in bimetric theory, yet with a resultant weaker condition due to the interaction terms. Intuitively, the Einstein tensor in the f-sector also satisfies the Bianchi identity $\tilde{\nabla}^\mu G_{\mu\nu} = 0$ where $\tilde{\nabla}$ is the covariant derivative with respect to f. Since the interaction potential remains invariant under simultaneous diffeomorphisms of g and f, the following relation applies

$$\sqrt{-g}\nabla^\mu V_{\mu\nu}{}^{\text{int},\,g}(S) + \sqrt{-f}\tilde{\nabla}^\mu V_{\mu\nu}{}^{\text{int},\,f}(S) = 0, \tag{4.10}$$

from which one infers the following general constraint on the sources

$$\sqrt{-g}\nabla^\mu T^g_{\mu\nu} + \sqrt{-f}\tilde{\nabla}^\mu T^f_{\mu\nu} = 0. \tag{4.11}$$

Still, we will impose that the two stress-tensors are conserved separately, which in other words means that each matter Lagrangian is separately invariant under diffeomorphisms. Here, we do not have any matter sources coupled to f. If the stress-energy tensors are conserved individually, this means that there are four constraints conveyed by the equations of motion in $d = 4$, namely

$$\nabla^\mu \sum_{n=0}^{3} (-1)^n \beta_n g_{\mu\rho} \left(Y^{(n)}\right)^{\rho}{}_{\nu}(S) = 0. \tag{4.12}$$

Through (4.10) one notices that they are the same in each sector. Equivalently, in d dimensions one obtains d constraints and these relations are known as *Bianchi constraints*. We will frequently refer to them in the sections which follow. It should be stressed that these conditions arise from the equations of motion. Hence, they do not provide any additional information. However, we will discover that it occasionally is more clear to replace some of the equations with constraints to obtain a fruitful result without lengthy calculations.

4.2 Background solutions

Having established the equations of motion, we will now investigate explicit cosmological solutions. Chiefly the discussion will be based on material in [6] interspersed with complimentary results. As our goal is to compare the phenomenological predictions to those of general relativity, we make an FLRW ansatz for the matter-coupled metric g,

$$\mathrm{d}s_g^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu = -N(t)^2\,\mathrm{d}t^2 + a(t)^2\,\mathrm{d}\boldsymbol{x}^2, \tag{4.13}$$

where $N(t) = 1$ for cosmic time and $N(t) = a(t)$ for conformal time following the notation introduced in [11]. For the f-metric, we assume a more general isotropic and homogeneous metric on the form

$$\mathrm{d}s_f^2 = f_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu = -X(t)^2\,\mathrm{d}t^2 + Y(t)^2\,\mathrm{d}\boldsymbol{x}^2, \tag{4.14}$$

where $X(t)$ and $Y(t)$ are arbitrary, smooth scalar functions. The spatial part $\mathrm{d}\boldsymbol{x}^2$ is given by

$$\mathrm{d}\boldsymbol{x}^2 = \frac{\mathrm{d}r^2}{1-kr^2} + r^2\,\mathrm{d}\Omega^2, \text{ with } \mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2, \tag{4.15}$$

as in ordinary FLRW. We will abbreviate the expressions below by not writing out the t-dependence of the relevant terms. If we insert these two metric ansätze into (4.12), one concludes that the time-component of the Bianchi constraint yields[1]

$$\frac{3m^2}{a}\left(\beta_1 + 2y\beta_2 + y^2\beta_3\right)\left(\dot{Y} - \dot{a}x\right) = 0, \tag{4.16}$$

where the dot indicates a time derivative and where we have introduced the variables $y = Y/a$ and $x = X/N$ following [11]. Obviously, there are two ways to ensure that this relation is satisfied, i.e. one parenthesis has to be equal to zero. If the first one gives a trivial result, it implies solutions where $y \propto$ constant. As a consequence, the calculation of the equations of motion will generate the ordinary relativistic equations of general relativity with a cosmological constant of order m^2 and, moreover, the mass of the massive spin-2 field will disappear [6]. Physically, this means that such a class of solutions cannot be distinguished from ordinary general relativity at least on the background level. This motivates us to examine solutions where the second parenthesis evaluates to zero, i.e. when

$$x = \frac{\dot{Y}}{\dot{a}} = \frac{\mathrm{d}Y}{\mathrm{d}a} = \frac{Ky}{H}, \tag{4.17}$$

with the defined combinations $\dot{Y}/Y = K$ and $\dot{a}/a = H$ which we will use frequently later on[2]. Since $K = \dot{y}/y + H$, the important functions for the background are the scale factor ratio $y = y(t)$ and the cosmic scale factor $a = a(t)$. As for the source terms, we assume a perfect fluid in thermal equilibrium, which implies that the stress-energy tensor appear as in (2.23), with diagonal components $T^0_{\ 0} = -\bar{\rho}$ and $T^i_{\ i} = \bar{p}$ in the rest frame (no sum implied). Taking this into account, calculating the equations of motion produces the modified Friedmann equations,

$$3H^2 - N^2m^2\left(\beta_3y^3 + 3\beta_2y^2 + 3\beta_1y + \beta_0\right) + \frac{kN^2}{a^2} = \frac{N^2\bar{\rho}}{M_g^2}, \tag{4.18}$$

in the g-sector and

$$3K^2 - N^2m^2x^2\left(\beta_4 + 3\beta_3y^{-1} + 3\beta_2y^{-2} + \beta_1y^{-3}\right) + \left(\frac{K}{H}\right)^2\frac{kN^2}{a^2} = 0, \tag{4.19}$$

in the f-sector [6] where we have used (4.17). If we restrict ourselves to a spatially flat universe, we can set $k = 0$ and eliminate the corresponding terms. Through further manipulation with

[1] This calculation is easily performed by writing appropriate `Mathematica`-modules while considering the background geometry.

[2] Some authors prefer to include this relation directly in the metric ansatz, see for instance [12].

help from (4.17), it is possible to subtract these two equations from one another and obtain a single quartic algebraic equation for y,

$$\beta_3 y^4 + (3\beta_2 - \beta_4)\, y^3 + 3\,(\beta_1 - \beta_3)\, y^2 + \left(\frac{\bar{\rho}}{m^2 M_g^2} + \beta_0 - 3\beta_2\right) y - \beta_1 = 0. \tag{4.20}$$

Hence, y is determined by ρ, but we can specify this further by demanding that the source obeys the continuity equation. If we assume that the system has an equation of state $\bar{p} = \omega\bar{\rho}(t)$, the equation is the same as in (2.28), where the relation $\dot{\bar{\rho}} = -3H\bar{\rho}$ applies for pressure-less dust such as in (2.29). Here, one can proceed along two paths, either solve for $\bar{\rho}$ directly from this fluid equation (2.28) or replace it in (4.20) and its differentiated version (for pressure-less dust). In the first case, one then solves for y using the quartic equation and then inserts this into either (4.18) or (4.19) to solve for H^2. The latter choice propels us to differential equation for y in terms of β_i, whereupon one of the remaining Friedman equations is used to solve for H^2. This method is for instance used in [77], [78], [11] and [12]. For a zero-curvature model with pressure-less dust, we readily read off $\bar{\rho}$ from (4.20) and are able to find expressions for $\dot{y}$ and H through

$$\bar{\rho} = m^2 M_g^2 \left[-\beta_3 y^3 + (\beta_4 - 3\beta_2) y^2 + 3(\beta_3 - \beta_1) y + 3\beta_2 - \beta_0 + \frac{\beta_1}{y}\right], \tag{4.21}$$

$$\frac{\dot{y}}{y} = -3H \frac{\beta_3 y^4 + (3\beta_2 - \beta_4) y^3 + 3(\beta_1 - \beta_3) y^2 + (\beta_0 - 3\beta_2) y - \beta_1}{3\beta_3 y^4 + 2(3\beta_2 - \beta_4) y^3 + 3(\beta_1 - \beta_3) y^2 + \beta_1}, \tag{4.22}$$

$$H^2 = m^2 N^2 \left[\frac{\beta_4}{3} y^2 + \beta_3 y + \beta_2 + \frac{\beta_1}{3}\frac{1}{y}\right]. \tag{4.23}$$

To facilitate comparisons with experimental data, one may normalize the β_i and the density $\bar{\rho}$ to the present day Hubble-rate H_0 following [79],

$$\left(\frac{m}{H_0}\right)^2 \beta_i \to \beta_i, \quad \left(\frac{1}{H_0}\right)^2 \bar{\rho} \to \bar{\rho}, \tag{4.24}$$

rendering the β_i dimensionless. As stated, to obtain an equation for $y(t)$ we can solve the differential equation or fix $\bar{\rho}$ which gives a solution directly. The latter choice is easier since we can put bounds on $\bar{\rho}$ directly and we do not have to translate this information to constrain the integration constants. Still, another subtlety arises due to the fact that equation (4.20) may have as many as four solutions which introduces a conundrum regrading which solution is valid for which times. By integrating the differential equation in (4.22) one ensures that a proper solution is reached and that a real solution stays real. In chapter 2, we noted the current density parameter values. Since we are interested in investigating whether bimetric gravity can supplant dark energy, we can simply insert the remaining densities into from (2.37) into $\bar{\rho}$ while neglecting the Ω_k in (2.36), because we have the same matter coupling. This leads to a relation[3]

$$\bar{\rho}(a) = \Omega_m a^{-3} + \Omega_\gamma a^{-4} \approx 0.317 a^{-3} + 10^{-5} a^{-4}, \tag{4.25}$$

corresponding to Ω_m and Ω_γ as measured by Planck. Here, we have assumed that the physical Planck mass and the g sector Planck mass are equal, $M_{\mathrm{PL}} = M_g$. Then, one inserts this

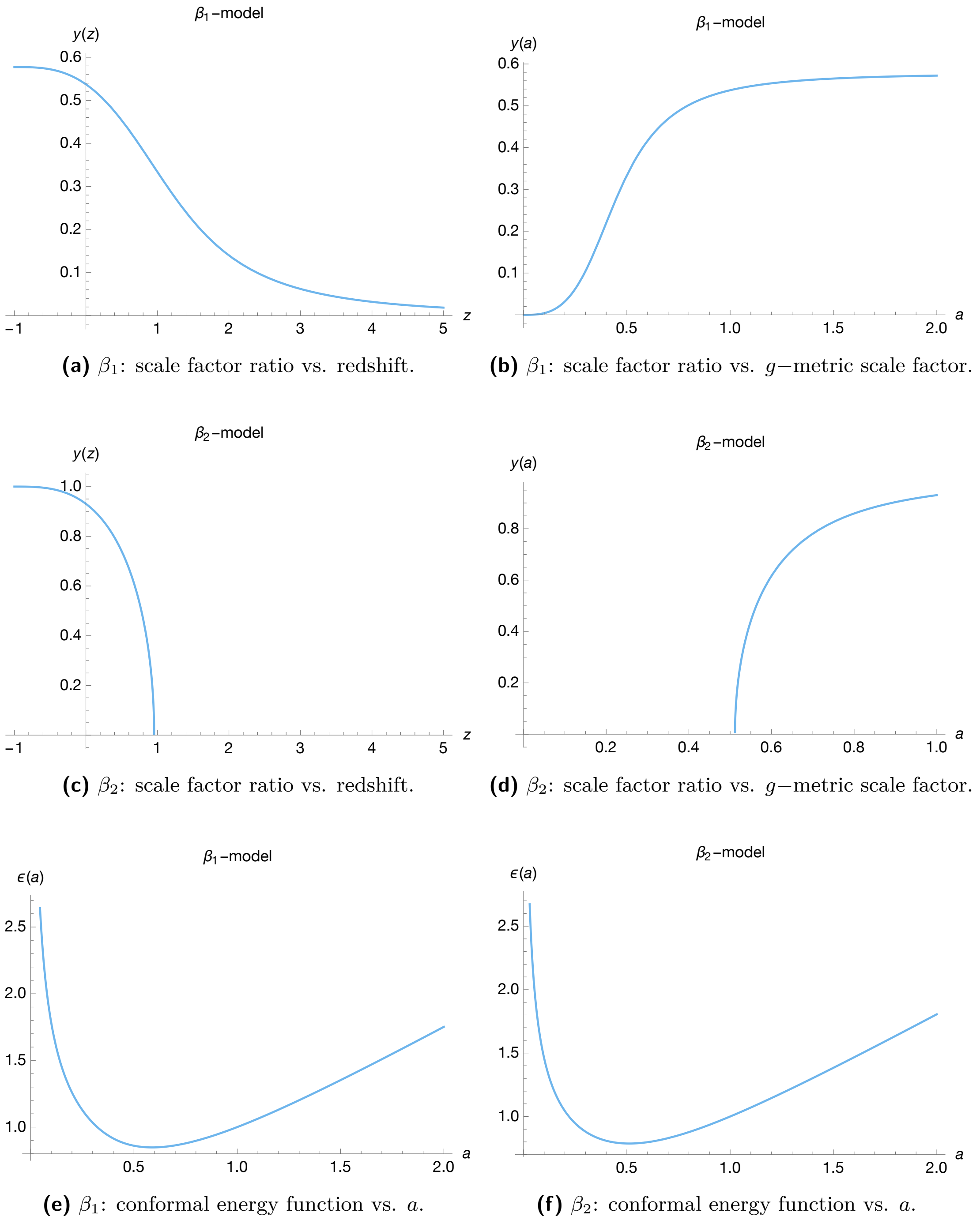
[3]The factor of three has been included in the definition of $\bar{\rho}$.

