Cosmological Tests of General Relativity

Theresa Baker
New College
University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2013
Abstract

Understanding the apparent accelerating expansion rate of the universe is a challenge for modern cosmology. One category of explanations is that we are using the wrong gravitational physics to study the observations. Our paradigmatic theory of gravity – Einstein’s theory of General Relativity – may be subsumed by a larger theory.

This thesis develops a selection of tools for testing General Relativity and the numerous alternative theories of gravity that have been put forward. I advocate that an elegant and efficient way to test this space of theories is through the use of parameterized frameworks. Inspired by the Parameterized Post-Newtonian framework I develop a new formalism, the Parameterized Post-Friedmann formalism, that aims to unify the linear cosmological perturbation theory of many alternatives to General Relativity.

Having introduced the Parameterized Post-Friedmann formalism and demonstrated its application via a suite of examples, I examine several issues surrounding parameterized tests of gravity. I first consider how the structure of a parameterization can influence the constraints obtainable from a given set of data. I then consider how to describe the growth of the large-scale structure of the universe in a parameterized manner. This leads to a convenient tool for calculating corrections to the growth rate of structure in modified theories, which can be used both with the Parameterized Post-Friedmann formalism or independently of it. I present forecasts for how well generalized deviations from General Relativity will be constrained by the next generation of galaxy surveys.

Throughout, this thesis aims to take a synoptic approach to theories of modified gravity, rather than focussing on specific models. A question yet to be answered is whether this approach is realistic in practical terms. The final part of this thesis takes the first steps towards an answer.
Acknowledgements

I would like to extend many thanks to everyone who has supported me over the past three years. Most importantly, I would like to thank my supervisor Pedro G. Ferreira for his unfailing encouragement, patience, and constant stream of good ideas. I am also extremely grateful to my collaborators Costas Skordis and Dimitrios Psaltis; I learn new things from every single conversation with them.

Many thanks also to everyone at the Oxford Astrophysics department (including those who are now based elsewhere), and in particular to the members of the cosmology research group, for providing such a friendly and stimulating environment. I am so glad that I can stay here a little longer to continue learning from you all.

Huge thanks to the inhabitants of office 512, ‘the tower’, and the larger student community for their friendship and camaraderie.

Finally, heartfelt thanks to my family for their continual love and support.
Publications Statement

The work in this thesis is based on four published papers and two that are in preparation:


I certify that the work presented here is my own, with the following details of coauthorship:

Chapter 2 is based on publications 1) and 4) above.

Chapter 3 is based on publication 2) above. The figures in this chapter were prepared by Joe Zuntz.

Chapter 4 is entirely my own work, based on publication 3) above.

Chapter 5 is based on a paper in preparation, as described in 5) above.

The figure presented in the conclusions is taken from a publication in preparation, described in 6) above. This figure was prepared in collaboration with D. Psaltis.
Notation and Conventions

Except where explicitly stated otherwise, equations in this thesis are written using a conformal time coordinate \( \eta \). Differentiation with respect to conformal time is denoted by a dot. For a substantial part of chapter 5 it is more convenient to use the logarithm of the cosmological scale factor as a time coordinate, which I denote as \( x = \ln a(\eta) \). Derivatives with respect to \( x \) are indicated by a prime.

Greek indices run over four dimensions of spacetime, whilst Latin indices run over three spatial dimensions (with the exception of §1.3.4, where they are used to indicate a general number of \( D \) dimensions). The metric signature convention used is \( \{-,+,+,+\} \).

Bold font is used to indicate four-vectors, for example, \( \mathbf{v} \). Arrows are used to indicate spatial three-vectors, for example, \( \vec{v} \).

I set the speed of light to unity throughout, \( c = 1 \). The inverse square of the reduced Planck mass is denoted by \( M_P^{-2} = 8\pi G \). Sometimes I will use the notation \( G_N \) or \( G_0 \) to explicitly refer to the Newtonian gravitational constant \((6.674 \text{ m}^3 \text{ kg}^{-1} \text{s}^{-2} \text{ to three decimal places})\), as opposed to any renormalized, effective gravitational constant. A subscript zero is used to denote a quantity evaluated ‘today’, at \( z \simeq 0 \).

To maintain consistency with the titles of published works (such as ‘the Parameterized Post-Friedmann formalism’) I have tried to use the American spelling of ‘parameterized’ and similar words throughout the text. I apologize for any inconsistencies and typographical errors that may have slipped through.
Contents

List of Figures ix

1 Introduction 1

1.1 General Relativity ................................. 1

1.1.1 Gravity in the Twentieth Century .................. 1

1.1.2 So Why Modify Gravity? .......................... 3

1.1.3 Fundamentals of General Relativity ............... 5

1.1.3.1 Curvature of a Manifold ....................... 5

1.1.3.2 The Gravitational Action ...................... 7

1.1.3.3 The Gravitational Field Equations .............. 9

1.1.4 Cosmological Perturbation Theory in an FRW Universe . 10

1.1.4.1 Linear Perturbation Equations ................. 10

1.1.4.2 The Origin of Matter Perturbations ............. 14

1.2 Tools for Testing Gravity ........................... 18

1.2.1 The Cosmic Microwave Background ................ 18

1.2.1.1 Multipoles and Method ....................... 18

1.2.1.2 The Multipoles at Last Scattering .............. 20

1.2.1.3 From Last Scattering to Today ................ 23

1.2.1.4 Observable Quantities ......................... 25

1.2.1.5 Peak Positions ............................... 26

1.2.1.6 Current Measurements ......................... 27

1.2.2 Galaxy Surveys ................................. 29

1.2.2.1 The Growth Rate and Redshift-Space Distortions 29

1.2.2.2 Baryon Acoustic Oscillations .................. 31

1.2.3 Gravitational Lensing ........................... 34

1.2.3.1 Cosmic Shear ................................. 34

1.2.3.2 CMB Lensing ................................. 36

1.3 Theoretical Properties of Modified Gravity ........... 38
## CONTENTS

1.3.1 Lovelock’s Theorem and the Landscape of Modified Gravity 38  
1.3.2 New Degrees of Freedom 39  
1.3.3 Higher Derivatives and Instabilities 42  
  1.3.3.1 The Ostrogradski Instability 42  
  1.3.3.2 Ghosts 45  
1.3.4 Higher-Dimensional Theories of Gravity 46  
1.3.5 Screening Mechanisms 48  
1.4 Thesis Layout 50  

2 The Parameterized Post-Friedmann Framework for Theories of Modified Gravity 53  
2.1 Introduction 53  
2.2 The PPF Formalism 57  
  2.2.1 Basic Principles 57  
  2.2.2 Background-Level Parameterization 59  
  2.2.3 Perturbation Variables 60  
    2.2.3.1 Metric Variables 60  
    2.2.3.2 New Degrees of Freedom 60  
    2.2.3.3 Matter Perturbations 62  
  2.2.4 Perturbation Framework 63  
  2.2.5 Constraint Equations 65  
2.3 Worked Examples 67  
  2.3.1 Scalar-Tensor Theory and $f(R)$ Gravity 68  
  2.3.2 Einstein-Aether Theory 71  
  2.3.3 DGP 72  
  2.3.4 Eddington-Born-Infeld Gravity 76  
  2.3.5 Hořava-Lifschitz Theory 78  
  2.3.6 Horndeski Theory 81  
  2.3.7 GR with a Dark Fluid 85  
2.4 The PPF Coefficients 88  
  2.4.1 Degeneracy 88  
  2.4.2 Signatures and Scale-Dependence 89  
  2.4.3 The Quasistatic Limit 91  
  2.4.4 Connection to Screening Parameterizations 93  
2.5 Conclusions 94  
2.6 Supporting Material for Chapter 2 96  
  2.6.1 Transformation of $f(R)$ Gravity into a Scalar-Tensor Theory 96
2.6.2 Is the Bianchi 0-Component Always the E.O.M.? A Plausibility Argument .............................................. 96

3 Ambiguities in Parameterizations of Modified Gravity 99
  3.1 Introduction .................................................. 99
  3.2 Parameterizations .......................................... 101
    3.2.1 Parameterization A ................................... 102
    3.2.2 Parameterization B ................................... 103
  3.3 Spectra ..................................................... 104
  3.4 Constraints ................................................. 107
    3.4.1 Parameter Estimation .................................. 107
    3.4.2 Insufficient Ansatzes ................................ 109
    3.4.3 Variant Parameterization .............................. 110
  3.5 Conclusions ............................................... 112

4 The Growth of Parameterized Perturbations in an Einstein-de Sitter Universe 115
  4.1 Introduction ................................................ 115
  4.2 Parameterization of Metric-Only Gravity Theories ............... 117
  4.3 Density Perturbations ..................................... 119
    4.3.1 Subhorizon Scales .................................. 120
    4.3.2 Superhorizon Scales ................................. 122
    4.3.3 Connection to Constraints ........................... 124
    4.3.4 Other Growth Observables ........................... 125
  4.4 Theories With Additional Degrees of Freedom ..................... 129
    4.4.1 Subhorizon Scales .................................. 132
    4.4.2 Superhorizon Scales ................................. 134
  4.5 Conserved Superhorizon Perturbations? ......................... 134
    4.5.1 Metric-Only Theories ................................. 136
    4.5.2 Theories with Additional Degrees of Freedom .......... 137
  4.6 Conclusions ............................................... 139
  4.7 Supporting Material for Chapter 4 .......................... 140
    4.7.1 System of Equations for Theories with Extra D.o.F. .... 140
    4.7.2 Conservation of Z in Metric-Only Theories .................. 141
    4.7.3 \( \delta_M \) and \( \delta_E \) in Theories with Extra D.o.F. ........ 144

5 Corrections to the Growth Rate in Modified Gravity 147
  5.1 Introduction ................................................ 147
  5.2 The Quasistatic Approximation .................................. 149
List of Figures

1.1 The matter power spectrum as measured by the Baryon Oscillation Spectroscopic Survey (BOSS). ........................................... 17
1.2 Temperature power spectrum of the CMB from the Planck experiment. .............................................................. 22
1.3 Cosmological models consistent with the first release of WMAP data, in the \( \Omega_M - \Omega_{\Lambda} \) plane. ................................................. 27
1.4 The redshift-space correlation function of the 2dF survey \([83, 149]\), plotted as a function of transverse (\( \sigma \)) and radial (\( \pi \)) separation between pairs of galaxies. ................................. 30
1.5 Measurements of the density-weighted growth rate, \( f\sigma_8 \), from several recent galaxy redshift surveys. ........................................... 32
1.6 The power spectrum of the lensing potential measured by the Planck satellite. .............................................................. 37
1.7 An incomplete visualization of the space of modified gravity theories. Theories have been principally classified according to the way in which they avoid the restrictions of Lovelock’s theorem. ........... 40

3.1 CMB power spectra for variations of the linear term of the Taylor expansion in \( \Omega_{\Lambda}(z) \). The Taylor expansion is the ‘level 2’ ansatz (see §3.1) for the function \( G_{\text{eff}} \) that appears in the Poisson equation, as defined in eq.(3.5). ......................................................... 106
3.2 The growth function \( f(z)\sigma_8(z) \) for the same models shown in Fig. 3.1.106
3.3 A demonstration of how extreme changes to the effective gravitational constant can be permitted by the CMB when they can be counteracted by a significant gravitational slip. ......................... 108
3.4 The growth rate \( f(z)\sigma_8(z) \) corresponding to the CMB spectra shown in figure 3.3. ......................................................... 109
3.5 Constraints on the free functions, evaluated at redshift zero, in a model restricted to linear evolution in \( \Omega_{\Lambda}(z) \) (filled contours) and one with cubic evolution (line contours). ........................................... 110
3.6 Joint constraints on the slip parameter $\zeta$ and $G_{\text{eff}}$ at $z = 0$, for parameterization A (black lines) and B (filled green). In both cases 68% and 95% contours are shown. ......................... 112

3.7 Likelihoods of the effective gravitational and slip parameters at $z = 0$, for parameterization A (thick blue) and B (thin green). .................. 113

4.1 Joint constraints on the slip parameter $\zeta$ and $\mu_P$ at $z = 0$, for parameterization A (black lines) and B (filled green). In both cases 68% and 95% contours are shown. ......................... 124

4.2 Low-$l$ power spectrum of the ISW effect induced by a theory of modified gravity constructed in parameterization A (see §4.2 for details). ......................... 128

5.1 Examples of how the growth rate is affected by different source terms in eq.(5.12). Effects on the growth rate are expressed as a percentage deviation from the GR prediction. ......................... 155

5.2 As Fig. 5.1, but for sustained deviations from GR. ...................... 158

5.3 Lower panels show the fractional deviation of the density-weighted growth rate from its $\Lambda$CDM+GR model caused by the source functions in the upper panels, see eq.(5.32). ......................... 160

5.4 These functions describe the extent to which modifications to the equation of state of the non-matter sector (ie. $\omega_X \neq -1$) source corrections to the growth rate of GR, see eqs.(5.34) and (5.38). ...................... 162

5.5 Forecast constraints on $\delta S$ for a typical next-generation galaxy survey, where $\delta S$ sources deviations from the GR growth rate (see eq.(5.22)). ......................... 168

6.1 A two-dimensional parameter space for gravitational regimes that exist in our universe, quantified in terms of their characteristic gravitational potential ($x$-axis) and approximate spacetime curvature ($y$-axis). ......................... 175
Chapter 1

Introduction

1.1 General Relativity

1.1.1 Gravity in the Twentieth Century

Albert Einstein’s theory of General Relativity (hereafter GR) is the paradigmatic theory of gravity. Almost a century after Einstein’s pivotal publications [119, 120, 121], GR is taught in universities around the world as part of the standard lore of physics – as immutable as the fundamentals of thermodynamics or Maxwell’s laws of electromagnetism.

It is tempting to assume that GR has always enjoyed this status. In fact this is not true; though the results of Sir Arthur Eddington’s 1919 eclipse expedition to Principe were in agreement with Einstein’s theory [117], with a measurement error of roughly 20% they could hardly be considered decisive [127]. One might well wonder how quickly GR would have been adopted if Eddington, a highly influential figure in British science at the time, had not given his support to the theory.

For the first sixty years of its existence GR remained an appealing but largely unconfirmed theory, predominantly because the technological capabilities to measure its subtle non-Newtonian predictions for the Solar System environment did not exist. However, theoreticians of the time did not idly wait for GR to be confirmed or refuted. Instead they began a mathematical exploration of the fertile new territory that Einstein had unveiled, and a crop of alternative gravity theories sprang up [226]. The majority of these adopted the basic ingredients of GR but added embellishments such as new dynamical fields (scalars, vectors, additional tensor fields, and any combination thereof) and sometimes prior geometric features (such as fixed second metrics and non-dynamical scalars) [309]. One of the most well-known of these theories is the scalar field theory of Jordan, Brans...
1. INTRODUCTION

and Dicke [56] that is still discussed today, albeit usually as a toy model or limit of a more complex theory. Some of this crop of alternative theories were later found to be in gross contradiction with experiment and are no longer considered, e.g. the conformally flat theories of Nordstrom [103] and Ni [227], which forbid the existence of gravitational lensing effects.

By the 1970s new satellite-based experiments meant that more theories fell foul of Solar System tests. But some are not yet fully consigned to the dusty annals of physics, thanks to the fresh wave of scrutiny that has fallen on GR during the past decade or so. Whilst most of the original gravity theories of the 1970s have been ruled out, they have spawned offspring that are considered today in cosmological contexts, as candidate explanations for the accelerated expansion of the universe. For example, one might argue that Rosen’s bimetric theory [263, 264] and Rastall’s theory [257] – both prior-geometric theories featuring a flat, non-dynamical metric – are the ancestors of the currently popular Massive Gravity theory (see [151] for a review). Similarly, Einstein-Aether theories [170, 171, 323] can trace their ancestry back to the vector field theories of Will & Nordvedt [310] and Hellings & Nordvedt [150].

There are several key observations to be drawn from reflecting upon this period of science history:

1. The space of mathematically-possible gravitational theories is much richer than General Relativity (see [81] for a comprehensive review).

2. This rich space of theories includes examples that become equivalent to GR in a particular regime, whilst displaying new phenomena in other regimes\(^1\). Therefore the apparent verification of a theory in one regime – such as the Solar System results mentioned above – does not guarantee it is the complete, correct theory.

3. Testing the predictions of gravitational theories in a wide range of physical systems is therefore crucial. Such tests should be attempted as soon as the relevant technology becomes available, in order to guide theoretical developments. Arguably some of more violently exotic theories described above would never have been constructed if even marginally better data been available at the time.

\(^1\)Based on examples from particle physics, one usually thinks of a theory being valid on energy scales below some cut-off scale. I have refrained from referring to energy scales here because it is possible that other variables may determine what regime or limit a system is in, see §1.3.5 and §6.
1.1.2 So Why Modify Gravity?

General Relativity has passed every Solar System-based test to date with flying colours [43, 197]. Why, then, has the past decade seen a revival in alternative gravitational theories?

The current renaissance was sparked by multiple lines of evidence suggesting that approximately 70% of the energy density of our universe exists in a non-matter form (some of these observations will be discussed in §1.2; see also [12] for useful discussions of each). The simplest explanation for this would be the existence of a cosmological constant term, \( \Lambda \), in the field equations of gravity (see eq.(1.20) to follow). In fact there are two ways a cosmological constant term can enter the Einstein field equations – as a geometric term appearing on the left-hand side, and on the right-hand side as an energy density belonging to the vacuum. Without an \emph{a priori} reason to set the geometric factor to zero, there will be some degree of cancellation between the two contributions on either side to produce an overall, effective cosmological constant. In order to reproduce the observed expansion rate of the universe, the effective cosmological constant would need to have a value of \( \Lambda \approx H_0^2 \) (where \( H_0 \) is the Hubble factor at redshift zero).

When expressed as an energy density this is equivalent to:

\[
\rho_{\text{vac}} = \Lambda M_P^2 \approx (10^{-47} \text{ GeV})^4 \approx 10^{-123} M_P^4 \tag{1.1}
\]

where \( M_P^2 = (8\pi G_N)^{-1} \) is the reduced Planck mass squared.