density into (4.21) and the y solution into (4.23). To proceed, we use relation (2.40). As we remarked previously, $\mathcal{E}(a = 1) = 1$, due to that we normalize by the present expansion rate. This allows us to constrain the β_i-values in (4.23). Regrettably, this leads to a solution which is not very transparent for arbitrary β_i. We will plot the solutions for some of the one-parameter models. In figure 4.1 the finite branch solutions for y in the one-parameter β_1- and β_2-models are presented. Equation (4.20) has multiple solutions and those which produce finite values for y in the far future, $a \to +\infty$ or equivalently $z \to -1$ following (2.10), are referred to as *finite branches*. The deep past corresponds to $a \to 0$, $z \to +\infty$, where the solution is undefined at $a = 0$. For the β_1-model, the ratio y is effectively zero for small a, as observed in figures 4.1(a) and 4.1(b). This is a property shared with other models and the cause can be found by studying the asymptotic solutions of (4.20). At early times the $\bar{\rho}(a)$ term is large, dominating all other terms which forces a solution $y \to 0$ as $a \to 0$. Concerning β_1, it is the one-parameter model which best fits supernovae data [79] [80], and is usually called the *minimal bimetric model*. However, it disagrees with observations for linear perturbations in the subhorizon limit [11] and produces instabilities far in the past [78].

The β_2-model has the nice feature of producing a constant ratio $y \approx 1$ at late times, which means that one obtains a de Sitter solution for the theory. A de Sitter solution is characterized by a constant H_0 in cosmic time and by a solution proportional to a in conformal time and we see that this is true in 4.1(f), where the late time curve's asymptot is a. Such a solution allows for an analytic[4] solution of the equations of motion for linear scalar perturbations [7]. Yet, it has a peculiar zero at $a \approx 0.5$ where it switches branches to the $y = 0$ solution, also permitted by (4.20). One may interpret this as if the past was solely governed by one metric g as in general relativity ($y = 0$ implies that there is no f-metric) until a point where the change $y \neq 0$ turns on the f-metric, from which onwards the whole system evolves towards a de Sitter universe. An idea is that this may be related to the filtering concept of degravitation, which we discussed in section 3.1. The transition in figures 4.1(c) and 4.1(d) is rather abrupt and thus one might consider including additional β_i-terms to mitigate it, fitting for a more viable physical theory. With our choice of $\bar{\rho}$ neither the β_3- nor the β_4-model provide valid solutions at $a = 1$.

Regarding the more interesting, multi-β_i models, we do not have sufficient information from the equations to fix both two parameters, but we have to express one in terms of the other. Validity conditions have been adressed in [80] and we will attempt to follow the analyses conducted in [11] and [12]. Still, their parameter values are obtained from comparison with supernovae and structure growth data whereas our experimental input comes through density factors. Hence, there will be a qualitative agreement and parameter values will differ slightly. At present, the infinite branch $\beta_1\beta_4$-model seems to be the two-parameter model which best fits experimental data while avoiding instabilities on the background level as well as for perturbations [11][12]. Its stability condition has been found as $0 < \beta_4 < 2\beta_1$ [12] with best fits comparing with growth data as $\beta_1 = 0.48$, $\beta_4 = 0.94$. To obtain the proper solution branch, we set $\beta_4 = 0.94$ in (4.21) and use the normalization condition in (4.23) to solve for β_1. With the Planck data, $\beta_4 = 0.94$ produces a solution $\beta_1 = 0.46$ which violates the stability argument and does not yield a y solution for late times, see figure 4.2(a). However, if we shift the values slightly by for instance setting $\beta_4 = 0.9$, an allowed β_1-value is acquired

[4] This can be surmised from the structure of equations (4.28)-(4.35) in section 4.3 as the time-derivatives of the background variables vanish and one can form linear combinations of the fields which yield one mode that couples to matter and one which does not.



(a) β_1: scale factor ratio vs. redshift.

(b) β_1: scale factor ratio vs. $g-$metric scale factor.

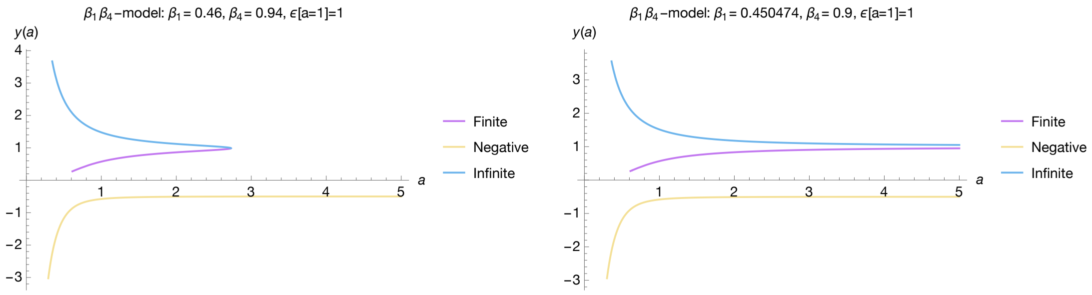
(c) β_2: scale factor ratio vs. redshift.

(d) β_2: scale factor ratio vs. $g-$metric scale factor.

(e) β_1: conformal energy function vs. a.

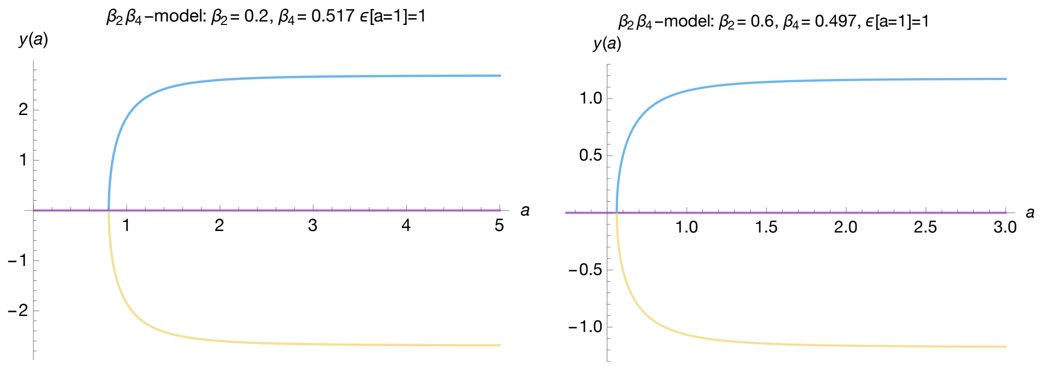
(f) β_2: conformal energy function vs. a.

Figure 4.1: Solutions for y in the one-parameter β_1-model with $\beta_1 \approx 1.27$ and the β_2-model with $\beta_2 \approx 0.79$. Note that $\mathcal{E}(a)$ in conformal time is $a \cdot \mathcal{E}(a)$ using ordinary cosmic time.



(a) $\beta_1\beta_4$: scale factor ratio solution branches. **(b)** $\beta_1\beta_4$: scale factor ratio solution branches.

Figure 4.2: Stable and unstable parameter combinations.



(a) $\beta_2\beta_4$: scale factor ratio solution branches. **(b)** $\beta_2\beta_4$: scale factor ratio solution branches.

Figure 4.3: The “on”-solution shifts with different sets of β_2- and β_4-values.

and the y solution will be stable in the future as illustrated in figure 4.2(b). We deem that the negative solutions are not viable as a negative scale factor would not make sense for the f-metric ansatz. As for the $\beta_2\beta_4$-model, we find that it features the same on/off-behavior as the pure β_2-model. However, different but still consistent parameter combinations could induce the transition to occur at different times, as seen in figures 4.3(a) and 4.3(b) in figure 4.3.

4.3 Scalar perturbations

Serving as a catalysis for structure formation, we will embark on the treatment of linear scalar perturbations in bimetric gravity. Similar expressions for the tensor perturbations are listed in appendix B. For the perturbed metrics we make an ansatz

$$\mathrm{d}s^2_{\bar{g}+\delta g^{(\mathrm{s})}} = -N^2\left(1+E_g\right)\mathrm{d}t^2 + 2Na\partial_i F_g\,\mathrm{d}t\,\mathrm{d}x^i + a^2\left[\left(1+A_g\right)\delta_{ij} + \partial_i\partial_j B_g\right]\mathrm{d}x^i\,\mathrm{d}x^j, \quad (4.26)$$

$$\mathrm{d}s^2_{\bar{f}+\delta f^{(\mathrm{s})}} = -X^2\left(1+E_f\right)\mathrm{d}t^2 + 2XY\partial_i F_f\,\mathrm{d}t\,\mathrm{d}x^i + Y^2\left[\left(1+A_f\right)\delta_{ij} + \partial_i\partial_j B_f\right]\mathrm{d}x^i\,\mathrm{d}x^j, \quad (4.27)$$

where the perturbation variables $\{E_{g,f}, F_{g,f}, A_{g,f}, B_{g,f}\}$ depend on both space and time. This ansatz connects to Weinberg's notation in [22], which we investigated in section 2.2.2, and was adapted to massive bigravity in [7] as well as in [11]. In the matter sector, we demand the same form for the perturbed stress-energy tensor as in (2.64) with F_g instead of F. We also define the velocity divergence as $\theta \overset{\text{def}}{=} \partial_i u^i$. Moreover, to facilitate the analysis of different wave modes we Fourier transform the spatial terms as before, $\partial_i \to -ik_i$, $k^2 = \boldsymbol{k}^2 = k_i^2 + k_j^2 + k_k^2$. We demand that the wave modes are isotropic. By varying the action (4.1), keeping terms up to linear order in the perturbations, and omitting the phase factor $\exp(-i\boldsymbol{k}\cdot\boldsymbol{x})$, we arrive at the following equations of motion [5]:

- $\delta g^{(\mathrm{s})}_{0-0}$:

$$\frac{3H}{N^2}\left(HE_g - \dot{A}_g\right) - k^2\left[\frac{A_g}{a^2} + \frac{H}{N}\left(\frac{2F_g}{a} - \frac{\dot{B}_g}{N}\right)\right] + \frac{m^2 yP}{2}\left(3\Delta A - k^2\Delta B\right) = \frac{\bar{\rho}\delta}{M_g^2}, \tag{4.28}$$