Quantum Field Theory tells us to consider the energy of the vacuum as arising from a simple harmonic oscillator of mass \( m \) at every point in spacetime, with ground-state energy \( E = \omega/2 = \sqrt{k^2 + m^2}/2 \) (in units \( \hbar = c = 1 \)). Integrating over all possible wavenumbers up to some cut-off scale \( k_{\text{max}} \), this gives an energy density of [307]:

\[
\tilde{\rho}_{\text{vac}} = \int_0^{k_{\text{max}}} \frac{d^3k}{(2\pi)^3} \frac{\sqrt{k^2 + m^2}}{2} \approx \frac{k_{\text{max}}^4}{16\pi^2} \tag{1.2}
\]

for some \( k_{\text{max}} \gg m \). If we believe that GR is the correct theory of gravity all the way up to Planck energies, then the integral above evaluates to a number of order \( 10^{-2} M_P^4 \) – a staggering discrepancy with the observed value in eq.(1.1). The problem is to explain why the ‘vacuum energy cosmological constant’ has been cancelled to such a fantastically high degree of accuracy by the ‘geometric cosmological constant’, when in principle they are totally independent of each other. One could consider using other energy scales from particle physics as the upper limit to eq.(1.2), eg. the QCD scale at \( \sim 0.1 \text{ GeV} \), but this does not resolve
1. INTRODUCTION

the discrepancy. This is the fine-tuning problem: data seem to indicate a value of the cosmological constant that is completely at odds with what we would expect from the robust theories of modern particle physics.

The situation is made worse still by the coincidence problem: the realization that the period of time for which the fractional energy densities of the effective cosmological constant and matter are comparable (as we observe today) is extremely short. The simple explanation that we are privileged observers, alive at a special time, is in opposition to the Cosmological Principle. Thus the cosmological constant explanation for the non-matter sector of our universe, though aesthetically appealing because of its simplicity, seems to be grossly inadequate.

It is for these reasons that we have begun to consider alternative, more complex descriptions of the non-matter sector. It is possible that there exists some new field in our universe that does not belong to the Standard Model of particle physics. Such theories fall under the umbrella description dark energy; see [84] for a review of the large body of work in this area. Another possible explanation for the observations is that Einstein’s theory of General Relativity is not, in fact, the complete description of gravity in all regimes of our universe. This is the category of modified theories of gravity, which I will introduce more fully in §1.3. A third option is that our mathematical description of the inhomogeneous matter distribution in the universe is too simplified, and the inferred acceleration is an artefact of this incorrect treatment\(^1\) [80, 256].

The distinction between dark energy and modified gravity theories is often blurred, since modifications to the field equations of GR have an equivalent representation on the left-hand side ('gravity/geometry') and the right-hand side ('the contents of the universe') [196]. For the purposes of this thesis I will implement the following definition: a dark energy model will be one in which the gravitational action is given solely by eq.(1.20) and all other fields (both matter fields and new non-GR fields) are minimally coupled to the metric. This definition is quite restrictive, limiting the domain of pure dark energy models largely to quintessence and k-essence theories (see §1.3.2), plus some unified dark sector models such as Chaplygin gases [140]. Strictly speaking, under this definition Einstein-Aether gravity (§2.3.2) is a dark energy model. However, due to its violation of Lorentz symmetry (a key property of GR) it is usually considered as belonging to the category of modified gravity theories\(^2\).

\(^1\)For example, in an inhomogeneous universe it is not guaranteed that the operations of volume-averaging a quantity and evolving it in time necessarily commute, at least not on all lengthscales. This concept is referred to as backreaction.

\(^2\)Also, once the constraint arising from the Lagrange multiplier has been used, the field equations look more akin to those of a modified gravity theory than they do to those of, say,
In §1.3 I will elaborate on the different categories of modified gravity theories. For the present I will simply state that, according to the definition above, a modified gravitational action might have some of the following features: a function coupling to the Einstein-Hilbert term, non-standard appearances of the various metric-related tensors (Riemann, Ricci, Einstein, Weyl...), contractions of these metric-related tensors with new fields or their derivatives, non-minimal coupling of matter fields to the metric, prior-geometric elements (such as fixed additional tensor fields or non-dynamical scalars), or contributions to the action requiring an integral over a number of dimensions different from four.

1.1.3 Fundamentals of General Relativity

1.1.3.1 Curvature of a Manifold

Perhaps Einstein’s most profound insight was to realize that gravitational physics admits a geometrical interpretation. This idea led him to postulate that spacetime could be described as a pseudo-Riemannian manifold\(^1\), that is, one on which a spacetime interval \(ds\) can be written in the form

\[
ds^2 = g_{\mu\nu}(x)dx^\mu \, dx^\nu \tag{1.3}
\]

where \(g_{\mu\nu}\) is the metric tensor and \(x^\mu\) are a system of coordinates on the manifold.

At each point on the manifold one can construct a tangent space with a set of basis vectors \(\{e_\mu\}\). In order to define a derivative on the manifold one needs to consider the values of quantities at two infinitesimally separated spacetime points; however, the basis vectors of the tangent spaces at these two points may be different. We therefore construct a covariant derivative that ‘takes into account’ the change of basis vectors as we move across the manifold. This is done by introducing an affine connection:

\[
\Gamma^\gamma_{\mu\nu} = e^\gamma \cdot \frac{\partial e_\mu}{\partial x^\nu} \tag{1.4}
\]

It is shown in numerous textbooks (e.g.,[155, 233]) that the affine connection can

\(^1\)A strictly Riemannian manifold requires \(ds^2 > 0\). This not true for all intervals in Special Relativity or GR due to the difference in sign between the spatial and temporal parts of the metric signature. On a pseudo-Riemannian manifold there is no restriction on the sign of \(ds^2\).
also be expressed in terms of derivatives of the metric as\(^1\):

\[
\Gamma^\gamma_{\mu\nu} = \frac{1}{2} g^{\gamma\delta} \left( \partial_\mu g_{\delta\nu} + \partial_\nu g_{\mu\delta} - \partial_\delta g_{\mu\nu} \right) \tag{1.5}
\]

Scalar quantities are independent of the basis vectors at any point, so their covariant derivatives are simply equal to their partial derivatives. The covariant derivatives of a four-vector with components \(v^\mu\) and rank-2 tensor with components \(M^{\mu\nu}\) are:

\[
\nabla_\alpha v^\mu = \partial_\alpha v^\mu + \Gamma^\mu_{\alpha\beta} v^\beta \tag{1.6}
\]

\[
\nabla_\alpha M^{\mu\nu} = \partial_\alpha M^{\mu\nu} + \Gamma^\mu_{\alpha\sigma} M^{\sigma\nu} + \Gamma^{\nu}_{\alpha\sigma} M^{\mu\sigma} \tag{1.7}
\]

Covariant derivatives of higher-rank tensors can be defined analogously.

The argument in eq.(1.3) indicates that the form of the metric elements will change if we pick a different system of coordinates. For example, we can choose a system of coordinates such that \(g_{\mu\nu}(x_0)\) becomes the Minkowski metric, \(\eta_{\mu\nu} = \text{diag}(-1,1,1,1)\), at the point \(x_0\). If a region of spacetime is flat, i.e. totally empty of all matter (including electromagnetic fields and similar), then there exists a choice of coordinates which allows the metric to be written in Minkowski form at all points in the spacetime. If the spacetime is not flat then it is only possible to write the metric in Minkowski form over some small region, termed a local inertial frame.

However, finding the correct coordinate choice to bring the metric to Minkowski form may be very difficult. How can we tell if a spacetime is genuinely curved, or simply appears that way because of a poor choice of coordinates? This information is contained in the Riemann curvature tensor.

I stated above that the covariant derivative encodes information about how the set of basis vectors changes as we move over the manifold. However, we know that in flat space no such change of basis vectors occurs (this is not a feature of Newtonian gravity). This suggests that, if we want to determine whether a spacetime is flat or curved, we should compare what happens when we transport a four-vector along two different paths. This leads to the definition of the Riemann curvature tensor as:

\[
\nabla_\beta \nabla_\gamma v^\delta - \nabla_\gamma \nabla_\beta v^\delta = R^\delta_{\alpha\beta\gamma} v^\alpha \tag{1.8}
\]

\(^1\)I will consider only torsionless manifolds, such that the affine connection is symmetric under exchange of its two lower indices. Formally, the connection defined in eq.(1.4) is the Levi-Civita connection. This is the only affine connection on a torsionless, pseudo-Riemannian manifold that satisfies the condition of metricity \((\nabla_\alpha g_{\mu\nu} = 0)\), so I will use the two names interchangeably.
for an arbitrary four-vector $v^\gamma$. The corresponding coordinate expression is:

$$R^\delta_{\alpha\beta\gamma} = \partial_\beta \Gamma^\delta_{\alpha\gamma} - \partial_\gamma \Gamma^\delta_{\alpha\beta} + \Gamma^\delta_{\sigma\beta} \Gamma^\sigma_{\gamma\alpha} - \Gamma^\delta_{\sigma\gamma} \Gamma^\sigma_{\beta\alpha}$$  \hspace{1cm} (1.9)

In a system of local inertial coordinates at point $x_0$ we have $g_{\mu\nu}(x_0) = \eta_{\mu\nu}$ and $\partial_\gamma g_{\mu\nu}|_{x=x_0} = 0$. However, we cannot choose a system of coordinates that sets the second derivatives of the metric to zero everywhere; therefore $R^\delta_{\alpha\beta\gamma}$ will have non-zero elements. The converse is also true: a spacetime is flat if and only if $R^\delta_{\alpha\beta\gamma} = 0$ everywhere, for all elements of the Riemann curvature tensor.

In order to formulate the Einstein field equations we will also need the Ricci tensor and Ricci scalar, which are obtained by contracting the Riemann curvature tensor as shown below:

$$R_\alpha^\gamma = R^\delta_{\alpha\delta\gamma} \hspace{1cm} R = g^{\mu\nu} R_{\mu\nu}$$  \hspace{1cm} (1.10)

### 1.1.3.2 The Gravitational Action

The actions for scalar fields, electromagnetism and a point particle are all quadratic in first-order derivatives of the dynamical degrees of freedom; this results in field equations that are second-order in derivatives. Naively, then, one might expect gravity to follow suit. However, when we try to build a covariant, scalar action from the metric and its first derivatives, we meet an unusual obstacle. Such an action will reduce to a trivial constant in a local inertial frame, where $g_{\mu\nu} = \eta_{\mu\nu}$ and $\partial_\alpha g_{\mu\nu} = 0$. Since a scalar quantity has the same value in all frames, this hypothetical action must then be a constant in all frames, and our attempt to construct an action for gravity has failed.

We are left with no choice but to allow second derivatives of the metric to appear in the gravitational action. This should ring alarm bells – as I will discuss in §1.3.3.1, such a theory will generally be energetically unstable. This eventuality can be avoided if the gravitational action is exactly linear in the second-order derivatives of the metric; initially, it appears that the dangerous second-order derivative can then be removed through an integration-by-parts. The only scalar quantity constructed from the metric which meets this second-derivative-linearity requirement is the Ricci scalar, $R$. One is then inevitably led to the action for General Relativity (shown here without an action for the matter fields):

$$S_g = \frac{M_P^2}{2} \int \sqrt{-g} \, d^4x \, (R - 2\Lambda)$$  \hspace{1cm} (1.11)

where I have included an allowable constant term, and chosen notation for the
1. INTRODUCTION

constants that anticipates an application to cosmology.

However, the integration-by-parts alone is an unsatisfactory method of removing the troublesome second derivative term from the action. Denoting a general dynamical variable by $q$, the vanishing of the resulting surface term requires that both $\delta q = 0$ and $\delta \dot{q} = 0$ on the boundary of the spacetime (where for this section only a dot represent a derivative with respect to physical time). It is not guaranteed that a classical solution which satisfies all the boundary conditions exists. Furthermore, specifying both a dynamical coordinate and its conjugate momentum on the boundary is problematic from a quantum viewpoint.

A solution to the apparent conundrum is provided by the addition of the Gibbons-Hawking-York counterterm \[137, 314\]. I will sketch only the basic idea here. The Einstein-Hilbert action can be split into a ‘safe’ part that is quadratic in first-order derivatives of the metric, and the remaining troublesome surface term (after the integration-by-parts has been performed):

$$S_g = S_{\text{quad}} + S_{\text{sur}}$$

$$S_{\text{sur}} = \frac{M_P^2}{2} \int_{t=\text{const}} d^3 x \frac{1}{\sqrt{-g}} \partial_{\mu} (g g^{\mu 0}) n_0$$ \hspace{1cm} (1.12)

where $n_0$ is a normal vector to the $t = \text{constant}$ surface. We decompose the metric as:

$$ds^2 = -(N dt)^2 + h_{ij} \left[ dx^i + N^i dt \right] \left[ dx^j + N^j dt \right]$$ \hspace{1cm} (1.13)

Now consider the following action, the Gibbons-Hawking-York counterterm:

$$\tilde{S}_{\text{sur}} = -M_P^2 \int_{t=\text{const}} d^3 x \sqrt{h} K$$ \hspace{1cm} (1.14)

where now specifically $n_{\mu} = -N \delta_{\mu}^0$, and $K = -\nabla_{\mu} n^\mu$ is the trace of the extrinsic curvature. It can be shown, after some straightforward but lengthy algebra \[233\], that:

$$S_{\text{sur}} - \tilde{S}_{\text{sur}} = \frac{M_P^2}{2} \int_{t=\text{const}} d^3 x \sqrt{h} \left[ \frac{\partial_i N^i}{N} \right]$$ \hspace{1cm} (1.15)

In a moment we will consider variations of the total gravitational action with $\delta g^{\mu \nu} = 0$ on the $t=\text{constant}$ boundary (i.e. we still enforce $\delta q = 0'$, but not $\delta \dot{q} = 0'$). Under this restriction none of the quantities in eq. (1.15) vary, and we obtain the result $\delta S_{\text{sur}} - \delta \tilde{S}_{\text{sur}} = 0^1$.

\[1\] Since $\partial_i N^i$ is a purely spatial derivative it can have no variation if the metric is held fixed everywhere on the $t=\text{constant}$ surface.
Putting everything together then, we take the new full gravitational action to be:

\[ \tilde{S}_g = S_g - \tilde{S}_{\text{sur}} \]
\[ = S^{\text{quad}} + S_{\text{sur}} - \tilde{S}_{\text{sur}} \] \hspace{1cm} (1.16)
\[ = \frac{M_P^2}{2} \int d^4x \left( R - 2\Lambda \right) + M_P^2 \int_{t=\text{const}} d^3x \sqrt{h} K \] \hspace{1cm} (1.17)

When we carry out the variation with the restrictions described above, the last two terms in eq.(1.16) cancel. We are left with only the safe, quadratic part of the Einstein-Hilbert action. However, it should be highlighted that the Gibbons-Hawking-York method is not completely failsafe. It is not true that \( S_{\text{sur}} = \tilde{S}_{\text{sur}} \) in general; a discussion of further caveats can be found in [233].

### 1.1.3.3 The Gravitational Field Equations

We can now proceed to obtain the full gravitational field equations. To do this we must couple eq.(1.17) (or eq.(1.11) if we are not concerned about boundary terms) to the action for matter fields. This is generally defined implicitly via the expression:

\[ T^{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\delta L_M}{\delta g_{\mu\nu}} \] \hspace{1cm} (1.18)

where \( T^{\mu\nu} \) is the energy-momentum tensor\(^1\). In §1.1.4 we will implement the case of a perfect fluid, for which the energy-momentum tensor has the form\(^2\):

\[ T^{\mu\nu} = (\rho + P)u^\mu u^\nu + g^{\mu\nu}P \] \hspace{1cm} (1.19)

where \( \rho \) is the energy density of the fluid, \( P \) is its pressure, and \( u^\mu = (-1, 0, 0, 0) \) is the fluid four-velocity in comoving coordinates.

The full Einstein field equations in the presence of matter are then given by the variation of the action \( S_{\text{Tot}} = \tilde{S}_g + \int L_M d^4x \):

\[ G_{\mu\nu} + \Lambda g_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = 8\pi G_N T_{\mu\nu} \] \hspace{1cm} (1.20)

where the first equality defines the Einstein tensor, \( G_{\mu\nu} \). For the present, let us

---

\(^1\)Note here that \( L_M \) is the Lagrangian density of the matter fields, related to the usual Lagrangian by \( L = \sqrt{-g} L \).

\(^2\)It is also possible to write explicit expressions for the energy-momentum tensors of imperfect fluids, charged fluids, etc., but eq.(1.19) will be sufficient for our purposes.
consider the case without a cosmological constant, \( \Lambda = 0 \). Then using the trace of these equations, \( R = -8\pi G_N T^\mu_\mu \), we see that the Ricci scalar is zero in vacuum (or in the presence of a radiation-like fluid with \( P = 1/3\rho \)). This means that in vacuum the gravitational field equations reduce to the simple form:

\[
R_{\mu\nu} = 0 \tag{1.21}
\]

This expression allows us to make a rather interesting comment about the number of dimensions of our universe. In four dimensions the symmetries of the Riemann curvature tensor imply that it has twenty independent components. Eq.(1.21) is equivalent to a set of ten constraint equations for these twenty components, leaving ten degrees of freedom in \( R^\delta_{\alpha\beta\gamma} \).

In contrast, in three dimensions the Riemann tensor has only six independent components, and eq.(1.21) is a set of six equations. Then the only way to satisfy \( R_{\mu\nu} = 0 \) in three dimensions is for all components of the Riemann tensor to be equal to zero. As discussed in §1.1.3.1, this is the requirement for flat space. Therefore it is only in four (or more) dimensions that we can have gravitational fields (ie. curved space) in a vacuum.

Another feature of the Einstein field equations which will be important in later chapters is the divergenceless nature of the Einstein tensor, that is, \( \nabla_\mu G^{\mu\nu} = 0 \). This in turn implies that the divergence of the energy-momentum tensor must vanish:

\[
\nabla_\mu T^{\mu\nu} = 0 \tag{1.22}
\]

The \( \nu = 0 \) component of this expression gives an evolution equation for the stress-energy content of a spacetime. We will use eq.(1.22) in §1.1.4.1 to determine how the energy densities of matter fields evolve in an expanding universe.

### 1.1.4 Cosmological Perturbation Theory in an FRW Universe

#### 1.1.4.1 Linear Perturbation Equations

The basic elements of the modern standard cosmological model, usually known as the ‘ΛCDM’ model, can be found in many texts, for example [111, 236, 306]. Its key features are that, to leading order, the expansion rate of our universe can be described through the evolution of the scale factor \( a(\eta) \) that appears in the in
the Friedmann-Robertson-Walker\(^1\) (FRW) metric as:

\[
ds^2 = a^2(\eta) \left[ -d\eta^2 + \frac{dr^2}{(1 - Kr^2)} + r^2 \left( d\theta^2 + \sin^2 \theta \, d\varphi^2 \right) \right]
\]

(1.23)

I have chosen a conformal time coordinate \(\eta\), which is related to the proper time \(t\) of an observer comoving with the cosmic expansion by \(dt = a(\eta)d\eta\). The parameter \(K\) specifies the curvature of spatial hypersurfaces; it can adopt values greater than, equal to, or less than zero, leading to closed, flat and open universes respectively\(^2\). The evolution of the scale factor is given by the Friedmann equation, which is obtained as the ‘time-time’ part of eqs.(1.20) when using the metric of eq.(1.23):

\[
\left( \frac{\dot{a}}{a} \right)^2 = \frac{8\pi G_N a^2}{3} \rho(a) - K + \frac{\Lambda a^2}{3}
\]

(1.24)

In this expression \(\rho(a)\) denotes the total energy density of all matter components in the universe, such as baryons, cold dark matter, photons and neutrinos. The evolution of the energy density of the \(i\)th matter component is given by eq.(1.22) evaluated with a perfect fluid stress-energy tensor and an FRW metric:

\[
\dot{\rho}_i = -3\mathcal{H} \rho_i (1 + \omega_i)
\]

(1.25)

where \(\omega_i = P_i / \rho_i\) is the equation of state parameter of the matter component and \(\mathcal{H} = \dot{a}/a\) is the conformal Hubble factor. Observations to date indicate that we live in a universe that is very close to being spatially flat (see Table 1.1), so I will set \(K = 0\) from here onwards.

It is clear that we do not live in an isotropic, homogeneous universe as described by eq.(1.23). The first step towards a more realistic description of the matter distribution in our universe is to introduce linear energy density perturbations about the homogeneous background. This linearized description turns out to be accurate for large scales, because long-wavelength density perturbations have only recently re-entered our cosmological horizon (having exited it during the inflation era). Perturbations do not grow when their wavelength is larger than the cosmological horizon scale, because the horizon essentially defines the greatest distance over which causal physics can operate. Therefore long-wavelength perturbations – though laid down during a very early inflationary epoch – have had

---

\(^1\)Or Friedmann-Lemaitre-Robertson-Walker, according to preference.

\(^2\)Note that it is common to rewrite the line element eq.(1.23) in terms of the scaled radial coordinate \(\tilde{r} = \sqrt{|K|} r\), so that \(Kr^2 \equiv k\tilde{r}^2\), where \(k\) can have the values +1, 0, −1.
insufficient time to grow beyond the regime of linear perturbation theory. In contrast, short-wavelength modes of the density field have exited the linear regime and require a fully non-linear solution of the gravitational field equations and fluid equations. Generally this is only possible numerically, so N-body computer simulations are an essential tool for studying small-scale phenomena.

I will present here the relevant equations for linear cosmological perturbation theory in GR; how these equations are modified in other theories of gravity is the raison d’être of this thesis. We begin by perturbing the metric elements. This excites ten degrees of freedom: four scalar perturbations, four vector perturbations (coming from two three-vectors with two constraint equations) and two tensor perturbations (arising from a symmetric $3 \times 3$ tensor with four constraint equations). However, the decomposition theorem shows that (at linear order only) scalar, vector and tensor perturbations are decoupled [111]. Furthermore, vector perturbations decay in an expanding universe, whilst tensor perturbations are relevant only for the description of gravitational waves – they are sourced only by the anisotropic stress of the matter distribution [305], which is negligibly small at late times in the universe. Only scalar perturbations exhibit the growing modes needed to cause the formation of large-scale structure in the universe, so I will focus exclusively on them.

The notation that I will use for scalar perturbations to the metric is displayed by the following line element [31, 284]:

$$ds^2 = a(\eta)^2 \left[ -(1 - 2\Xi) d\eta^2 - 2(\vec{\nabla}_i \epsilon) d\eta dx^i + \left( 1 + \frac{1}{3} \beta \right) \gamma_{ij} + D_{ij} \nu \right] dx^i dx^j$$

where $\gamma_{ij}$ is a flat spatial metric and $D_{ij}$ is a derivative operator that projects out the longitudinal, traceless, spatial part of the metric perturbation:

$$D_{ij} = \vec{\nabla}_i \vec{\nabla}_j - 1/3 \gamma_{ij} \vec{\nabla}^2$$

In Chapter 2 I will work with explicitly gauge-invariant variables, but here I will make contact with conventional expressions by working in the conformal Newtonian gauge. In this gauge one sets $\epsilon = \nu = 0$ and $-\Xi = \Psi, -\beta/6 = \Phi$. The linearly perturbed Einstein tensor is straightforwardly calculated by inserting the above metric into the definitions of §1.1.3.2. This must be equated to the linearly
perturbed stress-energy tensor, which has the following components:

\[
\begin{align*}
\delta T_0^0 &= -\rho \delta \\
\delta T_i^i &= 3\rho \Pi \\
\delta T_i^j - \frac{1}{3}\delta^i_j \delta T^k_k &= D_i^j [\rho (1 + \omega) \Sigma] \\
\end{align*}
\] (1.28)

Here the fluid variables should be interpreted as a sum over all cosmologically-relevant components, e.g. \(\rho \delta = \sum_a \rho_a \delta_a\) and similar. \(\delta\) denotes the fractional perturbation to the energy density of a fluid: \(\rho = \bar{\rho} (1 + \delta)\), where \(\bar{\rho}\) is the mean energy density. Note that in this thesis I will generally drop the bar and simply use \(\rho\) to denote the zeroth-order part of the energy density. \(\theta\) is the velocity potential, related to the velocity of the fluid by \(v_i = \nabla_i \theta\). \(\Pi = \delta P/\rho\) is a dimensionless pressure perturbation, where \(P\) is a zeroth-order component of the perfect fluid stress-energy tensor defined in eq.(1.19). \(\Sigma\) is the anisotropic stress, sourced by the quadrupole moment of the Boltzmann distribution of a system of particles; it is non-zero only for photons and neutrinos, but these constitute an almost-negligible fraction of the universe at late times. The pressure perturbation can be divided into adiabatic and non-adiabatic contributions. The non-adiabatic contribution is present only when the total effective \(\omega\) varies, and is described by:

\[
\omega \Gamma = \Pi - c_s^2 \delta, \quad c_s^2 = \frac{\dot{P}}{\rho} = \omega - \frac{1}{3H} \frac{\dot{\omega}}{1 + \omega} 
\] (1.30)

where \(c_s^2\) is the sound speed of the dominant fluid. The evolution of the density and velocity perturbations are given by the temporal and spatial components of the linearized version of eq.(1.22), usually referred to as the fluid conservation and Euler equations respectively [163, 213]:

\[
\begin{align*}
\frac{d}{d\eta} \left( \frac{\delta}{1 + \omega} \right) &= - \left( k^2 \theta - 3\dot{\Phi} \right) - \frac{3H \omega \Gamma}{1 + \omega} \\
\dot{\theta} &= -3H(1 - c_s^2) \theta + \frac{c_s^2 \delta}{1 + \omega} + \frac{\omega \Gamma}{1 + \omega} - \frac{2}{3} \omega \Sigma + \Psi
\end{align*}
\] (1.31) (1.32)

Finally, equating the linearized Einstein tensor to the linearized energy-momentum tensor, the linearized Einstein equations for an FRW metric in the
conformal Newtonian gauge are [213, 284]:

\[ 2 \nabla^2 \Phi - 6 \dot{\mathcal{H}} (\dot{\Phi} + \mathcal{H} \Psi) = 8 \pi G_N a^2 \rho \delta \]  
\[ 2 (\dot{\Phi} + \mathcal{H} \Psi) = 8 \pi G_N a^2 \rho (1 + \omega) \theta \]  
\[ 6 \left( \dot{\Phi} + \mathcal{H} \dot{\Psi} + 2 \dot{\mathcal{H}} \Phi + (2 \dot{\mathcal{H}} + \mathcal{H}^2) \Psi \right) - 2 \nabla^2 (\Phi - \Psi) = 24 \pi G_N a^2 \rho \Pi \]  
\[ \Phi - \Psi = 8 \pi G_N a^2 \rho (1 + \omega) \Sigma \]  

The third and fourth equations above originate from the trace of the Einstein equations and their traceless, longitudinal component respectively.

In this thesis I will make frequent use of the Poisson equation, which comes from combining eqs. (1.33) and (1.34):

\[ 2 \nabla^2 \Phi = 8 \pi G_N a^2 \rho \left[ \delta + 3 \mathcal{H} (1 + \omega) \theta \right] \equiv 8 \pi G_N a^2 \rho \Delta \]  

where the last equality defines the gauge-invariant density perturbation \( \Delta \).

1.1.4.2 The Origin of Matter Perturbations

In the standard model of slow-roll, single-field inflation [8, 145, 206] perturbations in the universe originate from quantum fluctuations of a scalar inflaton field. Fluctuations of the inflaton source perturbations in the matter energy density, leading to the growth of potential wells. The power spectrum of gravitational potential wells laid down during this early epoch is [111]:

\[ P_\Phi(k) = \frac{50 \pi^2}{9 k^3} \left( \frac{k}{H_0} \right)^{(n_s-1)} \delta_H^2 \Omega_{M0}^2 \]  

where \( \delta_H \) is the amplitude of scalar perturbations at horizon crossing (\( \sim 4.6 \times 10^{-5} \) for the \( \Lambda \)CDM model [65]) and \( \Omega_{M0} = 8 \pi G_N \rho_{M0}/3 H_0^2 \) is the present energy density fraction in pressureless matter. The quantity \( n_s \) is the spectral index, and it describes deviations from the scale-invariant Harrison-Zel’dovich spectrum of perturbations. In fact, a small deviation from \( n_s = 1 \) is expected even in the simplest models of inflation: the Hubble factor declines slightly during inflation, meaning that the large-\( k \) modes – which exit the horizon last – have a slightly smaller amplitude than the small-\( k \) modes\(^1\). This prediction is supported by re-

\(^1\)This can be seen using the gauge-invariant curvature perturbation \( \zeta = \Phi + \frac{H}{d\phi/dt} \), which has the power spectrum \( P_\zeta(k) = \left( \frac{H^2}{2 \pi \frac{d\phi}{dt}} \right)^2 \). During inflation \( \frac{d\phi}{dt} \) is essentially constant, whilst \( H \) declines a little [155].
cent measurements from the Planck satellite [250] (see §1.2.1.6), which measured $n_s = 0.9603 \pm 0.0073$.

Combining eqs. (1.33), (1.35) and (1.36) and taking the superhorizon limit ($k \ll H$) for a universe dominated by a fluid with a constant equation of state parameter (so $c_s^2 = \omega$), one obtains the following equation for the evolution of the potential [12]:

$$\ddot{\Phi} + 3H(1 + c_s^2)\dot{\Phi} = 0$$ \hspace{1cm} (1.39)

It is clear that $\Phi = \text{constant}$ solves the above equation; this turns out to be the dominant solution for $c_s^2 > -1$ (ie. the second solution is a decaying one). Using this solution in eq. (1.33) on superhorizon scales, using eq. (1.24) and neglecting anisotropic stress (so $\Phi = \Psi$):

$$-6H^2\Phi = 8\pi G_N a^2 \rho \delta = 3H^2\delta$$ \hspace{1cm} (1.40)

$$\Rightarrow \delta = -2\Phi = \text{constant}$$ \hspace{1cm} (1.41)

We see that energy density perturbations also remain frozen outside the horizon.

The behaviour of perturbation modes upon re-entering the horizon after the inflationary epoch depends upon the dominant component of the universe at the time of horizon entry. The mathematical details are given in [111], but in qualitative terms, potential modes that enter the horizon during the radiation era oscillate and decay, whilst those that enter in the matter era remain constant\(^1\). Therefore the transition between the matter and radiation epochs will leave a signature in the matter power spectrum.

Gravitational potentials today can be related to their primordial values just after inflation by the expression:

$$\Phi(k,a) = \frac{9}{10} \Phi_p(k) T(k) \frac{D(a)}{a}$$ \hspace{1cm} (1.42)

$T(k)$ is a *transfer function* that relates the primordial potential to its value after horizon-crossing and matter-radiation equality (as a function of $k$). It has to be calculated numerically; a well-known fitting function is the ‘BBKS’ function [35], whilst a more up-to-date fit is presented in [122]. $D(a)$ is the *growth factor* which describes the evolution of potentials at late times. The factor of $9/10$ is a convention to account for the fact that even the large-scale potentials, normally

\(^1\)There is no contradiction between the constancy of the potentials during an epoch of structure growth. This can be seen using the Poisson equation, eq. (1.37); during the matter era $\rho_M \propto a^{-3}$ and $\Delta_M \propto a$. 

15
constant, decline a little as the universe transitions through matter-radiation equality.

In the subhorizon regime eq.(1.37) can be rewritten as:

$$\delta(k,a) = - \frac{2k^2a}{3\Omega_{M0}H_0^2}\Phi(k,a)$$

(1.43)

Using the expression above and eqs.(1.38) and (1.42) the matter power spectrum can then be written as:

$$P_M(k,a) = \langle |\delta(k,a)|^2 \rangle = 2\pi^2\delta_H^2 \frac{k^n_s}{H_0^{n_s+3}} T(k)^2 D(a)^2$$

(1.44)

A closely-related quantity is the dimensionless power spectrum:

$$\tilde{\Delta}_M(k,a) = \frac{k^3}{2\pi^2} P_M(k,a)$$

(1.45)

Qualitatively speaking, $\tilde{\Delta}_M \sim |\delta M/M|^2$ can be thought of as the power spectrum of mass fluctuations (because multiplying a Fourier mode of the overdensity $\delta \rho/\rho$ by its corresponding ‘box size’ gives the mass contained therein).

On large scales, where the transfer function is approximately constant, the matter power spectrum scales as $P_M \propto k^{n_s}$; on small scales the transfer function is such that the power spectrum decreases as $P_M \propto k^{n_s-4}(\ln k)^2$ [111]. The transition between these two regimes – often called the ‘turnover’ – marks the scale just entering the horizon at the time of matter-radiation equality. Hence it is a probe of the fractional energy densities in matter, radiation, and the non-matter sector. $\Omega_M$ is lower in a universe with $\Omega_\Lambda \neq 0$, meaning that matter-radiation equality happens later than in the $\Omega_\Lambda = 0$ case, pushing the turnover to larger scales. Current galaxy surveys fall just short of being able to measure sufficiently large scales where we expect the turnover to lie. A recent measurement of the matter power spectrum is shown in Fig. 1.1.

It is estimated that pure linear perturbation theory alone is sufficient for wavenumbers up to $k \sim 0.05$ h Mpc$^{-1}$ at low redshifts [292] (where $h = H_0/100$). Wavenumbers in the range $0.05$ h Mpc$^{-1} \gtrsim k \gtrsim 0.15$ h Mpc$^{-1}$ exist in the ‘mildly non-linear’ regime, where (depending somewhat on the application in hand) linear perturbation theory can be augmented by techniques such as the Zel’dovich approximation [316], renormalized perturbation theory (RPT) [88], and the CO-moving Lagrangian Acceleration method (COLA) of [291, 292]. Small scales with $k \gtrsim 0.15$ h$^{-1}$ Mpc fall into the non-linear regime, and so generally require a nu-

---

1Since $\Delta = \delta + 3\Omega/(k(1+\omega)|\vec{v}| \approx \delta$ for $k \gg \Omega$. 
Figure 1.1: The matter power spectrum as measured by the Baryon Oscillation Spectroscopic Survey (BOSS) [7, 94], showing the best-fit model for scales $0.02 \, \text{h} \, \text{Mpc}^{-1} < k < 0.3 \, \text{h} \, \text{Mpc}^{-1}$. Data points beyond the dotted vertical lines were not used to determine the best-fit model. The inset shows the baryon acoustic oscillation scale (see §1.2.2.2); the plot is obtained by dividing both the data and model by an appropriate smooth spectrum with no BAO features. Figure taken from [18].

Numerical simulation for their accurate description, e.g. [173, 202, 272, 313]. Given the accuracies forecast for future galaxy surveys, one must be very careful when using observations that approach the limit of the linear regime – especially if one hopes to disentangle systematic effects from possible new gravitational phenomena. For example, though the baryon acoustic oscillations at $k \sim 0.06 \, \text{h} \, \text{Mpc}^{-1}$ (see §1.2.2.2) are well-described by linear theory, the effects of mode-coupling from the mildly non-linear and non-linear regimes introduce $\sim 1\%$ correction effects.

A further issue which complicates the information contained within the matter power spectrum is that of bias. We will see in the next few subsections that often we are interested in the total matter overdensity, $\delta_M$, which is dominated by the contribution from cold dark matter. However, we can only observe the luminous baryonic component of this overdensity (galaxies). The bias factor relates these
1. INTRODUCTION

quantities as:

$$\delta_{\text{gal}} = b(k, a) \delta_M$$  \hspace{1cm} (1.46)

such that the galaxy power spectrum is related to the matter power spectrum by:

$$P_{\text{gal}}(k) = b(k, a)^2 P_M(k)$$  \hspace{1cm} (1.47)

It is believed that on large scales the bias factor is approximately constant and scale-independent [278], but smaller scales again require more complex models and numerical treatment [48].

1.2 Tools for Testing Gravity

1.2.1 The Cosmic Microwave Background

A successful description of the thermodynamics and microphysics of the Cosmic Microwave Background (CMB) has been established using only the ingredients of the Standard Model of particle physics, cold dark matter, and GR [111, 115, 266, 306]. Most theories of modified gravity are constructed to reduce to GR at early times in the universe, in order to preserve these successes. However, this is not to say that the CMB offers no useful constraints on modified gravity. Because we understand the microphysics that leads to the characteristic length-scales involved in the CMB, the angular scales at which we observe its key features give information about the distance to the redshift surface at which CMB photons were last scattered. Therefore the CMB probes the expansion history of the universe, which is affected by modifications to GR. Furthermore, the largest angular scales of the CMB offer a separate probe of gravitational physics via the Integrated Sachs-Wolfe effect. I will first briefly outline the physics underlying the key CMB observable, the temperature power spectrum. I will then discuss how the CMB peak positions can be used as a probe of gravity and conclude with the current status of observations.