- $\delta g^{(\mathrm{s})}_{0-i}$:

$$ik_i\left[\frac{1}{N^2}\left(HE_g - \dot{A}_g\right) - m^2\frac{a}{N}\frac{Py}{x+y}\left(xF_f - yF_g\right)\right] = -\frac{\theta}{ik_i M_g^2}, \tag{4.29}$$

- $\delta g^{(\mathrm{s})}_{i-i}$ (spatial trace):

$$\begin{aligned}&\frac{1}{N^2}\left[\left(2\dot{H} + 3H^2 - 2\frac{\dot{N}}{N}H\right)E_g + H\dot{E}_g - \ddot{A}_g - 3H\dot{A}_g + \frac{\dot{N}}{N}\dot{A}_g\right] + \\ &m^2\left[\frac{xP}{2}\Delta E + yQ\Delta A\right] = 0,\end{aligned} \tag{4.30}$$

- $\delta g^{(\mathrm{s})}_{i-j}$ (off-diagonal components):

$$\frac{k^2}{2}\left[\frac{A_g + E_g}{a^2} + \frac{H}{N}\left(\frac{4F_g}{a} - \frac{3\dot{B}_g}{N}\right) + \frac{2\dot{F}_g}{aN} - \frac{1}{N^2}\left(\ddot{B}_g - \frac{\dot{N}}{N}\dot{B}_g\right) + m^2 yQ\Delta B\right] = 0, \tag{4.31}$$

- $\delta f^{(\mathrm{s})}_{0-0}$:

$$\frac{3K^2}{x^2N^2}\left(KE_f - \dot{A}_f\right) - k^2\left[\frac{A_f}{y^2a^2} + \frac{K}{xN}\left(\frac{2F_f}{ya} - \frac{\dot{B}_f}{xN}\right)\right] - \frac{m^2P}{2y^3}\left(3\Delta A - k^2\Delta B\right) = 0, \tag{4.32}$$

- $\delta f^{(\mathrm{s})}_{0-i}$:

$$-\frac{ik_i}{xN}\left[\frac{1}{xN}\left(KE_f - \dot{A}_f\right) + \frac{m^2P}{y^2}\frac{1}{x+y}\left(yF_g - xF_f\right)\right] = 0, \tag{4.33}$$

[5] In this report this calculation has been performed in `Mathematica` through the creation of a designated script involving components of the background geometry using the standard formulae for the connection, covariant derivative, Ricci tensor, scalar curvature etc. plus the additional interaction terms required in (4.8). Regarding the terms from the Einstein tensor in the $g-$sector, they are the same as those in general relativity, see equations (2.82)-(2.85). The expression for the perturbed square-root matrix was simplified using the expansion in (4.6). In total, this verifies the result published in [11].

- $\delta f^{(\mathrm{s})}_{i-i}$ (spatial trace):

$$\frac{1}{x^2N^2}\left[\left(2\dot{K}+3K^2-2\left(\frac{\dot{x}}{x}+\frac{\dot{N}}{N}\right)K\right)E_f+K\dot{E}_f-\ddot{A}_f-3K\dot{A}_f+\left(\frac{\dot{x}}{x}+\frac{\dot{N}}{N}\right)\dot{A}_f\right]+ \\ \frac{m^2}{xy^2}\left[\frac{P}{2}\Delta E+Q\Delta A\right]=0, \tag{4.34}$$

- $\delta f^{(\mathrm{s})}_{i-j}$ (off-diagonal components):

$$\begin{aligned}\frac{k^2}{2}\Bigg[&\frac{A_f+E_f}{y^2a^2}+\frac{K}{xN}\left(\frac{4F_f}{ya}-\frac{3\dot{B}_f}{xN}\right)+\frac{2\dot{F}_f}{xayN}- \\ &\frac{1}{x^2N^2}\left(\ddot{B}_f-\left(\frac{\dot{x}}{x}+\frac{\dot{N}}{N}\right)\dot{B}_f\right)-\frac{m^2Q}{xy^2}\Delta B\Bigg]=0,\end{aligned} \tag{4.35}$$

with

$$P \stackrel{\mathrm{def}}{=} \beta_1+2\beta_2y+\beta_3y^2, \tag{4.36}$$

$$Q \stackrel{\mathrm{def}}{=} \beta_1+\beta_2(x+y)+\beta_3xy, \tag{4.37}$$

$$\Delta A \stackrel{\mathrm{def}}{=} A_f-A_g, \tag{4.38}$$

$$\Delta E \stackrel{\mathrm{def}}{=} E_f-E_g, \tag{4.39}$$

$$\Delta B \stackrel{\mathrm{def}}{=} B_f-B_g. \tag{4.40}$$

The associated Bianchi constraints from (4.12) for these perturbations are

$$\begin{aligned}\nabla^{\mu}_{\mathrm{pert}}\sum_{n=0}^{3}(-1)^n\beta_n g^{\mathrm{pert}}_{\mu\rho}\left(Y^{(n)}\right)^{\rho}_{\ 0}(S_{\mathrm{pert}}) = m^2P\Bigg[&\frac{k^2x}{x+y}\frac{N}{a}\left(yF_f-xF_g\right)+ \\ \frac{1}{2}\left(3\left\{y\dot{A}_f-x\dot{A}_g\right\}-k^2\left\{y\dot{B}_f-x\dot{B}_g\right\}+yK\left\{3\Delta A-3\Delta E-k^2\Delta B\right\}\right)&\Bigg]=0,\end{aligned} \tag{4.41}$$

and

$$\begin{aligned}\nabla^{\mu}_{\mathrm{pert}}\sum_{n=0}^{3}(-1)^n\beta_n g^{\mathrm{pert}}_{\mu\rho}\left(Y^{(n)}\right)^{\rho}_{\ i}(S_{\mathrm{pert}}) = ik_im^2\Bigg[&-\frac{P}{2}\left(xE_f-yE_g\right)-yQ\Delta A \\ -\frac{y}{x+y}\frac{a}{N}&\left[P\left(\left\{x\dot{F}_f-y\dot{F}_g\right\}-4H\left\{xF_f-yF_g\right\}\right)+\dot{P}\left(xF_f-yF_g\right)\right] \\ &-\frac{P}{(x+y)^2}\frac{a}{N}\left(\dot{y}x^2F_f-\dot{y}y^2F_g-2\dot{y}xyF_g+\dot{x}y^2\left\{F_f+F_g\right\}\right) \\ +\frac{P}{x+y}\left[\left(\left(\frac{a}{N}\right)^3-\frac{N}{a}\right)\left(y^2\left\{KF_f-HF_g\right\}\right)-\left(\frac{a}{N}-\frac{N}{a}\right)Hx^2F_g\right]&\Bigg]=0,\end{aligned} \tag{4.42}$$

where the last line vanishes for conformal time. Their reciprocals expressed in terms of the stress-energy tensor defined in (2.64) are

$$\nabla_{\mu}T^{\mu}_{\ \ 0} = \dot{\delta}+\theta+\frac{3}{2}\dot{A}_g-\frac{k^2}{2}\dot{B}_g+k^2F_g=0, \tag{4.43}$$

and

$$\nabla_\mu T^\mu{}_i = \dot{\theta} + \left[\frac{\dot{N}}{N} - H + \left(\frac{a}{N}\right)^2 H\right]\theta + \frac{k^2}{2}E_g + H\frac{a}{N}\frac{k^2}{N}\left[N - a\right]F_g = 0, \tag{4.44}$$

where the last expression is reduced to

$$\nabla_\mu T^\mu{}_i = \dot{\theta} + H\theta + \frac{k^2}{2}E_g = 0, \tag{4.45}$$

in conformal time. With these expressions, we are now ready to start the solution machinery and, for convenience, we will work in conformal time.

4.3.1 Subhorizon solution and WKB approximation

The subhorizon case for these equations have been throughly explored in [11], where one deduces the limit as in general relativity in section 2.2, namely $(k/a)\Phi \gg H^2\Phi \sim H\dot{\Phi} \sim \ddot{\Phi}$, i.e. every differentiation supplies an additional factor of H, which will be more visible in the next chapter in equation (5.8) when we transform these derivatives to be of the scale factor a. Moreover, it is sensible that K should scale in the same way, i.e. $K \sim H$ and that terms $A_{g,\,f}$ and $E_{g,\,f}$ should be of the same order as $kF_{g,\,f}$ and $k^2B_{g,\,f}$ as they take this form in the line elements in (4.26) and it would be odd if it were otherwise. In Newton's gauge, where $F_g = B_g = 0$ for the g-metric exactly as in general relativity, taking this limit, neglecting the F_f-terms in the $\delta g^{(\mathrm{s})}_{0-i}$ and $\delta f^{(\mathrm{s})}_{0-i}$-equations, (4.29) and (4.33), while using (4.43) and (4.45) produces

$$\left(\frac{k}{a}\right)^2\left(A_g + \frac{m^2ya^2}{2}PB_f\right) - \frac{3m^2yP}{2}\Delta A = \frac{\bar{\rho}\delta}{M_g^2}, \tag{4.46}$$

from $\delta g^{(\mathrm{s})}_{0-0}$, (4.28), and for the trace $\delta g^{(\mathrm{s})}_{i-i}$, (4.30),

$$\left(\dot{H} - H^2 + \frac{a^2\bar{\rho}}{2M_g^2}\right)E_g + m^2a^2\left(\frac{xP}{2}\Delta E + yQ\Delta A\right) = 0, \tag{4.47}$$

and for the off-diagonal components $\delta g^{(\mathrm{s})}_{i-j}$, (4.31),

$$A_g + E_g + m^2a^2yQB_f = 0. \tag{4.48}$$

In the f-sector, the corresponding equations for $\delta f^{(\mathrm{s})}_{0-0}$, $\delta f^{(\mathrm{s})}_{i-i}$ and $\delta f^{(\mathrm{s})}_{i-j}$, (4.32), (4.34) and (4.35), are

$$\left(\frac{k}{a}\right)^2\left(A_f - \frac{m^2a^2}{2y}PB_f\right) + \frac{3m^2P}{2y}\Delta A = 0, \tag{4.49}$$

$$-\left[\dot{K} - \left(H + \frac{\dot{x}}{x}\right)K\right]E_f + m^2a^2\frac{x}{y}\left(\frac{P}{2}\Delta E + Q\Delta A\right) = 0, \tag{4.50}$$

$$A_f + E_f - \frac{m^2a^2}{x}QB_f = 0, \tag{4.51}$$

which means that we are left with a system of algebraic variables which source the evolution equation for δ in (4.43). Deviations from general relativity can be evaluated in the subhorizon large-scale structure context in terms of a couple of parameters which can be observed by experiments like `Euclid`. In [11], they investigate three primary probes highlighted by `Euclid`, [81]; the *growth rate* of large-scale structures $f(a,\,k)$ which is related to $\dot{\delta}$ through

$$f(a,\,k) \stackrel{\text{def}}{=} \frac{\mathrm{d}\log\delta}{\mathrm{d}\log a} \approx \Omega_m^\gamma, \tag{4.52}$$

where the γ is the *growth index* and Ω_m the matter density we discussed in (2.37), a modification $Q(a,\,k)$ of Newton's constant in (2.51) and (2.90) respectively as

$$k^2 A_g \stackrel{\text{def}}{=} Q(a,\,k)\frac{a^2\delta\bar{\rho}}{M_g^2}, \tag{4.53}$$

and an effective anisotropic stress

$$\eta \stackrel{\text{def}}{=} -\frac{A_g}{E_g}, \tag{4.54}$$

resulting from the fact that A_g and E_g do not have to be equal to one another in (4.48) whereas $\Phi = \Psi$ by default in (2.85) in Newton's gauge in GR. As stated the full analysis is provided in [11], with the conclusion that the $\beta_1\beta_4$-model is the one whose predictions for these parameters mostly differ from those of GR and hence it is probably the first model which one may rule out.