1.2.1.1 Multipoles and Method

Perturbations of a non-relativistic perfect fluid are described by two variables: an energy density perturbation $\delta$, and a velocity perturbation $v_i = \nabla_i \theta$ (defined in §1.1.4.1), along with an equation of state parameter that relates the pressure perturbation to $\delta$. These two variables are equivalent to the first two moments
of the perturbed phase-space distribution for the particles comprising the fluid. Relativistic particles such as photons and neutrinos require higher moments of the perturbed phase-space distribution for their complete description. The perturbed Bose-Einstein distribution (as appropriate for photons) is:

\[ f(\vec{x}, p, \hat{p}, \eta) = \left\{ \exp \left[ \frac{p}{k_B T(\eta)} \left[ 1 + \Theta(\vec{x}, \hat{p}, \eta) \right] \right] - 1 \right\}^{-1} \]

where \( \vec{p} = p\hat{p} \) is the photon momentum and \( \Theta \equiv \delta T/T \) is the fractional temperature perturbation of the photons. The perturbation to \( p \) is neglected because the magnitudes of photon momenta are scarcely altered during Compton interactions. In order to calculate the spectrum of perturbations at the time of last scattering, then, we need to find the quantity \( \Theta \). To do this we first decompose the Fourier-transformed perturbation into a hierarchy of multipole moments\(^1\):

\[ \Theta_\ell(k, \eta) = \frac{1}{(-i)^\ell} \int_{-1}^{1} \frac{1}{2} P_\ell(\mu) \Theta(\mu) d\mu \]

where \( P_\ell \) are the Legendre polynomials and \( \mu = \hat{p} \cdot \hat{k} \). In theory one needs to know the whole hierarchy of multipoles in order to fully characterize the photon distribution. In practice, knowledge of only the lowest-order multipoles is usually sufficient, with the higher-order moments contributing only negligible corrections. This is particularly true on large scales. The higher-\( \ell \) Legendre polynomials oscillate with increasing frequencies; so when considering long-wavelength temperature perturbations these oscillations average to zero, and only the monopole is required.

The calculation of the CMB power spectrum then proceeds in three stages:

1. Solution of the system of linearized gravitational field equations, evolution equations for baryons and cold dark matter, and the Boltzmann hierarchies\(^2\) for the photon and neutrino distributions. Fortunately this complex, coupled system of equations can be greatly simplified in particular limits, see below. For the CMB power spectrum we need to evaluate the solution for the temperature multipoles at the time of last scattering.

2. Calculation of how the primary anisotropies are modified during their propagation towards us from the surface of last scattering. Concretely, we want

\(^1\)For ease of notation I will use the same variable for a quantity and its Fourier transform; the surrounding mathematics should make it clear which is intended.

\(^2\)By Boltzmann hierarchies I mean the coupled system of equations describing the evolution of the photon multipole moments, \( \Theta_\ell(k, \eta) \).
to relate the multipoles we measure today to those determined in step 1.

3. Conversion of the photon temperature multipoles observed today into more relevant observational quantities, i.e. multipoles of the CMB temperature power spectrum (the ‘$C_\ell$ s’).

1.2.1.2 The Multipoles at Last Scattering

We begin with the long-wavelength Fourier modes. As explained above, for horizon-scale temperature perturbations we only need to consider the monopole of the photon temperature distribution. However, the monopole anisotropy that will be observed at late times is not $\Theta_0(k, \eta_*)$ alone, but $\Theta_0(k, \eta_*) + \Psi(k, \eta_*)$, where $\eta_*$ is the conformal time of last photon scattering (see eq.(1.57) and the surrounding text for a more precise definition). The second term above, $\Psi(k, \eta_*)$, accounts for the energy lost as photons climb out of potential wells. Using eqs.(1.41), (1.42) and the initial conditions from inflation, one finds that on large scales the solution of the coupled conservation, Boltzmann and Einstein equations is [111]:

$$\Theta_0(k, \eta_*) + \Psi(k, \eta_*) = -\frac{1}{6} \delta(\eta_*) \quad (1.50)$$

On intermediate scales, to the required degree of accuracy, consideration of the monopole and dipole moments suffices. Higher moments are suppressed by successive powers of $\tau^{-1}$, where $\tau$ is the optical depth in the plasma and is large during the tight-coupling regime. The relevant system of equations can be reduced to the form of a simple harmonic oscillator equation for $\Theta_0$, with the gravitational potentials playing the role of forcing terms:

$$\ddot{\Theta}_0 + \mathcal{H} \frac{R}{1 + R} \dot{\Theta}_0 + k^2 c_s^2 \Theta_0 = -\frac{k^2}{3} \Psi + \mathcal{H} \frac{R}{1 + R} \dot{\Phi} + \dot{\Phi} \quad (1.51)$$

Here $R = 3\rho_b/4\rho_\gamma$ is the baryon-to-photon density ratio (not to be confused with the Ricci scalar) and $c_s = (3[1 + R])^{-\frac{1}{2}}$ is the sound speed in the plasma. Note that the limit $R \rightarrow 0$ recovers the usual result for a radiation-like fluid, $c_s^2 = \omega = 1/3$.

---

1 Numerical codes like CAMB [201] and CMBFAST [274] solve for higher moments too, but all of the essential physics is captured by the monopole and dipole moments.

2 The universe is said to be in the tight-coupling regime when the rate for photons to scatter off electrons is much larger than the expansion rate.
The solution of eq.(1.51) is [166]:

$$\Theta_0(\eta) - \Phi(\eta) = [\Theta_0(0) - \Phi(0)] \cos(kr_s)$$

$$- \frac{k}{\sqrt{3}} \int_0^\eta d\eta' [\Phi(\eta') + \Psi(\eta')] \sin[k(r_s(\eta) - r_s(\eta'))]$$

where $r_s$ is the comoving sound horizon in the plasma (see eq.(1.65) below), and $\Theta_0(0)$ and $\Phi(0)$ are set by the inflationary initial conditions. With this solution for the monopole in hand, the dipole moment can be calculated via the relation (originating from the equations describing Boltzmann hierarchy):

$$\dot{\Theta}_0 + k\Theta_1 = \dot{\Phi}$$

Eq.(1.51) explains why the CMB power spectrum contains oscillatory features, see Fig. 1.2. The first peak corresponds to Fourier modes that enter the horizon at just the right time to have grown to their maximum value when recombination occurs\(^1\). The adjacent trough at slightly smaller scales corresponds to modes that entered the horizon at exactly the right time to have completed one half oscillation by recombination. That is, at the time of recombination they have zero amplitude, so there is a lack of power on these scales in the last scattering surface. This pattern continues: successively higher peaks and troughs mark modes that have completed half-integer or integer oscillations by the time of recombination. Also, from eq.(1.51) we can begin to see why the CMB provides constraints on cosmological parameters: for example, increasing $\Omega_b$ increases $R$ and therefore decreases the sound speed of the plasma. This will affect the frequency of the temperature perturbation oscillations and hence the spacing of the peaks in the CMB.

Finally we come to small scales, where the quadrupole moment of the photon distribution cannot be neglected. The effect of the quadrupole term is to cause Silk damping [280], (sometimes called diffusion damping). On average, photons in the plasma propagate a distance equal to their mean free path; perturbations on scales shorter than this are erased, because their energy is dissipated by the diffusion of the photons. Perturbations on lengthscales longer than the mean free path are preserved, because (on average) photons scatter before they can travel these distances. Hence energy is not transported (i.e. dissipated) over these scales.

Small-scale perturbations entered the cosmological horizon during the radi-\(^1\)This is a simplification; recombination is not an instantaneous event. Furthermore, it spans other transitions such as decoupling and the drag epoch, see §1.2.2.2.
1. INTRODUCTION

Figure 1.2: Temperature power spectrum of the CMB from the Planck experiment. The vertical scale is $\ell(\ell + 1)C_\ell/2\pi$ and the shaded area represents cosmic variance. The solid line is the six-parameter $\Lambda$CDM model that best fits the combination of Planck, WMAP and high-$\ell$ data from ACT and SPT (see references in the text). Figure taken from [250].

...tion era, when radiation pressure hindered their growth. As a result the gravitational potentials have small amplitudes on these scales, so the temperature multipole equations can be simplified by neglecting them. The mathematics of the solution is a little more involved than the previous cases; full details can be found in [111, 166]. The result is that the monopole and dipole moments of the photon distribution both have the form:

$$\Theta_0, \Theta_1 \sim \exp \left[ ik \int c_s d\eta \right] \exp \left[ -\frac{k^2}{k_D^2} \right]$$

(1.54)

where the damping scale $k_D$ is a function of the free electron density, the baryon density, and the Thomson cross-section for photon-electron scattering. The damping exponential in eq.(1.54) suppresses the CMB power spectrum on small scales. The onset scale of the damping offers another way to constrain cosmological parameters, since $k_D \propto \sqrt{\Omega_b}$. This makes sense physically: a universe with more baryons has more free electrons prior to recombination (assuming overall neutrality), so photon-electron scattering occurs more frequently. The mean free path of
a photon is therefore shorter, so perturbations are ‘washed out’ on smaller scales.

1.2.1.3 From Last Scattering to Today

After leaving the surface of last scattering, photons ‘free stream’ until their detection at \( z = 0 \). Since the tight-coupling condition \( \tau \gg 1 \) no longer holds, it is no longer true that higher photon multipoles remain suppressed. The monopole, dipole and quadrupole ‘leak’ into the higher moments. The evolution of the total photon temperature perturbation during free-streaming is described by the Boltzmann equation, which is:

\[
\dot{\Theta} + (ik\mu - \dot{\tau}) \Theta = \dot{\Phi} - i k \mu \Psi - \dot{\tau} [\Theta_0 + \mu v_b]
\]  

(1.55)

Here \( v_b \) is the magnitude of the baryon velocity perturbation, and I have neglected a small polarization term which is not relevant to the present discussion. The solution of this equation was first presented in [166]; here I give only the key results. The photon temperature multipoles today (indicated by the conformal time \( \eta_0 \)) are given by the following expression, which is accurate to approximately 10%:

\[
\Theta_{\ell}(k, \eta_0) \simeq [\Theta_0(k, \eta_*) + \Psi(k, \eta_*)] \, j_\ell [k(\eta_0 - \eta_*)] \\
+ 3\Theta_1(k, \eta_*) \left\{ j_{\ell-1} [k(\eta_0 - \eta_*)] - \frac{(\ell + 1) j_\ell [k(\eta_0 - \eta_*)]}{k(\eta_0 - \eta_*)} \right\} \\
+ \int_0^{\eta_0} d\eta e^{-\tau} \left[ \dot{\Psi}(k, \eta) + \dot{\Phi}(k, \eta) \right] j_\ell [k(\eta_0 - \eta)]
\]  

(1.56)

\( j_\ell(k\eta) \) are the spherical Bessel functions. We see that the temperature multipoles today depend on the monopole and dipole moments at the (conformal) time of last scattering. As discussed previously, events such as recombination and last scattering are not really instantaneous, so \( \eta_0 \) is taken to be the conformal time corresponding to the peak of the visibility function \( g(\eta) \):

\[
g(\eta) = -\dot{\tau} e^{-\tau}
\]  

(1.57)

Loosely speaking, the visibility function can be thought of as a window function describing the duration of last scattering.

The first term in eq.(1.56) is the Sachs-Wolfe term [265]; qualitatively, it describes the temperature the photons have after climbing out of potential wells at the time of recombination. When the final conversion to multipoles of the CMB
1. INTRODUCTION

The power spectrum is performed (see step 3 below), the Sachs-Wolfe contribution is:

\[ \ell (\ell + 1) C^\text{SW}_\ell = \frac{\pi}{2} \left( \frac{\Omega_M}{D(a = 1)} \right)^2 \delta_H^2 \]  

(1.58)

The amplitude \( \delta_H \) and growth factor \( D(a) \) were introduced in §1.1.4.2. Based on eq.(1.58) alone one would expect the CMB power spectrum at large scales should be a constant (one usually plots \( \ell (\ell + 1) C_\ell / 2\pi \)). However, the absolute height of the Sachs-Wolfe plateau cannot be used to constrain cosmological parameters because there is an overall unknown amplitude factor associated with the CMB. Instead one usually thinks of fixing the large-scale amplitude, and measuring the peak heights relative to the plateau.

In fact the accelerating expansion of the universe at late times imprints a further contribution on top of this constant plateau, due to the Integrated Sachs-Wolfe effect (hereafter the ISW effect). Qualitatively, the ISW effect arises as CMB photons propagate through the potential wells of structure forming in the universe. If these potential wells remain constant then the gravitational blueshift a photon acquires as it enters a well is exactly cancelled by the redshift acquired as it leaves, and there is zero net effect. But if the depth of the potential well changes whilst the photon is crossing it there is a net redshift or blueshift; averaged over the whole sky, this leads to an enhancement in the power spectrum of the CMB on angular scales corresponding to those potential wells.

The last term in eq.(1.56) represents the ISW effect. During the matter-dominated phase of a flat, GR-governed universe \( \dot{\Phi} = \dot{\Psi} = 0 \), and \( \Phi = \Psi \). Therefore in the \( \Lambda \)CDM+GR model it is only at late times in the universe, when the non-matter-like sector begins to dominate, that the ISW integral is non-zero (at linear order – there is always a small non-linear ISW effect around over-/underdensities). In the limit \( \ell \gg 1 \) the spherical Bessel function behaves as \( j_\ell(x) \approx 1/\ell (x/\ell)^{(\ell-0.5)} \), which is suppressed for large \( \ell \); hence the ISW integral only affects the largest scales in the CMB. Unfortunately these are the scales at which cosmic variance\(^1\) becomes significant, which makes a measurement of the ISW power spectrum at a statistically significant level difficult to achieve.

\(^1\)Cosmic variance is the fundamental uncertainty imposed by the fact that we can only make a small number of measurements of very low multipoles, because only a few of the corresponding wavelengths ‘fit into’ the sky we measure. Hence we will always suffer the errors of small-number statistics on these scales. Quantitatively, the uncertainty on the \( C_\ell \) is (see next section for definition of the \( C_\ell \)):

\[ \left( \frac{\Delta C_\ell}{C_\ell} \right)_{CV} = \sqrt{\frac{2}{2\ell + 1}} \]
easier to instead detect the ISW effect through its cross-correlation with matter density fluctuations $\delta$ (or in practical terms, galaxy density fluctuations $\delta_g$); the correlation is non-zero because both large-scale structure and the ISW effect trace the gravitational potential.

Finally, I note here that further secondary anisotropies are imprinted on the CMB power spectrum during free-streaming by gravitational lensing; these are discussed in §1.2.3.2.

### 1.2.1.4 Observable Quantities

The last stage of the calculation is to convert the photon temperature perturbation at $z = 0$ into a useful observable. What we really observe is a set of temperatures and directions; the amplitude and directionality of the fluctuations can be separated by decomposing the real-space $\Theta$ into spherical harmonics:

$$\Theta(\vec{x}, \hat{\mathbf{p}}, \eta) = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}(\vec{x}, \eta) Y_{\ell m}(\hat{\mathbf{p}})$$

(1.59)

where $\hat{\mathbf{p}}$ denotes the direction on the sky (equivalently, the unit photon momentum vector), and the spherical harmonics obey standard orthogonality and normalization relations. Eq.(1.59) can be inverted to solve for the amplitudes $a_{\ell m}$:

$$a_{\ell m}(\vec{x}, \eta) = \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k} \cdot \vec{x}} \int d\Omega Y^*_{\ell m}(\hat{\mathbf{p}}) \Theta(\vec{k}, \hat{\mathbf{p}}, \eta)$$

(1.60)

where I have switched to the Fourier-space $\Theta$. In a standard, single-field inflationary scenario, the quantities $a_{\ell m}$ are samples of a Gaussian distribution with zero mean. The particular realization of the CMB anisotropies that we observe carries essentially no information; instead we are interested in the variance of the $a_{\ell m}$, as this tells us about the width of the Gaussian that they are drawn from, which in turn is tied to inflationary physics. The variance of the $a_{\ell m}$ defines the usual CMB power spectrum multipoles:

$$\langle a_{\ell m} a^*_{\ell' m'} \rangle = \delta_{\ell \ell'} \delta_{mm'} C_\ell$$

(1.61)

As mentioned in the previous section, for low multiples our observed sky only contains a small number of samples (since $-\ell \leq m \leq \ell$), so the distribution for the amplitudes of large-scale fluctuations is poorly sampled, leading to the fundamental cosmic variance uncertainty limit.
Finally, the temperature perturbation appearing in eq.(1.60) is converted into multipoles and normalized to the matter density fluctuations (essentially to separate the final expression for the $C_\ell$ into a part depending on the initial conditions arising from inflation, and a part depending on the subsequent growth history):

$$C_\ell = \frac{2}{\pi} \int dk k^2 P_M(k) \left| \frac{\Theta_\ell(k)}{\delta_M(k)} \right|^2$$

(1.62)

The multipoles calculated from eq.(1.56) are substituted into the expression above, along with the matter energy densities calculated as described in §1.1.4.2.

### 1.2.1.5 Peak Positions

The first acoustic peak of the CMB represents the sound horizon of the photon-baryon plasma at the time of decoupling. Denoting its angular size by $\theta_{fp}$, the corresponding multipole in the CMB power spectrum is:

$$\ell_{fp} = \frac{\pi}{\theta_{fp}} = \frac{\pi d_A(z_{\text{dec}})}{r_s(z_{\text{dec}})}$$

(1.63)

where the comoving angular diameter distance to the redshift of decoupling is given by [12]:

$$d_A(z_{\text{dec}}) = \frac{c}{H_0} \frac{1}{\sqrt{\Omega_{M0}}} \left[ \sqrt{\Omega_{M0}} \sinh \left( \sqrt{\Omega_{K0}} \int_{0}^{z_{\text{dec}}} \frac{dz}{E(z)} \right) \right]$$

(1.64)

The quantity in square brackets is referred to as the CMB shift parameter, $\mathcal{R}$, and $E(z) = H(z)/H_0$. The sound horizon $r_s$ at the time of decoupling is:

$$r_s(z_{\text{dec}}) = \frac{c}{\sqrt{3}H_0} \int_{z_{\text{dec}}}^{\infty} \frac{dz}{E(z)\sqrt{1+\mathcal{R}}}$$

(1.65)

It is clear from the expressions above that the position of the first peak is sensitive to the intrinsic curvature of the universe and its matter content (note that the present value of the Hubble parameter cancels from the ratio $r_s/d_A$). The fractional energy density of the non-matter sector also enters these expressions through $E(z)$; its contribution will be irrelevant in the high-redshift integral of eq.(1.65), but not so in eq.(1.64).