Another method which has been suggested as a valid option in several cases for large k, is to solve the system generally to obtain two coupled second-order differential equations for the two remaining dynamical fields, in a set $\Xi = \{\Psi,\,\Phi\}$, and then substituting them with an ansatz of type $\Xi = \Xi_0 \exp{(i\omega N)}$ [12]. This is a type of *Wentzel–Kramers–Brillouin approximation (WKB)*, which is often used in quantum physics to locate approximate solutions to partial differential equations whose spatial coefficients vary. Here one assumes that ω changes slowly in time, i.e. obeys

$$\left|\frac{\dot{\omega}}{\omega^2}\right| \ll 1. \tag{4.55}$$

Such an approach enables one to express the (approximate) eigenfrequencies of the system as functions of the ferociously complicated general pre-factors, which are made up by background constituents, of the reduced system. This has been proposed as a sensible option to study instabilities of the linear perturbations. However, this are very new results and caution must be taken so that the approximation holds in the regimes which one would like to examine.

Chapter 5 Numerical methods and solution strategy

In this chapter we will discuss algebraic algorithms to solve the system of linear scalar perturbations in bimetric gravity and the numerics involved. Principally, we are interested in solving the system in such a way that we obtain two coupled second-order differential equations of two variables. This is a delicate problem since not all permitted gauge choices allow such a formulation in terms of the original system variables. Then, should one succeed in obtaining a solution, we specify a set of parameter values and provide a set of consistent initial conditions to start the numerical machinery to arrive at an expression for the dynamical evolution of the system. We will start this chapter with specifying which properties our solutions should have. To begin with, a physically viable solution should obey two criteria following [63],

(i) **Initial value variation:** A small variation of the initial values should produce a small variation in the solution over a fixed, compact region of spacetime. Mathematically, this connects to stability theory and the variation could be measured using various specific norms.

(ii) **Causal structure:** A change in the initial values in a specific region in spacetime should not yield change(s) in the solution beyond the causal future of this region.

Since the perturbations in this study appear on top of a homogeneous, isotropic background universe with decoupled wave modes, we are unable to investigate the second criteria, but the first will guid us towards stable solution.

5.1 First attempt: Differential-algebraic equations

By performing a Fourier transform and through decoupling of the different wave modes, the system of equations of motion has been reduced to a system of *differential-algebraic equations* (DAEs). These are ordinary differential equations which appear together with an algebraic constraint, more specifically a system such that

$$\begin{cases} \dot{q} & = f(q,\, t) \\ 0 & = g(q,\, t). \end{cases} \tag{5.1}$$

Constraints such as $0 = g(q,\, t)$, which can be expressed in terms of the generalized coordinates together with time involved are known as *holonomic constraints* [82]. If the constraint is expressed in terms of the velocities involved, it is usually a *nonholonomic constraint*, which cannot be expressed as $0 = g(q,\, t)$. Differential-algebraic problems appear frequently in numerous physical applications; in classical mechanics one, for instance, encounters

$$\begin{cases} m\ddot{x}(t) + \lambda(t)x(t) &= 0 \\ m\ddot{y}(t) + \lambda(t)y(t) &= -g \\ x(t)^2 + y(t)^2 &= L^2, \end{cases} \tag{5.2}$$

which is a mathematical description of the equations of motion for a pendulum with a point mass m at coordinates (x, y) and length L subjected to a gravitational acceleration g. The tension of the pendulum is encoded in $\lambda(t)$, which acts as a Lagrange multiplier in the equations.

An important property of DAEs is their *differentiate index*, which measures the distance to the corresponding ordinary differential equation, i.e. how many times the system must be differentiated in order to obtain a system of ODEs. In our case for the perturbations in bimetric theory we are left with a system with an algebraic constraint accompanied by several differential equations, where the highest derivatives are of order two. The system as a whole bears an index two in order to obtain a differential equation for each field.

The solution procedure when dealing with these types of systems is twofold; firstly a recasting of the problem through index reduction is necessary and secondly a consistent set of initial conditions must be provided by the user in order to calculate the trajectory. In `Mathematica` the routine `NDSolve` can handle some types of DAEs of index 1, but does not by default index reduce a system, which can been accessed through the option `"IndexReduction"`. This method works either through the graph-based Pantelides algorithm, `"Pantelides"`, or the slower Structural matrix method, `"StructuralMatrix"`, where the latter takes variable cancellations in the different equations into account [83]. After the reduction, it is possible introduce dummy derivatives to ensure that the system is not overdetermined or to project the solution onto a surface defined by the original constraints. However, performing an analysis of the gauge-invariant system in [7] using this framework proved to be a dead end, as Mathematica could not per see handle complexity of the nested equations without further user assistance, resulting in multiple kernel collapses as different initial values were tested. To be more precise, the task proved formidable to find initial values which satisfied all constraints of the system as these were not visible. Still, this approach may be more fruitful in the future as algorithms improve. Hence, a road forward could be deduced if one were able to recast the system into form of ordinary differential equations.

5.2 Second strategy: Redefining fields and eliminating auxiliary variables

To simplify the numerical problem solving, we reduce the system in order to obtain two coupled second-order differential equations, which are simpler to solve numerically provided that there are no singularities in the coefficients. In general relativity, this problem reduced to a single second order differential equation, (2.89) and (2.91), and hence we expect to find something similar given the shared properties of the theories. In order to solve the system, we can use the gauge freedom to set some of the redundant variables to zero. Of course, all information must still be contained in the system through the remaining variables and the gauge must be fixed completely. Principally, we will work with three options, which are listed below with key advantages and shortcomings:

(i) **Gauge invariant variables:** The equations of motion are solved using variable combinations which are gauge invariant, i.e. independent of gauge. The procedure to obtain such variables for general relativity following [23] was explained in section 2.2.2. However, the physical interpretations may not be transparent as results tend to be expressed for particular gauge choices.

(ii) **Conformal Newton gauge:** The choice $B_g = F_g = 0$, which is an ubiquitous option in ordinary relativistic cosmology. This allows us to make direct comparisons with predictions from general relativity without variable transformations and set initial conditions based on GR in the past, which fit cosmological observations.

(iii) **Longitudinal gauge:** Here, $F_f = F_g = 0$, which casts the $\delta f^{(\mathrm{s})}_{0-i}$-equation, (4.33), into a particularly easy form and also abbreviates (4.42) considerably.

It has turned out that there is by no means a unique way to achieve this sought coupled system. To complicate the matter further, several choices, among them the Newton gauge, have not yielded straightforward solutions of correct differential order. More specifically, for some choices the second order derivative of one field vanishes leaving one second order differential equation and one of first order. Fortunately, in the $F_g = F_f = 0$ gauge, one may obtain a second order system using conformal time following Könnig and Enander by redefining the fields as

$$\begin{aligned} \Phi_g &= A_g - H\dot{B}_g, \\ \Psi_g &= E_g - H\dot{B}_g - \ddot{B}_g, \end{aligned} \tag{5.3}$$

and in the f-sector,

$$\begin{aligned} \Phi_f &= A_f - \left(\frac{Y}{X}\right)^2 K\dot{B}_f, \\ \Psi_f &= E_f - \left(\frac{Y}{X}\right)^2 \left(\ddot{B}_f + 2K\dot{B}_f - \frac{\dot{X}}{X}\dot{B}_f\right), \\ B_f &\to B_g + \Delta B, \end{aligned} \tag{5.4}$$

which allows us to construct the combinations $\delta g^{(\mathrm{s})}_{0-0} + 3H\delta g^{(\mathrm{s})}_{0-i} + y^4(\delta f^{(\mathrm{s})}_{0-0} + 3K\delta f^{(\mathrm{s})}_{0-i})$ and the off-diagonal $\delta f^{(\mathrm{s})}_{i-j} + H/(Ky^4)g^{(\mathrm{s})}_{i-j}$ in equations (4.28), (4.29), (4.32) and (4.33) and (4.35) and (4.31), respectively which yield,

$$\frac{a^2\bar{\rho}}{M_g^2}(k^2\delta + 3H\theta) - k^4(y^4\Phi_f + \Phi_g) = 0, \tag{5.5}$$

$$\Phi_g + \Psi_g + \frac{K}{H}y^2(\Phi_f + \Psi_f) = 0, \tag{5.6}$$

where we can solve for Φ_f and Ψ_f in terms of $\{\Phi_g, \Psi_g, \delta, \theta\}$. The goal is to obtain two second-order differential equations involving the fields $\{\Phi_g, \Psi_g\}$. An equation for ΔB is acquired from $\delta g^{(\mathrm{s})}_{i-j}$ according to (4.31). To establish expressions for δ and θ we begin by solving for $\dot{\Phi}_f$ in $\delta f^{(\mathrm{s})}_{0-i}$, (4.33), and insert this into $\delta f^{(\mathrm{s})}_{0-0}$, (4.32), while substituting from the expressions for Ψ_f, Φ_f and ΔB. Then one simultaneously solves for δ and θ using this reduced $\delta f^{(\mathrm{s})}_{0-0}$-equation as well as the reduced $\delta g^{(\mathrm{s})}_{0-0}$-equation, (4.32) and (4.28) respectively.

In total, this leaves us with a system dependent on $\{\Phi_g,\, \Psi_g,\, B_g\}$ and their derivatives. To conclude, we seek to eliminate B_g. Fortunately for us, this is possible owing to the properties of the background, encoded in the Friedmann equations. We impose the constraint (4.17), the background relation (2.29) and solve for $\dot{y}$ and $\ddot{y}$ recursively using equation (4.22) and (4.23) where we at the same time replace all H-terms. Among the equations at hand in section 4.3, eight equations of motion, two stress-energy Bianchi constraints and two Einstein tensor + bimetric interaction term Bianchi constraints, with a possible maximum of eight linear independent equations, we can pick out three which we have not used so far, to solve the remainder of the system. In reality, one of them will constrain B_g and hence there are only two linear independent choices when the information is combined. A particularly nice choice is to pick the velocity equation $\delta f^{(\mathrm{s})}_{0-i}$, the trace $\delta g^{(\mathrm{s})}_{i-i}$ and the perturbed energy conservation equation in (4.33), (4.30) and (4.43) respectively. Inserting all previous information and field expressions into these three equations, one proceeds by Gauss elimination to produce solutions for $\{\ddot{\Phi}_g,\, \ddot{\Psi}_g,\, \dot{B}_g\}$. Remarkably, all non-differentiated B_g terms vanish and the $\dot{B}_g$ equation will take the following form

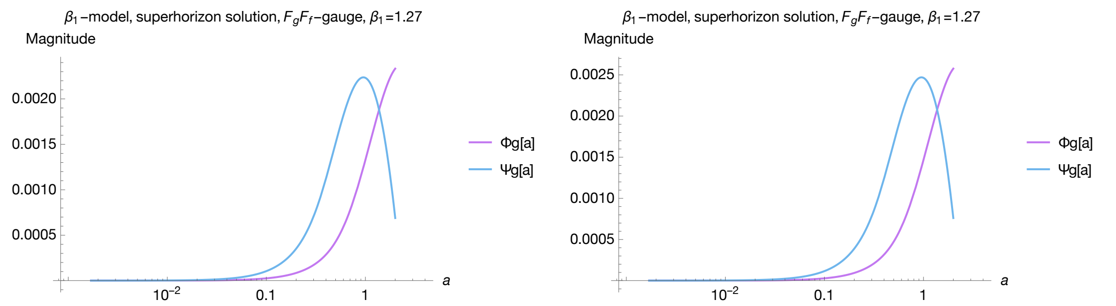
$$\dot{B}_g = \mathcal{N}\left(\dot{\Phi}_g,\, \dot{\Psi}_g,\, \Phi_g,\, \Psi_g,\, \text{background terms}\right), \tag{5.7}$$

which means that it can be differentiated to replace $\ddot{B}_g$ without introducing new derivates of higher order than $\{\ddot{\Phi}_g,\, \ddot{\Psi}_g\}$. Supplying this differentiated solution for $\ddot{B}_g$ into the solutions for $\{\ddot{\Phi}_g,\, \ddot{\Psi}_g\}$ gives two independent second-order differential equations for $\{\Phi_g,\, \Psi_g\}$. To analyze these, we insert sensible values for k or take the limit $k \to 0$ to obtain the superhorizon solution at the final stage and transform the normalized fields so that we may plot the quantities in terms of the scale factor $a(t)$ according to

$$\left(\frac{1}{H_0}\right)\dot{\Phi} \to a\mathcal{E}(a)\frac{\mathrm{d}}{\mathrm{d}a}\Phi, \tag{5.8}$$

with $\mathcal{E}(a)$ calculated from the background equations. A transformation rule for $\ddot{\Phi}$ is reached by applying (5.8) twice.