We see that in order to use measurements of $\ell_{fp}$ to tell us something about the non-matter sector, we need to know values of other cosmological parameters. If we use only the CMB this presents us with a problem due to the geometric
degeneracy: we can accommodate any value of $\Omega_{K0} = 1 - \Omega_{M0} - \Omega_{R0} - \Omega_{\Lambda 0}$ by adjusting $H_0$ suitably, and with a sufficiently large $\Omega_{K0}$ a model with $\Omega_{\Lambda 0} = 0$ can be made to fit the data. This is demonstrated by the coloured points in Fig. 1.3, which represent fits to data from the Wilkinson Microwave Anisotropy Probe (WMAP) [234].

However, when we combine the CMB with other data – eg. using measurements from the Hubble Key Project [130] to set even a weak prior on $H_0$ – the degeneracy is broken, and models with $\Omega_{\Lambda 0} = 0$ are strongly disfavoured.

1.2.1.6 Current Measurements

In March 2013 the first cosmology results from the European Space Agency’s Planck satellite were released ([247] and accompanying papers). The Planck results are the highest-resolution measurements of the CMB anisotropies to date up to multipoles of $\ell \sim 2000$. The sensitivity of Planck is approximately ten times that of WMAP, and its greater frequency coverage allows for better foreground
Table 1.1: Best-fit values and 68% confidence limits for the base six-parameter ΛCDM model from the Planck satellite. $z_*$ denotes the redshift at which the optical depth reaches unity (used to represent the redshift of the decoupling phase). See [247] for full details.

modelling and removal. For multipoles $\ell \geq 2000$ the most precise measurements come from the Atacama Cosmology Telescope (ACT) [68, 129, 279] and the South Pole Telescope (SPT) [73, 181, 269], which have both measured the range $300 \leq \ell \leq 10,000$.

The CMB temperature power spectrum of Planck is shown in Fig. 1.2. No evidence for extensions to the standard six-parameter ΛCDM model were detected; the best-fit parameter values are shown in Table 1.1.

In addition to results regarding the absence of non-gaussianity (with important implications for inflationary models) [246, 248] and several anomalies at low multipoles, Planck detected gravitational lensing of the CMB at a significance of 25 $\sigma$ [249], see §1.2.3. In [245] the ISW effect was detected via three different methods including, for the first time, cross-correlation with the CMB lensing potential. However, the ISW detection significance reported by the Planck collaboration is lower than other claims in the literature, at $\sim 2.5 \sigma$. The results are consistent with the cosmological constant value of the ΛCDM model, but do not provide useful constraints on $\Omega_\Lambda$. One of the methods employed – stacked aperture photometry of the CMB at the locations of known superstructures (very large clusters and voids) – yielded stronger detections of $\sim 3 - 4 \sigma$ significance. However, the amplitude of the cross-correlation is in very strong tension with the standard cosmological paradigm, being about ten times larger than could maximally be expected in the ΛCDM model. There are concerns that what has been detected
(both by Planck and by [142, 143]) is not genuinely the ISW effect, but is some other poorly-understood correlation or an artefact of incorrect void profiles being used. This issue is yet to be resolved [160, 169, 223].

1.2.2 Galaxy Surveys

1.2.2.1 The Growth Rate and Redshift-Space Distortions

One of the key probes for testing alternative gravity theories is the growth rate of large-scale structure, defined as:

$$f(a) = \frac{d \ln \Delta_M(a)}{d \ln a}$$

(1.66)

where $\Delta_M$ is the gauge-invariant matter density perturbation introduced in eq.(1.37). This makes intuitive sense; if a modification to gravity results in an effective ‘fifth force’ (eg. as occurs in scalar-tensor type theories), this will affect the rate of infall of dark matter towards over-dense regions, and hence the rate at which structure forms. In GR $f(a)$ (or equivalently, $f(z)$) is independent of scale, and its time-dependence is well-described by the approximation [207, 236]:

$$f(z) = \Omega_M(z)$$

(1.67)

where $\Omega_M(z) = 8\pi G_N a^2 \rho_M / 3H^2$ is the fractional energy density of the universe in matter at redshift $z$, and the index $\gamma \approx 6/11$ to a very good approximation in linear theory. In most theories of modified gravity the growth rate acquires a scale-dependence. In chapter 5 I will discuss in detail how the growth rate behaves in alternative gravity theories; for now, I focus on the observational tools used to measure $f(z)$.

The principal tool for measuring the growth rate is redshift-space distortions (RSDs). In most cases the distance to a galaxy is determined by obtaining its redshift (spectroscopically) and using a cosmology-dependent redshift-distance relation. However, there will be a systematic error on distances determined in this way, because a measured galaxy redshift does not result purely from the background expansion rate – it is also affected by the peculiar velocity of the galaxy, ie. its local motion with respect to a comoving observer.

On small scales, such as within the centre of a galaxy cluster, random motions of galaxies dominate. On larger scales there is coherency in the motion of galaxies, as they will tend to all be moving towards the nearest large potential well. These effects are visible in a plot of the correlation function (essentially the real-space...
1. INTRODUCTION

Figure 1.4: The redshift-space correlation function of the 2dF survey [83, 149], plotted as a function of transverse ($\sigma$) and radial ($\pi$) separation between pairs of galaxies. The image is constructed from four mirror-reflections of one quadrant (since the real data has $\sigma, \pi > 0$). In the absence of any RSD effects the contours would be circular. The contours plotted are $\xi = 10, 5, 2, 1, 0.5, 0.2, 0.1$. Figure from [235].

The galaxy correlation function measured by the 2 Degree Field (2dF) survey [83, 149] is plotted in Fig. 1.4. On small scales the high galaxy velocities result in a large spread of inferred distances along the line of sight; this is known as the fingers of god effect. On larger scales the correlation contours appear squashed – qualitatively speaking, we underestimate the distances to galaxies ‘behind’ an overdensity (because they are falling towards us, decreasing their apparent redshift) and overestimate the distances to galaxies ‘in front’ of an overdensity (because their peculiar velocity away from us increases their apparent redshift).

It is these larger scales, of order $10–30 \, h^{-1} \, \text{Mpc}$, from which useful information can be extracted. According to the classic Kaiser description of RSDs [179] the equivalent of the Fourier space matter power spectrum):

$$\xi(r_{ab}) = \langle \delta(r_a) \delta(r_b) \rangle$$  \hspace{1cm} (1.68)
matter overdensities measured in real space \((r)\) and redshift space \((s)\) are related by:

\[
\delta_s(k, z) = \delta_r(k, z) \left(1 + \beta \mu^2\right)
\]  

(1.69)

where the \(\delta_i\) are Fourier variables, \(\mu = (\vec{k} \cdot \vec{r})/kr\) gives the cosine of the angle between the wave vector \(\vec{k}\) and the line of sight, and \(\beta = f/b\). Forming the redshift-space power spectrum and taking angular averages results in [12]:

\[
P_s(k, z) = P_r(k, z) \left(1 + \frac{2}{3} \beta + \frac{1}{5} \beta^2\right)
\]  

(1.70)

In practice, extracting measurements of \(f(z)\) from the expression above is made difficult by uncertainties in the bias and the need to model non-linearities [173, 273]. Some of the sensitivity to bias can be removed by quoting measurements on the density-weighted growth rate, \(f(z)\sigma_8(z)\), where \(\sigma_8(z)\) is the root-mean-square of matter density fluctuations inside a sphere of radius \(8 \, h^{-1}\)Mpc.

Fig. 1.5 shows measurements of the growth rate from several recent surveys. This plot is set to improve in the near future with data from the ongoing BigBOSS [271] and DES [1] experiments, and further still with the planned launch of the ESA Euclid satellite in 2020 [15]. Euclid is designed to measure the growth rate in 14 redshift bins from \(z = 0.6 - 2.1\) with errors of \(1 - 3\%\) [217].

1.2.2.2 Baryon Acoustic Oscillations

At early times in the universe electrons and photons were tightly coupled to one another through Compton scattering. The electrons were further coupled to baryons through the Coulomb interaction, and all three species oscillated together as a plasma. As recombination took place the free electron density dropped until it became insufficient to couple the electrons to the photons; this severs the connection between the baryons and photons too (since baryons are only coupled to photons via the electrons) [122]. The characteristic scale of the plasma oscillations at this time of baryon-photon decoupling – the drag epoch – is left imprinted on the baryon distribution. It is exactly analogous to the first peak of the CMB power spectrum, which represents the horizon scale in the primordial plasma at the time of photon-electron decoupling (note the drag epoch occurs after decoupling, but both decoupling and the drag epoch take place during recombination).

\footnote{The situation here is worth clarifying. Recombination is the era during which free electrons combine with ionized nuclei to form stable atoms. At some point during this extended process}
Figure 1.5: Measurements of the density-weighted growth rate, $f\sigma_8$, from several recent galaxy redshift surveys. The thick solid curve is to the $\Lambda$CDM+GR prediction with WMAP9 parameters, whilst the dashed line uses Planck parameters instead. The dotted, dot-dashed, and dot-dot-dashed lines are theoretical curves from the self-accelerating branch of DGP ($\S$2.3.3) [131, 208], a coupled dark energy model [10, 243], and the Hu-Sawicki $f(R)$ model [165] (see also $\S$5.6.2) respectively. The growth rate predictions for these theories are taken from [105]. Figure from [99].

As the matter overdensities grow this characteristic scale is maintained, and ultimately becomes imprinted on the galaxy distribution. The feature is manifested as an enhancement in the galaxy correlation function (or the power spectrum if working in Fourier space, see Fig. 1.1), and can be used as a 'standard ruler': using the well-understood physics of recombination, we can calculate the physical size of this baryon acoustic oscillation (BAO) scale. Comparing this to the observed (angular) size of the feature at a given redshift then tells us the distance to that redshift. In this way the BAO scale acts as a geometric probe

the Compton rate for photons to scatter off electrons drops below the Hubble expansion rate, so that the majority of photons are decoupled from electrons. However, this does not mean that electrons have decoupled from photons: there are still sufficient photons left in the high-energy tail of the photon energy distribution for them to scatter off (photons greatly outnumber electrons at this time). It is only sometime after photon decoupling that the population of high-energy photons becomes too small for electrons to scatter off photons. This latter stage is the drag epoch.
of the background expansion rate of the universe, which in turn gives information about the effective $\omega$ and the energy density fraction constituted by the non-matter sector.

More concretely then, the sound horizon in the baryon-photon plasma at the drag epoch is given by an expression identical to eq.(1.65), but substituting $z_{\text{drag}}$ for $z_{\text{dec}}$. In principle, if one can accurately measure a the galaxy power spectrum in both the radial and transverse directions, the corresponding radial and transverse BAO features should be separately detectable. These would correspond to an angular scale and a redshift interval respectively, given by [12]:

$$
\theta_{\text{BAO}}(z) = \frac{r_s(z_{\text{drag}})}{(1 + z) d_A(z)} \quad (1.71)
$$

$$
\delta z_{\text{BAO}}(z) = \frac{r_s(z_{\text{drag}}) H(z)}{c} \quad (1.72)
$$

Until recently measurements have not been sufficiently accurate to determine separate radial and transverse power spectra. As a result we have been able to measure only a direction-averaged BAO scale (introduced in [123]):

$$
\left[ \theta_{\text{BAO}}(z)^2 \delta z_{\text{BAO}}(z) \right]^{\frac{1}{2}} = \frac{r_s(z_{\text{drag}})}{[(1 + z)^2 d_A(z)^2 c/H(z)]^{\frac{1}{2}}} = \frac{r_s(z_{\text{drag}})}{D_V(z)} \quad (1.73)
$$

where the last equality above defines the dilation scale $D_V(z)$. The WMAP data give $z_{\text{drag}} \approx 1020$ and $r_s(z_{\text{drag}}) \approx 153$ Mpc [192], whilst the Sloan Digital Sky Survey (SDSS) measured $D_V(0.35) \approx 1370$ Mpc [240]. Combined with measurements of $r_s(z_{\text{drag}})/D_V(0.2) = 0.1980$ and $r_s(z_{\text{drag}})/D_V(0.35) = 0.1094$ from the 2dF survey [239], these results favour the $\Lambda$CDM model over a universe with $\Omega_\Lambda = 0$. When combined with supernovae data, the resulting constraint on the equation of state parameter of the non-matter sector is $-1.098 < \omega < -0.841$ at 2 $\sigma$ confidence.

This situation changed in 2012 with the detection of separate transverse and radial BAO scales, first in the WiggleZ galaxy survey [47] and shortly afterwards confirmed by [18] using the Baryon Oscillation Spectroscopic Survey (BOSS). In this way one obtains independent measurements of $d_A(z)$ and $H(z)$ at a given redshift; both teams found results consistent with the $\Lambda$CDM model. However, the authors of [18] note that this analysis does not significantly improve constraints on cosmological parameters because most of the data used is at low redshifts, where the information contained in $d_A(z)$ and $H(z)$ is degenerate for most cosmological models.
1.2.3 Gravitational Lensing

A disadvantage of the probes described in the previous subsection is that they are sensitive to uncertainties in the bias relation between dark matter and luminous baryons. The phenomenon of gravitational lensing – the deflection of photon trajectories by massive objects – is free of this complication, since the corresponding observables depends upon the total energy density of a region of spacetime (not just the baryonic component)\(^1\). Gravitational lensing has been observed in five different scenarios to date:

- The positions of stars viewed near the solar limb shift during a solar eclipse;
- Microlensing by exoplanets or other small bodies causes a temporary increase in the brightness of stars within our galaxy;
- Strong lensing causes a distant source to be multiply imaged by a lower-redshift lens (e.g. a galaxy cluster can lens a high-redshift galaxy);
- Weak lensing by the large-scale matter distribution of the universe or multiple object-object systems causes tiny but coherent distortions in the observed ellipticities of a field of galaxies;
- Lensing of CMB photons by large-scale structures similarly results in a small distortion of the CMB anisotropies.

I will discuss the last two cases above, as these can be used as probes of gravitational physics on cosmological scales. Time delay measurements between strongly-lensed images of the same object can be used to calibrate the distance-redshift relation; however, this requires a treatment of non-linear scales and hence falls outside of the discussion of the later chapters of this thesis.

1.2.3.1 Cosmic Shear

Consider a perfectly spherical galaxy observed from a large distance. In a flat spacetime it appears as a circle in the plane of the sky. Now pick a point P somewhere in that circle, and imagine a vector connecting the point to the centre of the circle; let us denote this vector \( \theta^s = \{ \theta^s_x, \theta^s_y \} \).

Now we introduce a network of cold dark matter filaments and halos along the line of sight, which deflects a light ray emerging from point P as it propagates towards the observer. In the image plane seen by the observer the ray is detected

\(^1\)This is a somewhat over-idealistic statement. Lensing signals are often cross-correlated with other observables which are subject to bias-related uncertainties.
at point P', connected to the centre of the image by a new vector \( \vec{\theta}^{im} = \{\theta_x^{im}, \theta_y^{im}\} \).

The relationship between the vectors \( \vec{\theta}^s \) and \( \vec{\theta}^{im} \) is given by the transformation matrix [200]:

\[
A_{ij} = \frac{\partial \theta^s_i}{\partial \theta^{im}_j} \equiv \begin{pmatrix}
1 - \kappa - \gamma_1 & -\gamma_2 \\
-\gamma_2 & 1 - \kappa + \gamma_1
\end{pmatrix} \equiv \int_0^\chi \partial_i \partial_j (\Phi + \Psi) \, \tilde{\chi} \left(1 - \frac{\tilde{\chi}}{\chi}\right) \, d\tilde{\chi}
\]

(1.74)

where \( i, j = x, y \) and \( \chi(z) \) is the comoving distance to redshift \( z \). The quantity \( \kappa \) is called the *convergence* and describes the magnification of the original image. The quantities \( \gamma_1, \gamma_2 \) carry information about how the original vector \( \vec{\theta}^s \) has been rotated to a new position in the image plane. When the appropriate distortions are applied to all points in the image, the original circular image is distorted into an ellipse.

One cannot extract useful information from this shearing effect without knowing the shape of the original image (and of course all galaxies are not really perfect spheres). Instead one must measure the shapes of thousands of galaxies and isolate the correlation of their ellipticities. Assuming that elliptical galaxies in our universe are on average randomly oriented (which is not exactly true due to intrinsic alignments, see [187]), the correlations in ellipticities must then be due to the shearing caused by intervening matter structures.

The reason that galaxy weak lensing is a useful probe of gravity theories is the presence of the factor \( \Phi + \Psi \) in eq.(1.74). In a GR-governed matter-dominated epoch this is equal to \( 2\Phi \), since the relation \( \Phi = \Psi \) holds. A very common characteristic of modified gravity theories is that this equality no longer holds true, due to additional terms in the modified field equations. A common parameterization of the Poisson and slip relations (eqs.(1.37) and (1.36)) that we will meet in the following chapters is:

\[
-2k^2\Phi = 8\pi G_N a^2 \mu(k, a) \rho \Delta
\]

\[
\Phi = \gamma(k, a) \Psi
\]

(1.75) 

(1.76)

where \( \mu \) and \( \gamma \) are two functions of time and scale. See §2.4.3, §3.2.2 and [30, 31, 32, 325] for a discussion of this parameterization.