In order to determine the system we have to provide two initial values for the fields as well as two for their derivatives. We are interested in plotting the variables from the inflationary epoch, corresponding to roughly $a \sim 10^{-3}$, till today at $a = 1$. Since we have assumed that we are dealing with small perturbations, we must pick values for the fields which are small relative to the other parameters involved. Here, we will select values of the order 10^{-6} owing to our normalization of the other fields. Concerning physical solutions, which we discussed in the introduction of this chapter, we will vary these initial values of an order of 10 % and see how the results differ. To begin with, we study superhorizon solutions with the β_1-model we investigated in section 4.2 with $\beta_1 \approx 1.27$. The superhorizon solutions are plotted in figure 5.1 from $a = 0.0017$, which is approximately the earliest time `Mathematica` allowed due to the behavior of the pre-factors, assuming that the derivates at early times are negligible with a log a-axis. We vary the initial values slightly in figures 5.1(a) and 5.1(b), which affects the magnitudes of the fields slightly but not the shape of the curves. The potentials are frozen in at an early stage The ΛCDM solution with background data according to section 2.1 in Newton's gauge for equation (2.89) is plotted in figure 5.2, valid for all k-modes. Qualitatively, the shape for the curves do not differ but since the latter is valid for all k, caution must be taken concerning the magnitudes. Moreover, we are plotting in completely different gauges and a transformation must ensue to make a direct comparison. These are unresolved issues,



(a) $\Phi_g(a = 0.0017) = \Psi_g(a = 0.0017) = 10^{-6}$. **(b)** $\Phi_g(a = 0.0017) = \Psi_g(a = 0.0017) = 1.1 \cdot 10^{-6}$.

Figure 5.1: Superhorizon solutions for the minimal bimetric β_1-model in $F_g = F_f = 0$, both with $\dot{\Phi}_g(a = 0.0017) = \dot{\Psi}_g(a = 0.0017) = 0$ where the dot signifies a derivative with respect to a, but with different initial values for the fields.

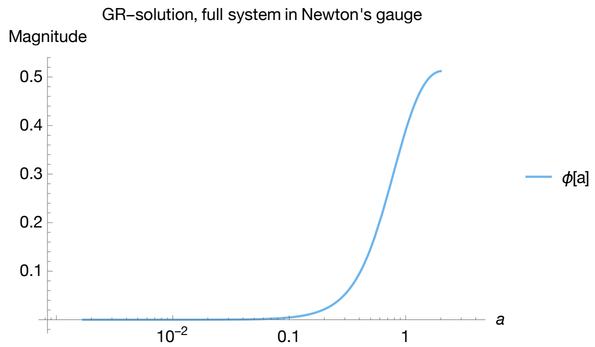


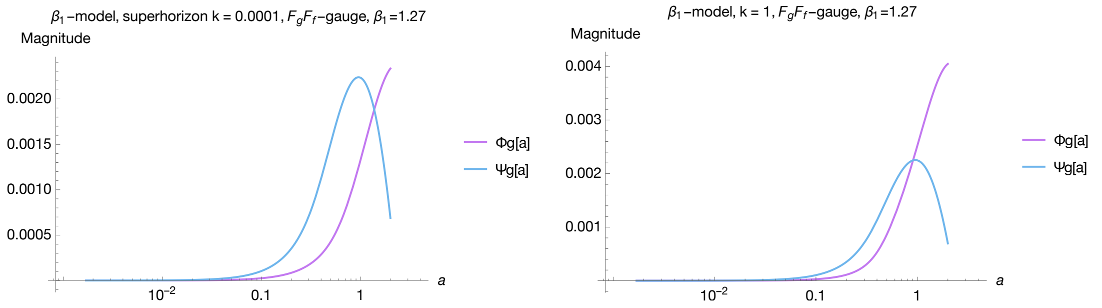
Figure 5.2: Solution for ΛCDM for pressure-less dust, following (2.89), with the same initial values as in 5.1(a).

which will be addressed in upcoming works. We have also checked to place in a small k-value, $k = 0.0001$ instead of taking the limit $k \to 0$ as well as a larger k-value $k = 1$. The results are found in figures 5.3(a) and 5.3(b), where one draws the conclusion that the approximation $k = 0$ is indeed value for the superhorizon case and that results starts to differ for larger k. As for the infinite-branch $\beta_1\beta_4$-model, the system has been solved using the consistent values in section 4.2 but the results have yet to be graphically illustrated.

5.2.1 Noether identities

A different approach[1], suggested by Solomon, suitable for modified gravity, to obtain gauge invariant combinations is to analyze the perturbed action with help from *Noether identities* [84]. These relate the equations of motion to one another and are direct consequences of the gauge invariance of the action. Firstly, one identifies the simultaneous transformation, i.e. gauge variation, of the perturbed variables which leaves the action invariant. In bimetric gravity, one proceeds in the same way as in general relativity as we did in section 2.2.2, since the bimetric action is invariant under simultaneous diffeomorphisms of g and f following (3.58) as we discussed in section 3.5. Under a coordinate transformation $x^\mu \to x^\mu + \xi^\mu(x)$,

[1]This method was partly used in [12] with a succinct description.



(a) $k = 0.0001$.

(b) $k = 1$.

Figure 5.3: Superhorizon solutions for the minimal bimetric β_1-model in $F_g = F_f = 0$, both with $\dot{\Phi}_g(a = 0.0017) = \dot{\Psi}_g(a = 0.0017) = 0$, $\Phi_g(a = 0.0017) = \Psi_g(a = 0.0017) = 10^{-5}$ but for different wave modes k.

where $\xi^\mu(x)$ are arbitrary first-order scalar functions as in section 2.2.2, the perturbed part of each metric transforms as in (2.72), namely

$$\delta g_{\mu\nu} \to \delta g_{\mu\nu} + \mathscr{L}_\xi \bar{g}_{\mu\nu}, \;\; \delta f_{\mu\nu} \to \delta f_{\mu\nu} + \mathscr{L}_\xi \bar{f}_{\mu\nu}. \tag{5.9}$$

Reminding ourselves that the scalar parts are obtained from the components ξ^S through the splitting ξ^i following the notation in [22], we arrive at

$$\Delta E_g = \frac{2}{N}\partial_t\left(\frac{\xi_0}{N}\right), \; \Delta A_g = \frac{2H}{N^2}\xi_0, \; \Delta B_g = -\frac{2}{a^2}\xi^S, \; \Delta F_g = -\frac{1}{Na}\left(\xi_0 + \dot{\xi}^S - 2H\xi^S\right), \tag{5.10}$$

$$\Delta E_f = \frac{2}{X}\partial_t\left(\frac{\xi_0}{X}\right), \; \Delta A_f = 2K\xi_0, \; \Delta B_f = -\frac{2}{X^2}\xi^S, \; \Delta F_f = -\frac{1}{XY}\left(\xi_0 + \dot{\xi}^S - 2K\xi^S\right), \tag{5.11}$$

where we have lowered the indices with the unperturbed metric in each sector, echoing (2.79). For the matter sector, one simply subjects the stress-energy tensor to the same treatment, and $\mathscr{L}_\xi \bar{T}_{\mu\nu}$ yields,

$$\Delta\delta = N\partial_t\left(\frac{\bar{\rho}}{N}\right)\xi_0, \;\; \Delta\theta = -\frac{1}{a^2}\partial_i\xi_0. \tag{5.12}$$

Here we have calculated the perturbations with lowered indices $\delta T_{\mu\nu}$ as in [22], i.e.

$$\delta T^\mu{}_\nu = \bar{g}^{\mu\lambda}\left[\delta T_{\lambda\nu} - \delta g_{\lambda\sigma}\bar{T}^\sigma{}_\lambda\right]. \tag{5.13}$$

Then, one considers an arbitrary variation of the action, where the components in the Lagrangian can be written as the variation of each field multiplied by the field's equation of motion. At this point, one replaces the variation of the field with its gauge variation which must render $\delta S = 0$ by default, i.e.

$$\delta S = \int \mathrm{d}^4x \left[\sum_i \mathcal{E}_{\phi_i}\Delta\phi_i\right] = 0, \tag{5.14}$$

where $\mathcal{E}_{\phi_i}$ is the equation of motion for each perturbation field ϕ_i and $\Delta\phi_i$ the variation presented in (5.10), (5.11) and (5.12). Through integration by parts one arrives at coefficients involving the different equations of motion multiplied by either ξ_0 or ξ^S. As stated these functions are arbitrary and hence their coefficients must equal zero for $\delta S = 0$ to hold. The relations produced are the Noether identities. If a field's equation of motion can be deduced from a combination of the other equations of motion involved in these identities, it is superfluous. This makes it easy to safely set some of these fields to zero with one option per Noether identity. In our case, the Noether identity for ξ_0 enables us to set one of the variables $\{A_g,\, A_f,\, \delta\}$ to zero and the one for ξ^S lets us choose one of $\{B_g,\, B_f\}$.[2] Moreover, by integrating the θ-variation spatially, we obtain an additional field to place in the identity.

To continue, the calculation can be further facilitated by splitting the components into *dynamical* fields and *auxiliary* fields [84]. In order to do so, we study the second-order action formed by inserting the first-order perturbed metrics and Taylor expanding up to second order in these perturbation variables. Naïvely, one would assume that one also has to take purely second-order perturbations into account, but since such terms only appear in the equations of motion multiplied by background variables, they vanish as the background equations are satisfied. The fields which have kinetic terms in the second order action are the dynamical ones, while the rest are auxiliary, having algebraic equations of motion. However, we must first fix the gauge before starting to eliminate the auxiliary fields, as some of the terms in the second-order action are relics (not physical degrees of freedom) which violate the requirement that the action is solely invariant under infinitesimal transformations of first order. To incorporate this fact, we can settle on a gauge based on the Noether identities and then recast the surviving auxiliary fields as combinations of the remaining dynamical fields.

In practice, for our bimetric setup we insert the perturbed metrics into the dynamical terms $\sqrt{-g}R(g)$ and $\sqrt{-f}R(f)$ in the action, Taylor expand and count the fields with kinetic terms. At first glance, this task might seem daunting but fortunately, we can use the first-order equations of motion to put some of the terms to zero or express them in terms of δ and θ. This leaves us with $\{A_g,\, A_f,\, B_g,\, B_f,\, \delta\}$ as dynamical and the residual fields are auxiliary.

Instead of setting a field in each set to zero, we can express them in terms of one another. A fitting choice which provides a solution to the system is to set $A_f = 0$ and

$$\delta = \frac{1}{2}\left(-3A_g + k^2 B_g\right), \tag{5.15}$$

which gives $\theta(t) = -k^2 F_g$ in the energy conservation equation (4.43). To proceed one solves for $\dot{A}_g$ in $\delta g^{(\mathrm{s})}_{0-i}$, (4.29), while implementing the θ solution and inserts the result into $\delta g^{(\mathrm{s})}_{0-0}$, (4.28). Then one extracts the auxiliary field E_f from $\delta f^{(\mathrm{s})}_{0-i}$, (4.33), and places the expression into $\delta f^{(\mathrm{s})}_{0-0}$, (4.32). With these two density equations we are able to obtain a simultaneous result for F_f and F_g, which can be implanted into the velocity equations, $\delta g^{(\mathrm{s})}_{0-i}$, (4.29), and $\delta f^{(\mathrm{s})}_{0-i}$, (4.33), to produce expressions for E_g and E_f. The remaining variables are the dynamical fields $\{A_g, B_g, B_f\}$. This time, it will be A_g that eventually drops out upon inspection of the last equations. To achieve this, we replace all auxiliary fields and background contributions in $\delta g^{(\mathrm{s})}_{i-i}$, $\delta g^{(\mathrm{s})}_{i-j}$, $\delta f^{(\mathrm{s})}_{i-i}$ and $\delta f^{(\mathrm{s})}_{i-f}$, (4.30), (4.31), (4.34) and (4.35), which will yield two second-order differential equations for $\{B_g, B_f\}$, the terms involving $\ddot{A}_g$ vanish, and a constraint equation for $A_g = \mathcal{N}(\dot{B}_g,\, \dot{B}_f,\, B_g,\, B_f, \text{background terms})$ is acquired. By

[2] One can easily recognize this by noting that the variations of these fields do not come with any derivatives of ξ_0 or ξ^S. Hence, there will be no differentiated terms of their equations of motion in the final identities and their equations of motion can be written in terms of the ones of the remaining fields.

differentiating this constraint we are able to remove the $\dot{A}_g$-terms and are left with two coupled equations depending on the fields $\{B_g, B_f\}$ and their derivatives. To demonstrate this in `Mathematica`, we have solved the system using this procedure for the case of the $\beta_1\beta_4$-model. However, the final pre-factors have been difficult to analyze due to their intrinsic complexity; it is not certain whether specific limits can be taken but further, more extensive studies will hopefully enlighten us in this task. The goal is to compare this solution with the one calculated in the $F_g = F_f = 0$-gauge where all expressions have been transformed to gauge invariant variables to serve as a crosscheck. Moreover, presumably even more efficient gauge choices for the solution strategy will be proposed in the near future.

Chapter 6 Outlook

This thesis has served as an introduction to relativistic perturbation theory as well as bimetric cosmology. What remains to be explored, which unfortunately did not make it into this thesis due to the limited schedule, is calculating the contributions for the ISW effect in bimetric gravity, which will originate both from the superhorizon translation provided by the comoving curvature perturbation, which we do not know whether it is conserved or not in this case, as well as the integral from recombination till today of the differentiated bimetric potentials interpreted in Newton's gauge carried out in conformal time. Of course, comparisons with experimental data must be performed using the setup established in section 2.3.1.

Equally intriguing is the comparison with the analytic Newtonian expressions and the further simplified WKB approximation for the general case. Upcoming experiments, such as `Euclid` will serve to put more narrow bounds on the parameters, testing the viability of the theory as well as of ΛCDM. Ideally, black holes, gravitational lensing and other probes could be combined to select the best parameters. The recently introduced ghost-free double-coupling with its associated cosmology should prove interesting in this context, providing valuable insights.

As pointed out, regarding instabilities, much remains to be resolved and one could imagine nonlinear contributions originating from the Vainshtein mechanism, which we briefly mentioned in section 3.2, to have significant impact on the interpretation, rendering previous discarded theories valid. Hence, it will take some time and careful reasoning to establish a set of guidelines. Of the more theoretical aspects, the location of a functioning degravitation mechanism would be most welcome. Naturally, a solid interpretation of the extra metric $f_{\mu\nu}$ as well as $g_{\mu\nu}$ in these models has to be provided, especially from the spacetime geometry point of view. Consequences connected to the weak and strong equivalence principles should also be investigated scrupulously.

The potential experimental biases from underlying assumptions based on ΛCDM will have to be settled statistically more throughly in several independent analyses. It could also be of great value, if one could modify existing numerical general relativity packages to accommodate bimetric gravity, which would provide augmented results with respect to `Mathematica`. Future comparisons with other promising theories should bring additional transparency when it comes to observables and help to establish a permanent foothold in the vast landscape of modified theories of gravity.

References

[1] **Supernova Search Team** Collaboration, A. G. Riess et al., *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, *Astron.J.* **116** (1998) 1009–1038, `arXiv:astro-ph/9805201 [astro-ph]`.

[2] **Supernova Cosmology Project** Collaboration, S. Perlmutter et al., *Measurements of Omega and Lambda from 42 high redshift supernovae*, *Astrophys.J.* **517** (1999) 565–586, `arXiv:astro-ph/9812133 [astro-ph]`.

[3] S. 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]`.

[4] S. 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]`.

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

[6] M. von Strauss, A. Schmidt-May, J. Enander, E. Mörtsell, and S. Hassan, *Cosmological Solutions in Bimetric Gravity and their Observational Tests*, *JCAP* **1203** (2012) 042, `arXiv:1111.1655 [gr-qc]`.

[7] M. Berg, I. Buchberger, J. Enander, E. Mörtsell, and S. Sjörs, *Growth Histories in Bimetric Massive Gravity*, *JCAP* **1212** (2012) 021, `arXiv:1206.3496 [gr-qc]`.

[8] K. Hinterbichler, *Theoretical Aspects of Massive Gravity*, *Rev.Mod.Phys.* **84** (2012) 671–710, `arXiv:1105.3735 [hep-th]`.

[9] C. de Rham, *Massive Gravity*, `arXiv:1401.4173 [hep-th]`.

[10] A. M. Schmidt-May, *Classically Consistent Theories of Interacting Spin-2 Fields*. PhD thesis, Stockholm University, Department of Physics, 2013.

[11] A. R. Solomon, Y. Akrami, and T. S. Koivisto, *Cosmological perturbations in massive bigravity: I. Linear growth of structures*, `arXiv:1404.4061 [astro-ph.CO]`.

[12] 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]`.

[13] M. S. Volkov, *Hairy black holes in the ghost-free bigravity theory*, *Phys.Rev.* **D85** (2012) 124043, `arXiv:1202.6682 [hep-th]`.

[14] R. Brito, V. Cardoso, and P. Pani, *Black holes with massive graviton hair*, *Phys.Rev.* **D88** (2013) 064006, `arXiv:1309.0818 [gr-qc]`.

[15] J. Enander and E. Mörtsell, *Strong lensing constraints on bimetric massive gravity*, *JHEP* **1310** (2013) 031, `arXiv:1306.1086 [astro-ph.CO]`.

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

[17] S. Hassan, A. Schmidt-May, and M. von Strauss, *Metric Formulation of Ghost-Free Multivielbein Theory*, arXiv:1204.5202 [hep-th].

[18] C. Deffayet, J. Mourad, and G. Zahariade, *A note on 'symmetric' vielbeins in bimetric, massive, perturbative and non perturbative gravities*, *JHEP* **1303** (2013) 086, arXiv:1208.4493 [gr-qc].

[19] A. Joyce, B. Jain, J. Khoury, and M. Trodden, *Beyond the Cosmological Standard Model*, arXiv:1407.0059 [astro-ph.CO].

[20] A. Einstein, *Die grundlage der allgemeinen relativitätstheorie*, *Annalen der Physik* **354** (1916), no. 7 769–822.

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

[22] S. Weinberg, *Cosmology.* Oxford Univ. Press, Oxford, 2008.

[23] V. F. Mukhanov, H. Feldman, and R. H. Brandenberger, *Theory of cosmological perturbations. Part 1. Classical perturbations. Part 2. Quantum theory of perturbations. Part 3. Extensions*, *Phys.Rept.* **215** (1992) 203–333.

[24] T. Padmanabhan, *Gravitation: Foundations and Frontiers.* Cambridge University Press, Cambridge, 2010.

[25] D. Baumann, "Cosmology: Part III Mathematical Tripos." University lecture notes, 2014. http://www.damtp.cam.ac.uk/user/db275/Cosmology/Lectures.pdf.

[26] H. Kurki-Suonio, "Cosmology i & ii." University lecture notes, 2005. http://www.helsinki.fi/~hkurkisu/cosmology/Cosmo0.pdf.

[27] C. Hirata, "Physics 217bc: The Standard Model - Cosmology." University lecture notes, 2008. http://www.tapir.caltech.edu/~chirata/ph217/.

[28] S. Dodelson, *Modern cosmology.* Academic Press, San Diego, CA, 2003.

[29] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time.* Cambridge monographs on mathematical physics. Cambridge Univ. Press, Cambridge, 1973.

[30] S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity.* Wiley, New York, NY, 1972.

[31] B. A. Bassett and R. Hlozek, *Baryon Acoustic Oscillations*, arXiv:0910.5224 [astro-ph.CO].

[32] **SDSS** Collaboration, D. J. Eisenstein et al., *Detection of the baryon acoustic peak in the large-scale correlation function of SDSS luminous red galaxies*, *Astrophys.J.* **633** (2005) 560–574, arXiv:astro-ph/0501171 [astro-ph].

[33] L. Amendola, M. Kunz, M. Motta, I. D. Saltas, and I. Sawicki, *Observables and unobservables in dark energy cosmologies*, *Phys.Rev.* **D87** (2013) 023501, arXiv:1210.0439 [astro-ph.CO].

[34] **EUCLID** Collaboration, R. Laureijs et al., *Euclid Definition Study Report*, arXiv:1110.3193 [astro-ph.CO].

[35] **Planck** Collaboration, P. Ade et al., *Planck 2013 results. I. Overview of products and scientific results*, arXiv:1303.5062 [astro-ph.CO].

[36] R. Sachs and A. Wolfe, *Perturbations of a cosmological model and angular variations of the microwave background*, *Astrophys.J.* **147** (1967) 73–90.

[37] **Planck** Collaboration, P. Ade et al., *Planck 2013 results. XIX. The integrated Sachs-Wolfe effect*, arXiv:1303.5079 [astro-ph.CO].

[38] W. Hu, *Dark synergy: Gravitational lensing and the CMB*, *Phys.Rev.* **D65** (2002) 023003, arXiv:astro-ph/0108090 [astro-ph].

[39] P. J. Mohr, B. N. Taylor, and D. B. Newell, *CODATA Recommended Values of the Fundamental Physical Constants: 2010*, *Rev.Mod.Phys.* **84** (2012) 1527–1605, `arXiv:1203.5425 [physics.atom-ph]`.

[40] S. Weinberg, *The Quantum Theory of Fields: Supersymmetry.* The Quantum Theory of Fields. Cambridge University Press, Cambridge, 2000.

[41] S. Weinberg, *The Cosmological Constant Problem*, *Rev. Mod. Phys.* **61** (1989) 1–23.

[42] G. 't Hooft, *Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking*, *NATO Adv.Study Inst.Ser.B Phys.* **59** (1980) 135.

[43] S. Dimopoulos and L. Susskind, *Mass Without Scalars*, *Nucl.Phys.* **B155** (1979) 237–252.

[44] C. de Rham, G. Gabadadze, L. Heisenberg, and D. Pirtskhalava, *Nonrenormalization and naturalness in a class of scalar-tensor theories*, *Phys.Rev.* **D87** (2013), no. 8 085017, `arXiv:1212.4128 [hep-th]`.

[45] C. de Rham, L. Heisenberg, and R. H. Ribeiro, *Quantum Corrections in Massive Gravity*, *Phys.Rev.* **D88** (2013) 084058, `arXiv:1307.7169 [hep-th]`.

[46] N. Arkani-Hamed, S. Dimopoulos, G. Dvali, and G. Gabadadze, *Nonlocal modification of gravity and the cosmological constant problem*, `arXiv:hep-th/0209227 [hep-th]`.