Combining eqs.(1.75) and (1.76) yields (suppressing some function arguments
1. INTRODUCTION

for clarity):

\[-2k^2(\Phi + \Psi) = -2k^2\Phi \left(1 + \frac{1}{\gamma}\right)\]  

\[= 8\pi G_N a^2 \mu \left(1 + \frac{1}{\gamma}\right) \rho \Delta \]  

\[= 8\pi G_N a^2 \Sigma(k, a) \rho \Delta \]  

(1.77)  

(1.78)  

(1.79)

The function \(\Sigma = \mu \left(1 + 1/\gamma\right)\) depends on properties of the background cosmology (which in a modified theory might include new degrees of freedom). Due to an unfortunate convention for notation, this \(\Sigma\) should not be confused with the anisotropic stress defined in eq.(1.29), which has been neglected in the above expressions (because it is vanishingly small during a matter-dominated era). One can therefore constrain the parameterization of eqs.(1.75) and (1.76) by maintaining this \(\Sigma\) factor in eq.(1.74), where the fiducial value is \(\Sigma = 1\). This \(\Sigma\) factor will propagate through into the expression for the power spectrum of the lensing potential, which is one of the final observables output from galaxy weak lensing surveys. For brevity I omit the derivation of the lensing potential power spectrum here, but it can be found in [111, 200].

1.2.3.2 CMB Lensing

As described in §1.2.1.5, the temperature power spectrum of the CMB cannot constrain \(\Omega_{\Lambda 0}\) when used alone. However, CMB photons should be deflected by large-scale matter structures in a manner exactly analogous to cosmic shear. The lensing deflections occur predominantly at low redshifts\(^1\), and imprint small distortions on top of the primary CMB anisotropies. If these small secondary anisotropies can be detected then the CMB becomes a probe of both low- and high-redshift physics, allowing it to break its own geometric degeneracy.

Gravitational lensing of the CMB was first detected at a significance of 3.4 \(\sigma\) by the authors of [92] through cross-correlation of WMAP data with radio sources. Using measurements of the CMB on small angular scales (up to \(\ell \sim 2000\)) from the ACT telescope, the authors of [277] were able to confirm \(\Omega_{\Lambda 0} \neq 0\) at 3.2 \(\sigma\) significance using the CMB alone. Subsequent measurements of CMB lensing from SPT were presented in [301]. The Planck experiment has now detected CMB lensing to extremely high significance, see §1.2.1.6. Fig 1.6 shows the power spectrum of the lensing potential measured by Planck, together with high-\(\ell\) data.

\(^1\)At least, low compared to the redshift of last scattering, \(z_* \sim 1000\). Whether \(z \sim 1 - 2\) qualifies as low redshift is debatable.
If one assumes a cosmological constant, CMB lensing is sensitive to the value of $\Omega_{\Lambda 0}$ through its dependence on the metric potentials (via an expression analogous to eq.(1.74)). Qualitatively, structures would grow more rapidly in a universe with a lower value of $\Omega_{\Lambda 0}$, leading to deeper potential wells and a larger amplitude of the lensing power spectrum. For the purposes of modified gravity (ie. dropping the assumption of a cosmological constant), the effects of a non-trivial relation between the two metric potentials are as discussed in the previous subsection.
1.3 Theoretical Properties of Modified Gravity

1.3.1 Lovelock’s Theorem and the Landscape of Modified Gravity

A celebrated work by Lovelock [211, 212] shows that the Einstein field equations with a cosmological constant term are the only Euler-Lagrange expressions obtainable from a Lagrangian density of the form $\mathcal{L}(g_{\mu\nu}, \partial g_{\mu\nu})$ which satisfy the following properties:

- They are rank (0, 2) tensor equations;
- They are symmetric under the exchange of indices;
- They are divergence-free (implying the conservation of the stress-energy tensor $T_{\mu\nu}$);
- They are second-order in derivatives of the metric (see §1.3.3 for clarification);
- They are local, that is, they involve quantities evaluated only at a spacetime point. A non-local field equation might involve a quantity integrated over a small region of spacetime or inverse derivative operators;
- The equations are formulated in four dimensions.

Formally, Lovelock’s theorem does not uniquely single out the Einstein-Hilbert Lagrangian; the most general possibility is:

$$
\mathcal{L} = \sqrt{-g} \left[ \alpha R - 2\Lambda + \beta \epsilon^{\mu\nu\rho\lambda} R_{\mu\nu}^{\alpha\beta} R_{\alpha\beta\rho\lambda} + \gamma \left( R^2 - 4 R_{\mu\nu} R^{\mu\nu} + R_{\mu\nu\rho\lambda} R^{\mu\nu\rho\lambda} \right) \right] \quad (1.80)
$$

where $\epsilon^{\mu\nu\rho\lambda}$ is the four-dimensional Levi-Civita symbol and $\alpha$, $\beta$ and $\gamma$ are constants. However, the third term in the equation above yields zero contribution when the Euler-Lagrange equations are taken [81]. The fourth term - known as the Gauss-Bonnet term [191, 230] - reduces to a surface term in four dimensions or fewer, i.e. it also does not contribute to the field equations. Therefore eq.(1.80) can be considered as an equivalent action for GR.

The consequence of Lovelock’s theorem is that to build a gravity theory that is different from GR one must implement one or more of the following fundamental changes:

---

1. As discussed in §1.1.3.2, the Einstein-Hilbert Lagrangian contains a term that features a second derivative of the metric, but this part of the Lagrangian does not ultimately contribute to the Einstein field equations.

2. $\epsilon^{\mu\nu\rho\lambda} = +1$ for even permutations of its indices, and is equal to $-1$ for odd permutations.
1. Introduce new degrees of freedom mediating gravitational forces. In the majority of theories this is achieved by introducing new fields into the gravitational action. However, there are a few cases – such as \( f(R) \) gravity (§2.3.1) – in which no new fields appear explicitly. Instead, the action for these cases is a non-Einstein-Hilbert function of only the metric (and related tensors such as \( R_{\mu\nu}, G_{\mu\nu} \), etc.). When the field equations are derived, it is found that more than two degrees of freedom of the metric are propagating, i.e. some modes that are non-dynamical in GR have become dynamical.

2. Consider a number of spacetime dimensions different from four;

3. Construct a theory for which the field equations contain higher than second-order derivatives of the metric, yet is energetically stable (see §1.3.3.2);

4. Employ a non-local Lagrangian;

5. Relinquish the principle of least action.

The last point above constitutes a radical departure from established physical theory – the idea that the equations governing gravitational physics might not be derivable from an action – and will not be discussed further in this thesis. The remaining four options can be used to initiate a classification scheme for theories of modified gravity. Fig.1.7 shows a depiction of how some example theories fit into this ‘landscape’ of theory space. I will now discuss the dominant three corners of this landscape in turn.

### 1.3.2 New Degrees of Freedom

This category is the most populated in current research on modified gravity. New scalar, vector or tensor fields (or any combination thereof) are introduced into the universe. As per our definition of modified gravity in §1.1.2 the new fields couple in a non-minimal way to the metric (if they are minimally coupled to the metric then our definition would classify the set-up as a dark energy model instead). Coupling of the new field(s) to ordinary matter fields is possible but not essential.

Examples involving each type of new field will be presented in §2.3, so I will not go into theory-specific details here. In general, though, one expects the extra field(s) to introduce new kinetic term(s) into the gravitational action (if it/they are dynamical), and there may also be associated mass terms and/or interaction

---

1The existing body of work on non-local theories is comparatively small, so I will not elaborate on them here. See [104] for an example.
Theories have been principally classified according to the way in which they avoid the restrictions of Lovelock’s theorem. I have introduced some further subdivisions for ease of classification. The arrows indicate connections that exist between some theories: for example, \( f(R) \) gravity can be mapped onto a scalar-tensor theory (see §2.6.1), and string compactifications give rise to scalar ‘moduli’ fields in the resulting four-dimensional effective theory.

This creates several opportunities for new non-GR coupling constants and arbitrary functions to appear in the Lagrangian, which can usually be tuned to produce an accelerating universe with an effective equation of state parameter \( \omega_{\text{eff}} \approx -1 \) for the non-matter sector.

In some sense, then, new fields are an easy way to replace the cosmological constant. The major difficulties associated with them are i) creating a theory that is consistent not only with observations of the background expansion rate, but also with observables that depend on the linear perturbations of a gravity theory (see §1.2), and ii) finding a motivated physical explanation for the forms/values of the new functions/ constants of the theory. Without this one is arguably guilty of fine-tuning, i.e. constructing a model specifically to fit the data with little or
no physical basis.

A recent series of papers [38, 39, 54, 138, 144] has developed a new technique for constructing theories of modified gravity containing non-GR field content. The approach is inspired by the effective field theory (hereafter EFT) approach used in particle physics and borrows heavily from work by [79, 275]. In this new technique one first specifies a field content, then writes down a Lagrangian containing all possible terms that are quadratic in perturbations of those fields and respect certain symmetries (e.g., Lorentz invariance). The qualitative argument is that all possible symmetry-preserving terms will be generated as loop diagrams correcting the tree-level action, and hence should be accounted for.

The approach has been dubbed ‘the effective field theory of modified gravity/dark energy’, but this is perhaps something of a misnomer. A true EFT approach would instruct one to write down all possible terms in the Lagrangian that could remain after integrating out modes above some cut-off energy scale $M$. This allows one to organize the Lagrangian terms into a power series in $E/M$, where $E \ll M$ is an energy scale relevant to the system in question. One can then discard terms which are highly suppressed, or treat them as successively small corrections to the tree-level result, etc.

The problem with applying this philosophy to dark energy or modified gravity models that are designed to explain cosmic acceleration is that the natural energy scale of the problem is the Hubble scale. If one considers models containing a single scalar field (for example), and the cut-off scale is taken to be the Planck mass, nearly every term that one could write down is hugely suppressed by factors of $H_0/M_P^1$. The only relevant terms that survive are those of the standard scalar-tensor Lagrangian: a non-minimally coupled Einstein-Hilbert term, a kinetic term for the scalar field, and a potential. Scalar-tensor type theories have been studied extensively, and are already tightly constrained.

To be pedantic, then, it is perhaps better to think of the technique employed in [38, 39, 54, 138, 144] as a tool for writing down perturbed Lagrangians rather than an EFT in the strictest sense. Effectively, this unified perspective on modified gravity Lagrangians is the action-level equivalent of the Parameterized Post-Friedmann formalism that I will describe in Chapter 2. It should be noted that the technique typically results in actions that are extremely complicated but very general – they subsume many other theories as particular limits/choices of the free functions multiplying each term in the action.

\[\text{Note that this is not the case for the EFT of inflation, since the Hubble factor at early times is many orders of magnitude closer to the Planck mass.}\]
1.3.3 Higher Derivatives and Instabilities

An alternative method for evading Lovelock’s theorem is to construct a theory that results in higher-derivative field equations, i.e. field equations containing spatial or temporal derivative(s) of order three or greater. Viable theories of this nature are difficult to find, due to the danger of the Ostrogradski instability described below. However, I wish to emphasize that the Ostrogradski argument is not watertight; there are ways to avoid the instability, but it places tight restrictions on the allowable forms of higher-derivative theories.

A well-known example of a higher-derivative theory is $f(R)$ gravity, in which the Einstein-Hilbert action is extended to be a more general function of $R$. This results in field equations that contain fourth-order derivatives; $f(R)$ gravity will be discussed in more detail in §2.3.1. If one is prepared to sacrifice the principle of Lorentz invariance, it is also possible to construct theories that contain higher-order spatial derivatives but remain second-order in time derivatives, such as Hořava-Lifschitz theory (§2.3.5). In this way one escapes the consequences of both the Lovelock and Ostrogradski theorems; it is a matter of taste whether the loss of Lorentz invariance is a reasonable price to pay for this.

1.3.3.1 The Ostrogradski Instability

All known laws of physics to date – including GR – are at most second-order in time derivatives. In principle one can construct a gravity theory different from GR if time derivatives greater than second-order are permitted in the gravitational field equations. In practice, it is extremely difficult to construct such a theory which is energetically stable, due to a classic result of Lagrangian mechanics known as Ostrogradski’s theorem [231].

First, let us examine the usual second-order case. The discussion here is based upon [312], and will be limited to the classical case. For this subsection only, bold font will be used to denote N-dimensional vectors of configuration space variables or phase space variables.

Consider a Lagrangian that depends on the configuration variables $\mathbf{q} = \{q_i\}$ and their first derivatives, $L(\mathbf{q}, \dot{\mathbf{q}})$ (note that the index $i$ here is used to label the dynamical variables, and should not be confused with a spacetime index). We assume that integration by parts has already been used to reduce the derivative order of the action as far as possible (e.g. rewriting terms $q_i \ddot{q}_i \rightarrow -\dddot{q}_i^2 +$ boundary terms$^4$). The equations of motion for the $q_i$ are obtained using the Euler-Lagrange

\[^4\text{It is assumed that } \delta q_i = \delta \dot{q}_i = 0 \text{ on the boundary, as usual for variational approaches.}\]
\[
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0
\]  \hspace{1cm} (1.81)

If the Lagrangian is \textit{non-degenerate} – meaning that \( \frac{\partial L}{\partial \dot{q}_i} \) depends on \( \dot{q}_i \) – then the resulting equations of motion and their solution have the form:

\[
\ddot{q}_i = f(q, \dot{q}) \quad \Rightarrow \quad q_i(t) = g(t, q_0, \dot{q}_0)
\]  \hspace{1cm} (1.82)

where \( f \) and \( g \) are functions of the configuration-space variables and their derivatives, and boundary conditions are given at time \( t_0 \) by \( q_0 = q(t = t_0) \) and \( \dot{q}_0 = \dot{q}(t = t_0) \). The Hamiltonian is constructed via the usual transformation to phase space:

\[
Q_i = q_i \hspace{1cm} P_i = \frac{\partial L}{\partial \dot{q}_i}
\]  \hspace{1cm} (1.83)

The importance of the Lagrangian being non-degenerate is that it means that the second expression in eq.(1.83) can be inverted to express all the \( \dot{q}_i \) in terms of \( Q_i \) and \( P_i \), i.e. there exists a vector of functions \( v \) such that \( \dot{q}_i = v(Q, P) \). Then we can construct the Hamiltonian:

\[
H(Q, P) = P . v(Q, P) - L(Q, v(Q, P))
\]  \hspace{1cm} (1.84)

The crucial feature of this Hamiltonian is that it depends non-trivially on the vector of conjugate momenta \( P \). The Hamiltonian represents the total energy of a system (except for cases where \( L \) depends explicitly on \( t \)).

Now we repeat these steps for the simplest possible higher-derivative Lagrangian, \( \tilde{L}(q, \dot{q}, \ddot{q}) \). The appropriate Euler-Lagrange equations acquire an additional term:

\[
-\frac{d^2}{dt^2} \left( \frac{\partial \tilde{L}}{\partial \ddot{q}_i} \right) + \frac{d}{dt} \left( \frac{\partial \tilde{L}}{\partial \dot{q}_i} \right) - \frac{\partial \tilde{L}}{\partial q_i} = 0
\]  \hspace{1cm} (1.85)

The condition of non-degeneracy now implies that \( \frac{\partial \tilde{L}}{\partial \ddot{q}_i} \) is not independent of \( \ddot{q}_i \), so the equations of motion and their solution have the form:

\[
q_i^{(4)} = \tilde{f}(q, \dot{q}, \ddot{q}, q^{(3)}) \quad \Rightarrow \quad \ddot{q}_i(t) = \tilde{g}(t, q_0, \dot{q}_0, \ddot{q}_0, q_0^{(3)})
\]  \hspace{1cm} (1.86)
where $q_i^{(n)}$ denotes a $n$th time derivative. The presence of additional boundary conditions in the solution for the $q_i$ means that more phase space variables are needed to describe the system:

$$Q_{1,1} = q_i$$  \hspace{1cm}  \frac{P_{1,1}}{} = \frac{\partial \tilde{L}}{\partial \dot{q}_i} - \frac{d}{dt} \left( \frac{\partial \tilde{L}}{\partial \ddot{q}_i} \right) \tag{1.87}$$

$$Q_{1,2} = q_i$$  \hspace{1cm}  \frac{P_{1,2}}{} = \frac{\partial \tilde{L}}{\partial \ddot{q}_i} \tag{1.88}$$

Again, the condition of non-degeneracy means that the eqs.(1.88) can be inverted to express $\ddot{q}$ in terms of other variables, i.e. the function $\tilde{v}$ exists such that $\ddot{q} = \tilde{v}(Q_1, Q_2, P_2)$. Only three variables are needed in $\tilde{v}$ because $\tilde{L}$ only depends on three variables.

Then the Hamiltonian is:

$$H(Q_1, Q_2, P_1, P_2) = P_1 Q_2 + P_2 \tilde{v}(Q_1, Q_2, P_2) - \tilde{L}(Q_1, Q_2, \tilde{v}(Q_2, Q_1, P_2)) \tag{1.89}$$

The Hamiltonian above displays the typical danger of higher-derivative theories. It has a simple linear dependence on the momentum $P_1$. This means that other modes of the system can be excited to very high energy states, and provided that $P_1$ is made sufficiently negative to compensate, energy will be conserved (recall that $H$ represents the total energy of the system). Nothing prevents two modes of the system running away to (equal and opposite) infinitely high and low energies because – unlike the second-order case – the Hamiltonian of the system is unbounded from below. Assuming there is a unique or small number of ground states for the system, the runaway excitation to high-energy states will be highly entropically favoured.

The arguments continue to hold for even higher-derivative Lagrangians; in fact the instability becomes progressively worse with increasing derivative order (that is, there are more ways for runaway to occur, so it is even more entropically favoured). For a non-degenerate Lagrangian that depends on the $N$th derivative of $q$ there will be $2N$ phase-space variables. The Hamiltonian will be linear in the conjugate momenta $P_1, \ldots, P_{(N-1)}$; only the ‘uppermost’ $P_N$ will be bounded from below.

Given these very general arguments, how is it possible to build a stable higher-derivative theory of gravity at all? The well-known case of $f(R)$ gravity is perhaps the simplest to understand. The index contractions involved in the Ricci scalar $R$ are such that only a single component of the metric is acted upon by
higher derivatives in the field equations. This excites one new degree of freedom, which partners an existing degree of freedom from the Einstein-Hilbert term – the gravitational potential – in the Hamiltonian. The energy associated with the gravitational potential is negative, and the energy associated with the new degree of freedom is positive. The only way for the new degree of freedom to run away to infinitely high energies would be for the gravitational potential to simultaneously decay to infinitely negative energy. With the exception of collapse to a black hole, this cannot occur. Effectively the new degree of freedom is tied down by a constraint relation.

### 1.3.3.2 Ghosts

The term *ghost* is used to refer to a scalar field whose kinetic terms have the opposite sign to the canonical case\(^1\). A ghost contributes negatively to the Hamiltonian of the theory, thus risking a runaway instability as discussed in the previous section. Generally ghosts are regarded as a sickness of a modified gravity theory; for example, the initial excitement caused by the self-accelerating branch of DGP gravity (§2.3.3) subsided somewhat after that branch of the theory was found to be haunted by a ghost [75, 139, 194].

However, other authors have deliberately considered theories based on ghost fields. Careful construction is required to ensure that the Hamiltonian remains bounded from below, despite the negative contribution of the ghost. For example, the action for the ghost condensate model of [148] is:

\[
S = \int d^4x \sqrt{-g} \left[ \frac{M_p^2}{2} R - X + \frac{X^2}{M^4} \right] + S_M
\]  

(1.90)

where \(X = -\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi\) and \(M\) is a mass scale. (Note that, according to our definition, this is formally a dark energy model rather than a theory of modified gravity. It is a member of the k-essence family of models). The effective equation of state parameter of the ghost is:

\[
\omega_\phi = \frac{1 - X/M^4}{1 - 3X/M^4}
\]

(1.91)

Acceleration is possible because \(-1 < \omega_\phi < -1/3\) for \(1/2 < X/M^4 < 2/3\). The \(X^2\) potential term in eq.(1.90) provides the necessary correction terms to ensure the Hamiltonian remains positive.

For clarity, I would like to draw the reader’s attention to the naturally confused

\(^1\)The canonical kinetic term for a scalar field is \(-\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi\).
1. INTRODUCTION

terms ghost (a scalar field with the ‘wrong’-sign kinetic term) and phantom (a field that results in an effective equation of state parameter less than $-1$) [70].

1.3.4 Higher-Dimensional Theories of Gravity

Although GR is formulated in 3 + 1 dimensions, the Riemannian geometry upon which it is based is more general than this. It is therefore possible to formulate gravitational field equations in more (or fewer) than four spacetime dimensions, and argue that our perceived four-dimensional universe arises merely as an effective description. I will discuss only the possibility of additional spatial dimensions. There has been a small amount of work on additional temporal dimensions (eg. [37]), but such models are generally disfavoured due to their potential for harbouring closed timelike curves and the ensuing paradoxes.

Broadly speaking, there are three classes of higher-dimensional theories:

a) Models in which the additional dimension(s) are compactified, that is, curled up to an unobservably small size relative to the three large spatial dimensions.

b) Models in which the additional dimension(s) are large or infinite in extent. Our four-dimensional universe is then referred to as a (mem)brane existing in the higher-dimensional bulk spacetime.

c) Models based on the Gauss-Bonnet term and higher-dimensional generalizations of the Einstein-Hilbert action. Such terms can arise in models involving compact extra dimensions, as described in a) above. However, I have listed them separately here because they do not rely on any assumptions about the size of the new dimensions a priori.

Let us first discuss case a). Such theories were originally considered by Kaluza and Klein [180, 188] in an attempt to unify the forces of gravity and electromagnetism by postulating a fifth dimension that is curled up into a circle. The five-dimensional metric tensor can be decomposed into the usual four-dimensional metric of GR, a gauge field $A_\mu$ and a scalar field [81]:

$$\gamma_{\mu\nu} = e^{2\alpha\phi} g_{\mu\nu} + e^{2\beta\phi} A_\mu A_\nu \quad \gamma_{\mu z} = e^{2\beta\phi} A_\mu \quad \gamma_{zz} = e^{2\beta\phi}$$

where $\alpha$ and $\beta$ are constants and the index $z$ denotes the fifth dimension. One finds that the field equations separate elegantly into GR-like expressions for $g_{\mu\nu}$ and Maxwell-like equations for the gauge field $A_\mu$, hence the superficial promise as a unification theory. However, the appearance of the superfluous scalar field
\( \phi \) led the original Kaluza-Klein theory to be quickly discarded; see [28, 232] for reviews.

The treatment of compact dimensions in this way experienced a revival with the advent of string theory. Instead of the simple circle of Kaluza-Klein theory, the additional six dimensions are compactified on geometrically complex spaces known as Calabi-Yau manifolds. Higher-dimensional analogies of the decomposition in eq.(1.92) give rise to the presence of multiple scalar fields in the four-dimensional effective theory, referred to as moduli. Progress in string theory was initially halted by the lack of a stabilization method for the moduli, as there appeared to be no energy barrier to prevent them from running to infinite field values. As some of the moduli are related to the size of cycles in the Calabi-Yau manifolds, such runaway could lead to the rapid expansion of the compact dimensions and hence render the theory non-viable. It was later found by [177, 178] that introducing higher-dimensional analogues of electromagnetic fields on the Calabi-Yau manifolds – fields known as fluxes – generates a potential for the moduli, keeping them stabilized near their respective minima. An overview of flux compactifications can be found in [141].

In terms of modified gravity, then, the result of compact extra dimensions is generally an exotic form of scalar-tensor theory or quintessence model (see §1.3.2 and §2.3.1). An example is dilatonic quintessence [244], in which the dilaton – a modulus that controls the string coupling constant – is free to propagate, whilst the other moduli remain frozen. Compact dimensions can also be constrained non-gravitationally: particle accelerator experiments, now reaching energies of order TeV, have yet to detect any evidence for their existence. This places an upper limit of \( \sim 10^{-19} \text{m} \) on their size.

In contrast, models involving large extra dimensions are not subject to the constraints from collider experiments. A key feature of such theories is that all Standard Model fields are confined our four-dimensional brane, with only gravity able to propagate into the bulk dimensions. In \( D \) dimensions the gravitational potential should fall off differently from \( 1/r^2 \), a prediction testable by torsion-balance experiments\(^1\). So far no deviation from the standard inverse-square force law has been found when probing scales down to \( \sim 0.1 \text{mm} \).

One of the most well-known braneworld theories, Dvali-Gabadadze-Porrati gravity [75, 116, 126, 131, 268, 272] will be discussed in §2.3.3. Other celebrated examples include the ADD model and the two Randall-Sundrum braneworld models – detailed reviews can be found in [81, 214].

\(^1\)The exact form of the altered fall-off depends on the theory in question. In some scenarios it still has a simple power-law form, whilst in others there are Yukawa-like corrections [6].
1. INTRODUCTION

Let us finally say a few words about category c) above. As stated in §1.3.1, in four dimensions the Einstein tensor and cosmological constant terms exhaust all possible field equations that possess the properties of symmetry, divergencelessness, etc. (see previous discussion). In a higher number of dimensions this is no longer true, because the Gauss-Bonnet term in eq.(1.80) no longer reduces to a topological surface term. For \( D = 5 \) or 6 dimensions the most general action possessing the required properties is:

\[
S_D = \frac{1}{16\pi G_D} \int d^Dx \sqrt{-\gamma} \left( \mathcal{R} - 2\Lambda + \hat{\alpha} \hat{G} \right) + S_M
\]  

(1.93)

where \( \gamma_{ab} \) is the D-dimensional metric \((a, b = 0, \ldots, D - 1)\), \( G_D \) is the D-dimensional Newton’s constant and \( \hat{\alpha} \) is a constant. The resulting field equations are:

\[
S_{ab} + \Lambda \gamma_{ab} + \hat{\alpha} \mathcal{H}_{ab} = 8\pi G_D \mathcal{T}_{ab}
\]

(1.94)

where \( S_{ab} \) is the D-dimensional Einstein tensor, \( \mathcal{T}_{ab} \) is the higher-dimensional stress-energy tensor of matter fields and \( \mathcal{H}_{ab} \) is the Lovelock tensor given by (script \( \mathcal{R} \)s referring to the D-dimensional equivalents of the the Ricci scalar, Ricci tensor and Riemann curvature tensor):

\[
\mathcal{H}_{ab} = 2 \mathcal{R} \mathcal{R}_{ab} - 4 \mathcal{R}_{ac} \mathcal{R}^c_b - 4 \mathcal{R}_{abcd} \mathcal{R}^{cd} + 2 \mathcal{R}_{acde} \mathcal{R}^{cde}_b - \frac{1}{2} \gamma_{ab} \hat{G}
\]

(1.95)

As mentioned above, Gauss-Bonnet theories are primarily of interest as a low-energy limit of a heterotic string action [204]. Performing a perturbative expansion of a typical string action in the inverse string tension \( \mu^{-1} \) and keeping only the two largest terms results in an action of the form of eq.(1.93) (after some field redefinitions). In this scenario the parameter \( \hat{\alpha} \) is related to the string tension as \( \mu^{-1} \sim 2\pi \hat{\alpha} \), and only positive values of \( \alpha \) are considered.

1.3.5 Screening Mechanisms

Modifications to the gravitational action cannot be undertaken lightly; the successes of GR must be preserved. A viable theory must contain some ‘scaling’ effects which make it virtually equivalent to GR in our Solar System whilst permitting modifications on cosmological scales. For example, one could imagine constructing theories that ‘run’ with length scale, energy scale, the ambient energy density of the environment, spacetime curvature, the depth of a potential well, etc. (clearly some of these concepts are closely related). An example would
be the $f(R)$ model of [165]:

$$
S_{HS} = \frac{M_p^2}{2} \int \left[ R - \frac{m^2 c_1 \left( \frac{R}{m^2} \right)^n}{1 + c_2 \left( \frac{R}{m^2} \right)^n} \right] d^4 x + S_M
$$

(1.96)

where the mass scale $m$ is set by $m^2 = 8\pi G_N \rho (a = 1)$, and $c_1$, $c_2$ are constants. At high curvatures ($R \gg m^2$) this action reduces to a cosmological constant for the case $c_1/c_2^2 \to 0$ with $c_1/c_2 = \text{const}$. Thus the $\Lambda$CDM+GR model can be recovered in the early universe and in the Solar System for particular parameter choices.

Over the past several years there has been a surge of interest in phenomena called screening mechanisms [182]. The term refers to a situation in which a gravity theory exhibits some qualitatively new feature that suppresses the modification to GR under certain conditions (as opposed to a more gradual ‘running’ of parameters as described above). At present three screening mechanisms are widely recognized, all based on scalar fields:

1. **The chameleon mechanism.**

Chameleon screening [183, 184] is implemented using a scalar field whose mass depends on the mean energy density of the gravitational system in question. In dense environments such as the Solar System the mass of the scalar is large, making its Compton wavelength extremely small. Its gravitational effects are then too short-range to be detectable. But on cosmological distance scales the mean energy density is much smaller and the scalar is able to propagate, mediating a ‘fifth force’ of gravitational strength. However, recent work by [304] has found that the need to achieve screening of the Sun and/or the Galaxy places strong restrictions on the chameleonic scalar field. Firstly, its Compton wavelength at low redshifts is limited to be of order 1 Mpc, a scale which is in the non-linear regime; therefore such a field would have no impact on the large-scale growth of structure. Secondly, the conformal factor that relates the Einstein and Jordan frames in this theory must be nearly constant over the past Hubble time (assuming that all matter species are universally coupled). This means that chameleonic fields cannot be the cause of the accelerated expansion (essentially, if the Einstein frame requires a cosmological constant or quintessence-like field to accelerate, then so does the Jordan frame).

2. **The symmetron mechanism.**

Like the chameleon mechanism, the symmetron mechanism [152, 153] relies

\[\text{There has also been an initial attempt at implementing a screening mechanism for gravity theories based on vector fields by [42].}\]
on a scalar field with a density-dependent potential. This time it is not the mass of the scalar that varies, but its coupling to matter. In high-density regions the vacuum expectation value (VEV) of the potential is at small field values, with the result that the scalar is uncoupled to matter fields. As the ambient energy density is decreased the potential undergoes a symmetry-breaking phase to a Higgs-like potential, with a non-zero VEV. The coupling of the scalar field to matter is then of gravitational strength, once again resulting in a fifth-force effect.

3. The Vainshtein mechanism.

The Vainshtein mechanism [25, 300] is qualitatively different to the previous two cases. Instead of relying on a special form of potential, a new phenomenon arises due to derivative couplings of the scalar field present in the Lagrangian and field equations (for example, operators like $\nabla_\mu \phi \nabla^\mu \phi \Box \phi$). These terms become large near massive sources. When the action is canonically normalized, this has the effect of suppressing the interactions between the scalar field and matter fields, thereby effectively recovering GR. DGP gravity (§2.3.3), Galileon theories (§2.3.6) [98] and Massive Gravity (for example [26, 27]) all display Vainshtein screening.

By construction these three mechanisms all operate in the regime of non-linear perturbations. Hence they will not be considered further in this thesis, which focuses on the effects of modified gravity on cosmological scales where perturbations are linear.

1.4 Thesis Layout

The structure of this thesis is as follows: in chapter 2 I will present the Parameterized Post-Friedmann formalism (PPF) [31], a framework which unifies the description of linear perturbation theory in most of the modified gravity theories present in current research literature. As well as discussing the assumptions and fundamental principles of the formalism, I will demonstrate its application to a suite of example gravity theories.

Chapters 3 and 4 use simplified versions of the full PPF formalism to explore some of the challenges related to parameterized approaches to modified gravity. In chapter 3 I demonstrate how the form of parameterization used can impact the constraints obtained from cosmological data sets. In chapter 4 I calculate the growth of perturbations during an Einstein-de Sitter epoch using a simplified formalism (a prototype of PPF), both on subhorizon and near-horizon scales. This
chapter should be considered as a trial investigation, which laid the groundwork for the more sophisticated analysis of chapter 5.

The growth of the large-scale structure of the universe is further pursued in chapter 5. There I present a very general and efficient method for calculating the density-weighted growth rate, $f\sigma_8$, that should be applicable to most gravity theories. I use this calculation to forecast the constraints on modified gravity obtainable from next-generation galaxy surveys. The work presented in this chapter can be connected to the PPF formalism (see §5.7), but does not rely upon it.

In chapter 6 I conclude by discussing the regimes in which General Relativity has (and has not) been tested, and the implications of this for alternative gravity theories.
Chapter 2

The Parameterized Post-Friedmann Framework for Theories of Modified Gravity

A unified framework for theories of modified gravity will be an essential tool for interpreting the forthcoming deluge of cosmological data. In this chapter I present a formalism that my collaborators and I have developed [31], the Parameterized Post-Friedmann framework (PPF), which parameterizes the cosmological perturbation theory of a wide variety of modified gravity models. I first explain the basic principles utilized in the construction of PPF, then demonstrate its application to a number of example modified gravity theories. I conclude by discussing some preliminary issues related to constraining the parameters of the formalism with observations, e.g. its simplification in certain regimes and degeneracy between its parameters. These issues will be taken up in more detail in later chapters of this thesis.

2.1 Introduction

Einstein’s theory of General Relativity has been required to defend its title as the true theory of gravitation since its birth. During most of the 20th century modifications to General Relativity (GR) were largely an abstract venture into the realms of mathematical possibility. Today the issue is a more pressing one that we have been forced to entertain by experimental results.

It is unfortunate, then, that our methods of testing GR have become less efficient. Constraining modified theories on an individual basis is likely to be an infinite process, unless our ingenuity at constructing new theories wanes [81]. We
need a fast way to test and rule out theories if we are to drive their population into decline.

As discussed in §1.1.1, an analogous situation existed in the 1970s, when the question at hand was “Is GR the correct description of gravity in the Solar System?” A compelling answer in the affirmative was provided by the Parameterized Post-Newtonian framework (PPN) [296, 308, 311]. In this formalism, competing gravitational theories were mapped onto a unified parameterization and constrained simultaneously using data from laboratory tests, lunar laser ranging and early satellite experiments.

We propose to revive this approach by constructing a parameterized framework that can be used to test the concept of cosmological modified gravity in a very general, model-independent way (we will explain in §6 why the PPN framework itself cannot be applied on cosmological scales, though see [229] for related ideas). A parametric approach is nothing new; there has been a substantial body of work along these lines in recent years [30, 38, 40, 45, 53, 54, 90, 113, 167, 208, 221, 252, 284, 295, 325]. However, most approaches have considered modifications to the field equations of GR that are motivated by simplicity and their relevance to a limited number of cases. Model-builders have moved beyond simple scalar-field theories, and there is need for a parameterization that can handle more sophisticated theories.

In this chapter we present a new formalism called the ‘Parameterized Post-Friedmann’ framework (PPF), which systematically accounts for the limited number of ways in which the Einstein field equations can be modified at the linearized level. This means that it encapsulates a very wide variety of theories, without the use of approximations or numerical solutions. Table 2.1 gives a non-exhaustive list of theories that are covered by the PPF framework.

We note that the name PPF has previously been employed to refer to a different formalism. The phrase ‘Parameterized Post-Friedmann’ was first introduced in [293], then later used by W. Hu and I. Sawicki [164, 167].\footnote{We have chosen to recycle the name yet again here (with kind permission from W. Hu) because we believe an analogy with PPN is a concise and broadly accurate way of describing our formalism. However, we must stress that the analogy with PPN should not be taken to extremes – hopefully the differences between PPN and PPF are made clear in this thesis. We advise the reader to take care which usage of the name PPF is being referred to in other works.} Like the PPF framework presented in this chapter, the formalism of Hu and Sawicki is a parameterization of modified gravitational field equations; it is close in spirit to the phenomenological parameterizations we will describe in chapter 3. The Hu & Sawicki parameterization explicitly divides the universe into three regimes: superhorizon scales, linear/quasistatic scales, and nonlinear scales. It employs...
two free parameters to describe the length-scales of transition between the three regimes. It also introduces three free functions to describe the ratio of the metric potentials and the modified gravitational constant in the horizon-scale and linear regimes separately. It was shown in [167] that by making appropriate choices for the time-dependence of these free functions and interpolating between the three scale regimes, the model-specific behaviour of several $f(R)$ theories and DGP gravity could be reproduced.

The parameterization presented here differs in philosophy from that of Hu & Sawicki (as well as practically, in the number and definition of the free functions used). The PPF framework presented in this chapter does not use ansatzes for the time-dependence of free functions (at least for analytic calculations), or interpolation between different regimes. It also explicitly parameterizes terms arising from perturbations of a new degree of freedom (i.e. terms like $\delta \phi$). This means that one does not need to solve model-specific field equations to check that a particular gravity theory can be represented by the parameterization; any theory subject to the conditions of Table 2.2 is automatically encompassed. Also, we do not a priori implement the boundary conditions that superhorizon curvature perturbations must be conserved, and that all modifications to gravity must be suppressed in the non-linear regime. Whilst we expect this to be the case, this suppression may be the result of non-linear phenomena that do not affect the linearized field equations.