[47] G. Dvali, S. Hofmann, and J. Khoury, *Degravitation of the cosmological constant and graviton width*, *Phys.Rev.* **D76** (2007) 084006, `arXiv:hep-th/0703027 [hep-th]`.

[48] M. Fierz and W. Pauli, *On relativistic wave equations for particles of arbitrary spin in an electromagnetic field*, *Proc. Roy. Soc. Lond.* **A173** (1939) 211–232.

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

[50] S. Weinberg, *The Quantum Theory of Fields: Foundations.* The Quantum Theory of Fields. Cambridge University Press, Cambridge, 1995.

[51] H. van Dam and M. Veltman, *Massive and massless Yang-Mills and gravitational fields*, *Nucl.Phys.* **B22** (1970) 397–411.

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

[53] A. Vainshtein, *To the problem of nonvanishing gravitation mass*, *Phys.Lett.* **B39** (1972) 393–394.

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

[55] D. Boulware and S. Deser, *Inconsistency of finite range gravitation*, *Phys.Lett.* **B40** (1972) 227–229.

[56] D. Boulware and S. Deser, *Can gravitation have a finite range?*, *Phys.Rev.* **D6** (1972) 3368–3382.

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

[58] N. Arkani-Hamed, H. Georgi, and M. D. Schwartz, *Effective field theory for massive gravitons and gravity in theory space*, *Annals Phys.* **305** (2003) 96–118, `arXiv:hep-th/0210184 [hep-th]`.

[59] C. de Rham and G. Gabadadze, *Generalization of the Fierz-Pauli Action*, *Phys. Rev.* **D82** (2010) 044020, `arXiv:1007.0443 [hep-th]`.

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

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

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

[63] R. M. Wald, *General relativity.* Chicago Univ. Press, Chicago, IL, 1984.

[64] R. L. Arnowitt, S. Deser, and C. W. Misner, *The Dynamics of General Relativity*, *Gen.Rel.Grav.* **40** (2008) 1997–2027, arXiv:gr-qc/0405109 [gr-qc].

[65] G. Date, *Lectures on Constrained Systems*, arXiv:1010.2062 [gr-qc].

[66] J. Khoury, G. E. Miller, and A. J. Tolley, *Spatially Covariant Theories of a Transverse, Traceless Graviton, Part I: Formalism*, *Phys.Rev.* **D85** (2012) 084002, arXiv:1108.1397 [hep-th].

[67] S. Hassan, A. Schmidt-May, and M. von Strauss, *Proof of Consistency of Nonlinear Massive Gravity in the Stückelberg Formulation*, *Phys.Lett.* **B715** (2012) 335–339, arXiv:1203.5283 [hep-th].

[68] S. 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].

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

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

[71] 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].

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

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

[74] J. Enander, A. R. Solomon, Y. Akrami, and E. Mörtsell, *Cosmic expansion histories in doubly-coupled, ghost-free massive bigravity*, arXiv:1409.2860 [gr-qc].

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

[76] J. Enander, *Cosmological tests of massive bigravity.* Licentiate thesis, Stockholm University, Department of Physics, 2012.

[77] Y. Akrami, T. S. Koivisto, and M. Sandstad, *Cosmological constraints on ghost-free bigravity: background dynamics and late-time acceleration*, arXiv:1302.5268 [astro-ph.CO].

[78] F. Könnig and L. Amendola, *Instability in a minimal bimetric gravity model*, *Phys.Rev.* **D90** (2014) 044030, arXiv:1402.1988 [astro-ph.CO].

[79] 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].

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

[81] **Euclid Theory Working Group** Collaboration, L. Amendola et al., *Cosmology and fundamental physics with the Euclid satellite*, *Living Rev.Rel.* **16** (2013) 6, `arXiv:1206.1225 [astro-ph.CO]`.

[82] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*. Pearson Education, Limited, 2013.

[83] M. Sofroniou and R. Knapp, *Advanced Numerical Differential Equation Solving in Mathematica*, tech. rep., Wolfram Research, Inc, 2008.

[84] M. Lagos, M. Bañados, P. G. Ferreira, and S. García-Sáenz, *Noether Identities and Gauge-Fixing the Action for Cosmological Perturbations*, *Phys.Rev.* **D89** (2014), no. 2 024034, `arXiv:1311.3828 [gr-qc]`.

[85] P. Borwein and T. Erdelyi, *Polynomials and Polynomial Inequalities*. Springer New York, New York, 1995.

[86] C. de Rham, G. Gabadadze, and A. J. Tolley, *Ghost free Massive Gravity in the Stückelberg language*, *Phys.Lett.* **B711** (2012) 190–195, `arXiv:1107.3820 [hep-th]`.

[87] S. Alexandrov, *Canonical structure of Tetrad Bimetric Gravity*, *Gen.Rel.Grav.* **46** (2014) 1639, `arXiv:1308.6586 [hep-th]`.

Chapter A Elementary symmetric polynomials

This appendix elaborates on the polynomials $e_n(S)$ in the potential term of bimetric theory, namely

$$2m^d \sum_{n=0}^{d} \beta_n e_n(S), \text{ where } S = \sqrt{g^{-1}f}. \tag{A.1}$$

Given d variables, an *elementary symmetric polynomial* of degree n is formed by adding together all distinct products of n variables, $n \leqslant d$, i.e. a polynomial which is symmetric under any permutation of its variables. In an example with variables x and $d = 3$, the elementary symmetric polynomials are

$$\begin{aligned}
e_0\left(x_1,\, x_2,\, x_3\right) &= 1, \\
e_1\left(x_1,\, x_2,\, x_3\right) &= x_1 + x_2 + x_3, \\
e_2\left(x_1,\, x_2,\, x_3\right) &= x_1x_2 + x_2x_3 + x_1x_3, \\
e_3\left(x_1,\, x_2,\, x_3\right) &= x_1x_2x_3.
\end{aligned} \tag{A.2}$$

The elementary symmetric polynomials of a $d \times d$ matrix S are defined through,

$$e_n(S) \overset{\text{def}}{=} \frac{1}{d!}\binom{d}{n} \epsilon^{\mu_1 \ldots \mu_n \lambda_{n+1} \ldots \lambda_d} \epsilon_{\mu_1 \ldots \mu_n \lambda_{n+1} \ldots \lambda_d} S^{\nu_1}{}_{\mu_1} \, \ldots \, S^{\nu_n}{}_{\mu_n}, \tag{A.3}$$

with $e_0(S) = 1$ and where $n \leqslant d$, from which we readily read off

$$e_d(S) = \frac{1}{d!} \epsilon^{\mu_1 \ldots \mu_d} \epsilon_{\mu_1 \ldots \mu_d} S^{\nu_1}{}_{\mu_1} \, \ldots \, S^{\nu_d}{}_{\mu_d} = \det S, \tag{A.4}$$

and $e_1(S) = \operatorname{Tr} S$. In other words, the elementary symmetric polynomials constitute a "deformed determinant" with the d variables as the row/column entries of S where the Levi-Civita tensors keep track of the permutations of distinct products of the chosen n row/column entries. The elementary symmetric polynomials can be related to one another via the recursive formula

$$e_n(S) = \frac{(-1)^{n+1}}{n} \sum_{k=0}^{n-1} (-1)^k \operatorname{Tr}\left(S^{n-k}\right) e_k(S), \tag{A.5}$$

derived from Newton's identities [85], where its corresponding relation usually is presented as

$$e_n(S) = \frac{1}{n} \sum_{k=1}^{n} (-1)^{k-1} e_{k-n}(S) p_k(S), \tag{A.6}$$

where $p_k(S)$ is the k:th power sum of the eigenvalues λ_i of S labeled by $i = 1, \ldots, d$, i.e.

$$p_k(S) = \sum_{i=1}^{d} \lambda_i^k. \tag{A.7}$$

Clearly $p_1(S) = \operatorname{Tr} S$ and the subsequent higher powers can be written as $p_k(S) = \operatorname{Tr}(S^k)$, which is a rudimentary result in linear algebra. One then arrives at (A.5) by redefining the summation variable.

It is possible to write the expansion of the determinant in terms of elementary symmetric polynomials as

$$\det(\mathbb{1} + S) = \sum_{n=0}^{d} e_n(S), \tag{A.8}$$

which can be deduced by replacing S with $\mathbb{1}+S$ in (A.4) in all terms but the polynomial one. An interesting symmetry relation also arises directly from the definition,

$$e_n(S) = \det S \cdot e_{d-n}\left(S^{-1}\right). \tag{A.9}$$

When varying the action, we obtain the related matrix terms $Y^{(n)}$. They can be expressed in terms of the elementary symmetric polynomials through

$$\left(Y^{(n)}\right)^{\mu}{}_{\nu} = \sum_{k=0}^{n}(-1)^k e_k(S)\left(S^{n-k}\right)^{\mu}{}_{\nu}. \tag{A.10}$$

Writing out the terms in matrix form for $n \leqslant 3$ which appear in our expressions, one arrives at

$$\begin{aligned}
Y^{(0)}(S) &= \mathbb{1}\\
Y^{(1)}(S) &= S - [S]\mathbb{1}\\
Y^{(2)}(S) &= S^2 - [S]S + \frac{1}{2}\left([S]^2 - [S^2]\right)\mathbb{1}\\
Y^{(3)}(S) &= \frac{1}{6}\left(6S^3 + 3S[S]^2 - [S]^3\mathbb{1} - 3S[S^2] + [S](-6S^2 + 3[S^2]\mathbb{1}) - 2[S^3]\mathbb{1}\right)
\end{aligned} \tag{A.11}$$

Chapter B Tensor perturbations in bimetric gravity

As in general relativity, the linear tensor perturbations are the trace- and divergenceless constituents. These requirements could be accommodated through an ansatz,

$$\mathrm{d}s^2_{\bar{g}+\delta g^{(\mathrm{t})}} = -N^2\,\mathrm{d}t^2 + a^2\left[\mathrm{d}\boldsymbol{x}^2 + 2h^{(g)}_{xy}(t,\,z)\,\mathrm{d}x\,\mathrm{d}y + h^{(g)}_{xx}(t,\,z)\left(\mathrm{d}x^2 - \,\mathrm{d}y^2\right)\right], \tag{B.1}$$

$$\mathrm{d}s^2_{\bar{f}+\delta f^{(\mathrm{t})}} = -X^2\,\mathrm{d}t^2 + Y^2\left[\mathrm{d}\boldsymbol{x}^2 + 2h^{(f)}_{xy}(t,\,z)\,\mathrm{d}x\,\mathrm{d}y + h^{(f)}_{xx}(t,\,z)\left(\mathrm{d}x^2 - \,\mathrm{d}y^2\right)\right], \tag{B.2}$$

for gravitational waves propagating in the z-direction, first presented in [7] with a de Sitter background, where the tensors are symmetric under the exchange of spatial indices. Note that the relation detailed in (4.17) still applies since the unperturbed metric is the same. The tensor perturbations are pure physical propagating modes and suffer hence from no gauge redundancy. We will restrict ourselves to conformal time as the equations become slightly more transparent. Moreover, we will express the result in Fourier modes while dropping the phases. Since no time components are present for the perturbations, the $0-0$ and $0-i$ equations are trivial, as the $4-4$ due to that no part of the perturbations travels parallel to the axis of propagation. After some calculations we arrive at the following,

- $\delta g^{(\mathrm{t})}_{i-i}$:

$$\pm\frac{1}{2}\left[\frac{1}{a^2}\left(k^2 h^{(g)}_{xx} + 2H\dot{h}^{(g)}_{xx} + \ddot{h}^{(g)}_{xx}\right) + m^2 Q y\left(h^{(g)}_{xx} - h^{(f)}_{xx}\right)\right] = 0, \tag{B.3}$$

- $\delta g^{(\mathrm{t})}_{i-j}$:

$$\frac{1}{2a^2}\left[\frac{1}{a^2}\left(\left(k^2 - 2\dot{H}\right)h^{(g)}_{xy} - 2H\dot{h}^{(g)}_{xy} + \ddot{h}^{(g)}_{xy}\right) + \frac{m^2 Q}{y}\left(y^2 h^{(g)}_{xy} - h^{(f)}_{xy}\right)\right] = 0, \tag{B.4}$$

- $\delta f^{(\mathrm{t})}_{i-i}$:

$$\pm\frac{1}{2}\left[\frac{1}{a^2}\left(\frac{k^2}{y^2}h^{(f)}_{xx} + \frac{1}{x^2}\left(\left\{2H - \frac{\dot{x}}{x} + 3\frac{\dot{y}}{y}\right\}\dot{h}^{(f)}_{xx} + \ddot{h}^{(f)}_{xx}\right)\right) - \frac{m^2 Q}{xy^2}\left(h^{(g)}_{xx} - h^{(f)}_{xx}\right)\right] = 0, \tag{B.5}$$

- $\delta f^{(\mathrm{t})}_{i-j}$:

$$\frac{1}{2a^4}\left[\frac{1}{x^2y^2}\left(\ddot{h}^{(f)}_{xy} - \left\{2H + \frac{\dot{x}}{x} + \frac{\dot{y}}{y}\right\}\dot{h}^{(f)}_{xy}\right) + \left(\frac{k^2}{y^4} + \frac{2}{x^2y^2}\left\{\frac{\dot{x}}{x}H + \frac{\dot{x}}{x}\frac{\dot{y}}{y} - \dot{H} - \frac{\dot{y}}{y}H - \frac{\ddot{y}}{y}\right\}\right)h^{(f)}_{xy}\right] - \frac{m^2 Q}{2a^2xy^4}\left(y^2 h^{(g)}_{xy} - h^{(f)}_{xy}\right) = 0. \tag{B.6}$$

In the de Sitter case, $x = y = \text{constant} = \alpha$ and it easy to form linear combinations of the two sectors to obtain for equations of motion for the two propagating massive modes and the two massless modes respectively as,

$$\begin{aligned} h^{(+)}_{xx} &\stackrel{\text{def}}{=} h^{(g)}_{xx} + \alpha^2 h^{(f)}_{xx}, \quad h^{(-)}_{xx} \stackrel{\text{def}}{=} h^{(g)}_{xx} - h^{(f)}_{xx}, \\ h^{(+)}_{xy} &\stackrel{\text{def}}{=} h^{(g)}_{xy} + \alpha^2 h^{(f)}_{xy}, \quad h^{(-)}_{xy} \stackrel{\text{def}}{=} h^{(g)}_{xy} - h^{(f)}_{xy}. \end{aligned} \tag{B.7}$$

Future analyses related to superhorizon effects should also focus on these in order to calculate additional contributions, such as the prevalence of E-modes and B-modes in bimetric gravity, measuring the magnitude of primordial gravitational waves.

Chapter C A redefined shift

In this appendix we present the steps given in [4] to redefine the shift variable, relation (3.36) in section 3.4, in a consistent manner in massive gravity. We will arrive at the same modified variables given in [10] and provide some comments based on the review [9].

Following our brief discussion in section 3.4, we seek to match the terms matrix by matrix in (3.37) and (3.38) to obtain the proper redefinition. In terms of the ADM parametrizations of g^{-1} and f in (3.19) and (3.31) respectively, (3.37) turns into

$$N^2 g^{-1} f = \begin{pmatrix} \tilde{N}^2 - \tilde{N}_l \tilde{h}^{lk} \tilde{N}_k + N^l \tilde{N}_l & -\tilde{N}_j + N^l \tilde{h}_{lj} \\ N^2 h^{il} \tilde{N}_l - N^i \left(\tilde{N}^2 - \tilde{N}_l \tilde{h}^{lk} \tilde{N}_k + N^l \tilde{N}_l \right) & N^2 h^{il} \tilde{h}_{lj} - N^i \left(-\tilde{N}_j + N^l \tilde{h}_{lj} \right) \end{pmatrix}, \tag{C.1}$$

where the contents in the parentheses on the second row matches the terms on the row above. If we decompose the g metric shift N^i as in (3.36) and introduce a new variable,

$$a_m \stackrel{\text{def}}{=} \frac{1}{\tilde{N}} \left(-f_{0m} + c_1^l f_{lm} \right), \tag{C.2}$$

the combination in (3.37) yields

$$\mathbb{E}_0 = \tilde{N} \begin{pmatrix} a_0 & a_j \\ -a_0 c_1^i & -c_1^i a_j \end{pmatrix}, \quad \mathbb{E}_2 = \begin{pmatrix} 0 & 0 \\ \left(h^{il} - c_2^i c_2^l \right) \tilde{N}_l & \left(h^{il} - c_2^i c_2^l \right) \tilde{h}_{lj} \end{pmatrix}, \tag{C.3}$$

and the last matrix $\mathbb{E}_1$ becomes

$$\mathbb{E}_1 = \begin{pmatrix} c_2^l \tilde{N}_l & c_2^l \tilde{h}_{lj} \\ -\left(\tilde{N} a_0 c_2^i + c_2^l \tilde{N}_l c_1^i \right) & -\left(\tilde{N} c_2^i a_j + c_1^i c_2^l \tilde{h}_{lj} \right) \end{pmatrix}, \tag{C.4}$$

where we have moved the scalar terms around in all expressions. Then the match with equation (3.38), i.e. $\mathbb{A}^2 = \mathbb{E}_0$, $\mathbb{B}^2 = \mathbb{E}_2$ and $\mathbb{A}\mathbb{B} + \mathbb{B}\mathbb{A} = \mathbb{E}_1$, implies

$$\mathbb{A} = \frac{1}{\tilde{N}\sqrt{x}} \begin{pmatrix} a_0 & a_j \\ -a_0 c_1^i & -c_1^i a_j \end{pmatrix}, \quad \mathbb{B} = \sqrt{x} \begin{pmatrix} 0 & 0 \\ D^i{}_k \tilde{N}^k & D^i{}_j \end{pmatrix}, \tag{C.5}$$

with the new variables

$$x \stackrel{\text{def}}{=} \frac{1}{\tilde{N}^3} \left(a_0 - c_1^l a_l \right), \quad \sqrt{x} D^i{}_k \stackrel{\text{def}}{=} \sqrt{\left(h^{il} - c_2^i c_2^l \right) \tilde{h}_{lk}}. \tag{C.6}$$

The form of the matrix $\mathbb{A}$ is easy to read off since squaring $\mathbb{E}_0$ gives $\mathbb{E}_0^2 = x\mathbb{E}_0$. At this point, we would like to establish a relation between the two coefficient functions c_1 and c_2. To accomplish this, we first observe that the matrix $D = \sqrt{H\tilde{h}}$ where both H and $\tilde{h}$ are symmetric matrices[1]. This means that we can find an expression for its inverse by expanding

[1] The variable x is a scalar.

the square-root. One commences by rewriting $D = \sqrt{\mathbb{1} + (H\tilde{h} - \mathbb{1})}$ where $\mathbb{1}$ is the identity matrix, whose expansion in powers of $H\tilde{h} - \mathbb{1}$ is

$$D = \mathbb{1} + \frac{1}{2}(H\tilde{h} - \mathbb{1}) - \frac{1}{8}(H\tilde{h} - \mathbb{1})^2 + \dots \tag{C.7}$$

Now, if we multiply this expansion by $\tilde{h}$ and compare it with the expansion of the transpose product, $(\tilde{h}D)^{\mathrm{T}}$, we realize that they are identical owing to that $\tilde{h}$ and H are symmetric. In component form, this translates to

$$\tilde{h}_{ik}D^k{}_j = \tilde{h}_{jk}D^k{}_i. \tag{C.8}$$

Using this relation and comparing the expressions in (C.5) and (C.4) through $\mathbb{AB} + \mathbb{BA} = \mathbb{E}_1$, one obtains

$$c_2^i = \frac{1}{\tilde{N}}D^i{}_k\left(c_1^k - \tilde{N}^k\right). \tag{C.9}$$

If we introduce our redefined shift variable n^i as

$$n^k \stackrel{\text{def}}{=} \frac{1}{\tilde{N}}\left(c_1^k - \tilde{N}^k\right), \tag{C.10}$$

this reduces to

$$c_2^i = D^i{}_k n^k. \tag{C.11}$$

By inserting this result into equation (C.6) and using (C.8), one arrives at a matrix equation for D,

$$\sqrt{x}D = \sqrt{\left(h^{-1} - Dn(Dn)^{\mathrm{T}}\right)\tilde{h}}, \tag{C.12}$$

which is solvable in terms of n^i, h and $\tilde{h}$. In other words all variables c_1, c_2 and D can be written as combinations of these two variables, from which we infer that there exists shift variables which render $N\sqrt{g^{-1}f}$ linear in N. Explicitly, we solve for D by squaring (C.12). Once again using relation (C.8) and collecting terms with D, this implies in component form

$$D^i{}_j Q^j{}_k D^k{}_l = h^{ij}\tilde{h}_{jl},\ \ Q^j{}_k = x\delta^j{}_k + n^j n^m \tilde{h}_{mk}. \tag{C.13}$$

Multiplying both sides with the matrix Q from the right, rearranging terms and taking the square-root, this yields

$$(DQ)^2 = h^{-1}\tilde{h}Q \ \Rightarrow\ D = \left(\sqrt{h^{-1}\tilde{h}Q}\right)Q^{-1}. \tag{C.14}$$

Here, we would like to calculate the inverse matrix Q^{-1} to establish D and it is easy to do through $QQ^{-1} = \mathbb{1}$ where we observe that $(nn^{\mathrm{T}}\tilde{h})^2 = (n^{\mathrm{T}}\tilde{h}n)nn^{\mathrm{T}}\tilde{h}$ since $\tilde{h}$ is a symmetric matrix. This propels us towards the conclusion,

$$Q^{-1} = \frac{1}{x}\left(\mathbb{1} - nn^{\mathrm{T}}\tilde{h}\right). \tag{C.15}$$

Indeed, D can be determined as a function of n^i, h and $\tilde{h}$. Expressed with our new shift variable n^i, equation (C.2) corresponds to

$$a_0 = \tilde{N} + n^k\tilde{N}_k,\ \ a_i = n^k\tilde{h}_{ki}. \tag{C.16}$$

To summarize, this calculation shows that it is possible to rewrite the shift variable N^i of the g-sector as

$$N^i = \tilde{N} n^i + \tilde{N}^i + N D^i{}_k n^k, \tag{C.17}$$

which gives a Lagrangian linear in N which allows us to remove one of the Bouleware-Deser ghost's degrees of freedom, *provided* that no higher order interaction terms introduce nonlinearities in N (this is why we investigate the properties of the potential and find interaction terms which adhere to this). The last degree of freedom is eliminated by demanding this constraint to be constant in time [62]. As previously stated, this calculation is analogous in pure bimetric theory, but the required variable definitions and relations are slightly more complicated [5]. It is also possible to construct a proof in the Stückelberg formalism, [86] [67], and in the vielbien formulation of massive/bimetric gravity, [16], with a full analysis in [87].

In massive gravity, the extra constraint coming from the redefined shift can also be inferred from an analysis of the Hessian of the Hamiltonian, $\mathcal{H}$, of the theory, [60] [9], which is calculated through

$$H_{\mu\nu} = \frac{\partial^2 \mathcal{H}}{\partial N^\mu \partial N^\nu}. \tag{C.18}$$

If $\det H_{\mu\nu} = 0$, this guarantees that the Hessian cannot be inverted and hence we cannot solve the equations of motion for all shift variables and the lapse variable. One of them must be deduced as a combination of the others, implying a constraint. Of course, $\det H_{\mu\nu} = 0$, is not achieved by default for a generic Hamiltonian, so we must choose potential terms which obey this condition.