The key feature of our parameterization is that it maintains a direct correspondence between the parameters\(^1\) of the formalism and ‘known’ theories. In this thesis we will refer to a ‘known’ theory as an established model for which field equations can be written down analytically. These are usually derived directly from a covariant action, though knowledge of the action itself is not required for PPF. A known model is represented by a point in the space of all possible theories (or a small region if the theory contains variable parameters). In contrast, we will use the description ‘unknown’ for a point in theory-space for which we do not possess the corresponding action. An unknown theory is characterized purely in terms of its PPF parameters. PPF can be used to make statements about unknown regions of theory-space in addition to the testing of known theories. Such statements could be of use in guiding model-builders to the most relevant regions of theory-space.

Any parameterization has limits of applicability, and PPF is no exception. The (fairly mild) assumptions underlying our formalism are stated in Table 2.2.

\(^1\)We will see in the next section that in an expanding universe one is forced to use time-dependent functions as ‘parameters’, rather than constant numbers.
2. THE PARAMETERIZED POST-FRIEDMANN FRAMEWORK FOR THEORIES OF MODIFIED GRAVITY

### Table 2.1: A non-exhaustive list of theories that are suitable for PPF parameterization. We will not treat all of these explicitly here. \( \mathcal{S} = R^2 - 4R_{\mu\nu}R^{\mu\nu} + R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} \) is the Gauss-Bonnet term.

<table>
<thead>
<tr>
<th>Category</th>
<th>Theory</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horndeski Theories</td>
<td>Scalar-Tensor theory (incl. Brans-Dicke)</td>
<td>[56, 125]</td>
</tr>
<tr>
<td></td>
<td>( f(R) ) gravity</td>
<td>[96, 288]</td>
</tr>
<tr>
<td></td>
<td>( f(G) ) theories</td>
<td>[13, 191, 230]</td>
</tr>
<tr>
<td></td>
<td>Covariant Galileons</td>
<td>[66, 133, 228]</td>
</tr>
<tr>
<td></td>
<td>The Fab Four</td>
<td>[62, 76, 78, 85]</td>
</tr>
<tr>
<td></td>
<td>K-inflation and K-essence</td>
<td>[19, 20]</td>
</tr>
<tr>
<td></td>
<td>Generalized G-inflation</td>
<td>[102, 189]</td>
</tr>
<tr>
<td></td>
<td>Kinetic Gravity Braiding</td>
<td>[101, 186]</td>
</tr>
<tr>
<td></td>
<td>Quintessence (incl. universally coupled models)</td>
<td>[10, 71, 243, 289]</td>
</tr>
<tr>
<td>Lorentz-Violating theories</td>
<td>Effective dark fluid</td>
<td>[163]</td>
</tr>
<tr>
<td>&gt; 2 new degrees of freedom</td>
<td>Einstein-Aether theory</td>
<td>[170, 171, 323, 324]</td>
</tr>
<tr>
<td></td>
<td>Hořava-Lifschitz theory</td>
<td>[161, 287]</td>
</tr>
<tr>
<td></td>
<td>DGP (4D effective theory)</td>
<td>[116, 126]</td>
</tr>
<tr>
<td></td>
<td>EBI gravity</td>
<td>[21, 22, 23, 24, 124]</td>
</tr>
<tr>
<td></td>
<td>TeVeS</td>
<td>[41, 282, 283]</td>
</tr>
</tbody>
</table>

PPN and PPF are highly complementary in their coverage of different accessible gravitational regimes. PPN is restricted to weak-field regimes on scales sufficiently small that linear perturbation theory about the Minkowski metric is an accurate description of the spacetime. Unlike PPN, PPF is valid for arbitrary background metrics (such as the FRW metric) provided that perturbations to the curvature scalar remain small. PPF also assumes the validity of linear perturbation theory, so it is applicable to large length-scales on which matter perturbations have not yet crossed the non-linear threshold (indicated by \( \delta_{\text{M}}(k_nl) \sim 1 \)); note that this boundary evolves with redshift.

Perturbative expansions like PPN and PPF cannot be used in the non-linear, strong-field regime inhabited by compact objects. However, this regime can still be subjected to parameterized tests of gravity via electromagnetic observations [174, 175] and the Parameterized Post-Einsteinian framework (PPE) for gravitational waveforms [86, 315]. Note that despite the similarity in nomenclature, PPE is somewhat different to PPN and PPF, being a parameterization of observables rather than theories themselves.

The purpose of this chapter is to present the formalism that will be used in later chapters of this thesis and future works [32], and demonstrate its use through a number of worked examples.
Table 2.2: A summary of the assumptions underlying the PPF formalism.

<table>
<thead>
<tr>
<th>Assumptions and Restrictions of the PPF Parameterization</th>
</tr>
</thead>
<tbody>
<tr>
<td>All field equations are second-order or lower in time derivatives (but $f(R)$ gravity is still treatable – see text).</td>
</tr>
<tr>
<td>There exists an FRW solution for the background cosmology.†</td>
</tr>
<tr>
<td>There exists a frame in which matter components obey their ordinary conservation equations, that is, $\nabla_\mu T^\mu_\nu = 0$.‡</td>
</tr>
<tr>
<td>The field equations of a gravitational theory are gauge-invariant.</td>
</tr>
<tr>
<td>Nonlinear perturbations are negligible at the length scales under consideration.</td>
</tr>
<tr>
<td>If $N$ non-GR fields are present and $N &gt; 2$, then $N - 2$ relations between the new fields must be specified. If $N &lt; 2$ (the majority of cases) then no additional relations are required.</td>
</tr>
</tbody>
</table>

† This potentially poses a problem for Massive Gravity, in which exact, flat FRW solutions do not exist [89]. However, it may still be possible to map the slightly perturbed background solutions of [302] onto our parameterization.

‡ Models that posit a universal coupling between a quintessence field and matter components are treatable, see §2.2.3.3. Models which implement non-universal couplings are not.

We will use the notation $\kappa = \frac{M_P^2}{\kappa} = 8\pi G$ and set $c = 1$ unless stated otherwise. Our convention for the metric signature is $(-, +, +, +)$. Dots will be used to indicate differentiation with respect to conformal time and hatted variables indicate gauge-invariant combinations, which are formed by adding appropriate metric fluctuations to a perturbed quantity (see §2.2.3). Note that this means $\hat{\chi} \neq \chi$.

2.2 The PPF Formalism

2.2.1 Basic Principles

As stated in the introduction, the PPF framework systematically accounts for allowable extensions to the Einstein field equations, whilst remaining agnostic about their precise form. A particularly important extension is to permit the existence of new scalar degrees of freedom (hereafter d.o.f.) that are not present in GR (due to the need to evade Lovelock’s theorem, see §1.3.1).

We emphasize that scalar d.o.f. are not synonymous with scalar fields, but can also arise from the spin-0 perturbations of vectors and tensors. In addition to these possible new d.o.f., there could also be modifications involving the d.o.f. already present in GR. Therefore we need to allow for perturbations of the spacetime metric to appear in the linearized field equations in a non-standard way.
With these concepts in mind, the construction of the PPF parameterization proceeds as follows: first, we add to the linearized gravitational field equations of GR new terms containing perturbations of all the scalar degrees of freedom that are present. Schematically:

\[
\delta G^\mu_\nu = \kappa \delta T^\mu_\nu + \delta U^\mu_\nu \tag{2.1}
\]

where

\[
\delta U^\mu_\nu = \delta U^{\text{metric}}_\mu_\nu (\delta g_{\rho\sigma}) + \delta U^{\text{d.o.f.}}_\mu_\nu (\delta \phi, V, \nu...) + \delta U^{\text{matter}}_\mu_\nu (\delta M, \theta_M...)
\]

In the expression above the tensor of modifications $\delta U_{\mu\nu}$ has been decomposed into parts containing scalar perturbations of the metric, scalar perturbations of new d.o.f., and matter perturbations. We will see in §2.2.3.3 that the matter term can be eliminated in favour of additional contributions to the first two terms of eq.(2.1). Possible contributions to $\delta U^{\text{d.o.f.}}_{\mu\nu}$ from new vector and tensor fields are indicated – schematically for the present – by $\delta A_i = (\vec{\nabla}_i V)/a$ and $2\nu \gamma_{ij} = -\delta \tilde{g}_{ij}$ respectively, where $\gamma_{ij}$ is a spatial matrix, $A_\mu$ is a new vector field and $\nu$ is a spin-0 perturbation of a new tensor field $\tilde{g}_{\rho\sigma}$. It is necessary to specify the number of new d.o.f. a priori, but one can remain completely general as to their physical origin.

Likewise we also need to choose the derivative order of the parameterization, that is, the highest number of time derivatives that appear in the field equations. We will choose this to be two, as Ostrogradski’s theorem [231] generically leads to the existence of instabilities in higher-derivative theories. Under special circumstances these instabilities can be avoided (see §1.3.3.1), such as occurs in the popular class of $f(R)$ gravity theories. However, any $f(R)$ theory can be mapped onto a second-order scalar-tensor theory via a Legendre transformation [288], so our parameterization is still applicable to $f(R)$ gravity in this form. A reminder of this transformation is given in §2.6.1.

The coefficient of each perturbation appearing in the PPF parameterization can only be a function of zeroth order ‘background’ quantities; for an isotropic and homogeneous universe this means that they are only functions of time and scale (through Fourier wavenumber $k$). The scale-dependence of these coefficient functions is not completely arbitrary – it is related to the derivative order of theory (see §2.4.2). We will find that for many known theories it has a simple polynomial form.

Having accounted for all terms permitted in a second-order theory, we will impose gauge invariance on the parameterization\(^1\). That is to say, we will ensure

\(^1\)In previous works we have referred to the stronger principle of gauge form-invariance, which guarantees the invariance of an equation under any canonical transformation. However, for the
that our parameterized, linearized field equations are explicitly invariant under the coordinate transformation \( x^\mu \rightarrow x^\mu + \xi^\mu \), where \( \xi^\mu \) is an arbitrary four-vector. In principle this is not absolutely necessary; one could consider parameterizing the field equations in a fixed gauge. However, we want to constrain the PPF framework with numerous different observables (e.g. CMB data, weak lensing data, structure growth data etc.), and particularly convenient gauge choices are often associated with each of these. Therefore we choose to maximize the flexibility of our parameterization by writing it in an explicitly gauge-invariant form. We implement this by using gauge-invariant variables as far as possible, and using a gauge-invariance-fixing term to guarantee that any remaining gauge-variant pieces cancel via the background field equations.

One can imagine that the parameterization dictated by the above principles could potentially contain large numbers of free functions. Fortunately these are not all independent; for most cases there exist constraint equations that reduce their number considerably. These are discussed in §2.2.5 below.

Finally we highlight that as the PPF framework is based on the linearized field equations, it is applicable to scale regimes that are well above those affected by screening mechanisms. Effects such as the chameleon, symmetron or Vainshtein mechanisms [61, 152, 153, 182, 183, 184, 300] are not present in the parameterization.

### 2.2.2 Background-Level Parameterization

In eq.(2.1) we introduced a tensor of modifications to the field equations of GR. Let the zeroth-order components of this tensor be \( U^0_i = -X, U^j_i = Y \delta^j_i \), such that the modified Friedmann and Raychaudhuri equations are:

\[
3 \left( \mathcal{H}^2 + K \right) = \kappa a^2 \rho_M + a^2 X \tag{2.2}
\]

\[
2 \left( \mathcal{H}^2 - \dot{\mathcal{H}} + K \right) = \kappa a^2 \rho_M (1 + \omega_M) + a^2 (X + Y) \tag{2.3}
\]

where \( \mathcal{H} \) is the conformal Hubble factor and dots denote derivatives with respect to conformal time. Hereafter we specialize to flat cosmologies, setting \( K = 0 \). \( \rho_M \) denotes a sum over standard matter components, including cold dark matter, with the effective total equation of state \( \omega_M \). Modifications to GR and exotic fluids/dark energy are formally indistinguishable at the unperturbed level, so \( X \) and \( Y \) can be regarded as an effective energy density and pressure respectively.

purposes of linear cosmological perturbation theory, the only really relevant transformation is the one described above.
2. THE PARAMETERIZED POST-FRIEDMANN FRAMEWORK FOR THEORIES OF MODIFIED GRAVITY

2.2.3 Perturbation Variables

2.2.3.1 Metric Variables

We write the perturbed line element as (showing scalar perturbations only):

\[
 ds^2 = a(\eta)^2 \left[ -(1 - 2\Xi)d\eta^2 - 2(\nabla_i \epsilon)d\eta dx^i + \left( 1 + \frac{1}{3}\beta \right) \gamma_{ij} + D_{ij} \nu \right] dx^i dx^j 
\]

where \( \gamma_{ij} \) is a flat spatial metric and \( D_{ij} \) is a derivative operator that projects out the longitudinal, traceless, spatial part of the perturbation:

\[
 D_{ij} = \nabla_i \nabla_j - \frac{1}{3} \gamma_{ij} \nabla^2 \quad (2.5)
\]

Bardeen introduced two gauge-invariant combinations of the metric perturbations [36]:

\[
 \hat{\Phi} = \frac{-1}{6}(\beta - \nabla^2 \nu) + \frac{1}{2} \mathcal{H}(\dot{\nu} + 2\epsilon) \quad (2.6)
\]

\[
 \hat{\Psi} = -\Xi - \frac{1}{2}(\dot{\nu} + 2\dot{\epsilon}) - \frac{1}{2} \mathcal{H}(\dot{\nu} + 2\epsilon) \quad (2.7)
\]

These reduce to the familiar potentials \( \Phi \) and \( \Psi \) of the conformal Newtonian gauge upon setting \( \epsilon = \nu = 0 \) – though beware different sign conventions. For our purposes it will be more convenient to use gauge-invariant variables of the same derivative order (note that \( \hat{\Psi} \) is second-order in time derivatives, whilst \( \hat{\Phi} \) is first-order). For this reason we define the following combination, in which the second-order terms cancel:

\[
 \hat{\Gamma} = \frac{1}{k} \left( \dot{\hat{\Phi}} + \mathcal{H} \dot{\hat{\Psi}} \right) \quad (2.8)
\]

\( \hat{\Phi} \) and \( \hat{\Gamma} \) will be the basic building blocks of the metric sector of our parameterization.

2.2.3.2 New Degrees of Freedom

We also need to introduce a gauge-invariant way of parameterizing the new scalar degrees of freedom. First consider a gauge transformation \( x^\mu \rightarrow x^\mu + \xi^\mu \) generated by the vector \( \xi^\mu = 1/a (\xi^0, \nabla^i \psi) \). The most general way that a dimensionless scalar perturbation \( \chi \) can transform under this co-ordinate shift is:

\[
 \chi \rightarrow \chi + \frac{1}{a} \left( G_1 \xi^0 + G_2 \dot{\xi}^0 + G_3 \psi + G_4 \dot{\psi} \right) \quad (2.9)
\]
Table 2.3: The behaviour of perturbations encountered in this chapter under the gauge transformation $x^\mu \rightarrow x^\mu + \xi^\mu$, where $\xi^\mu = 1/a(\xi^0, \nabla_i \psi)$.

\[ \omega_1 = G_2, \quad \omega_2 = -\frac{1}{6\mathcal{H}}(G_1 + G_4) \]
\[ \omega_3 = G_4, \quad \omega_4 = -\frac{1}{2}(G_3 + \mathcal{H}G_4) - \frac{k^2}{6\mathcal{H}}(G_1 + G_4) \]
In the case of a standard scalar field one obtains $\tilde{\chi} = \delta \phi - \frac{\dot{\phi}}{6\kappa}(\beta + k^2\nu)$. Eqs. (2.10) and (2.11) should be thought of a ‘template’ gauge-invariant variable that can be adapted to play the role of the new d.o.f. in a given theory. Note that if one uses a perturbation of the spacetime metric as an input to this prescription the result vanishes, since $\tilde{\chi}$ represents new d.o.f. only.

### 2.2.3.3 Matter Perturbations

A coupling between new d.o.f. and matter can give rise to an effective evolving and/or scale-dependent gravitational constant, e.g. in a typical scalar-tensor theory $G_{\text{eff}} = G/\phi$. This means that when we write the modified field equations in the form of eq. (2.1) $\delta U_{\mu\nu}^\alpha$ will contain matter variables. For example, consider a case where the modified field equations have the following form:

$$\delta G_{\mu\nu} = \frac{\kappa G}{f(\phi, A_{\mu\nu}, \tilde{g}_{\rho\sigma}, \ldots)} \delta T_{\mu\nu} + \left[ \delta U_{\mu\nu}^{\text{dof}} + \delta U_{\mu\nu}^{\text{metric}} \right]$$  \hspace{1cm} (2.12)

where once more we have indicated that the modifications could arise from new scalar, vector or tensor fields, and some arguments have been suppressed. The renormalized matter terms are taken into $\delta U_{\mu\nu}$:

$$\delta U_{\mu\nu} = \left( \frac{1}{f(\phi, A_{\mu\nu}, \tilde{g}_{\rho\sigma}, \ldots)} - 1 \right) \kappa \delta T_{\mu\nu} + \left[ \delta U_{\mu\nu}^{\text{dof}} + \delta U_{\mu\nu}^{\text{metric}} \right]$$  \hspace{1cm} (2.13)

$\delta T_{\mu\nu}$ can then be eliminated from the expression above using eq. (2.1). Rearranging, we then have $\delta U_{\mu\nu}$ expressed entirely in terms of metric variables and new d.o.f.:

$$\delta U_{\mu\nu} = [1 - f] \delta G_{\mu\nu} + f \left[ \delta U_{\mu\nu}^{\text{dof}} + \delta U_{\mu\nu}^{\text{metric}} \right]$$  \hspace{1cm} (2.14)

A theory that posits a universal coupling between a new d.o.f. and matter species can be transformed to a frame where the coupling is absent. These are not necessarily the standard Einstein and Jordan frames, as the field equations do not have to match GR in either frame. In the uncoupled frame the modification tensor $U_{\mu\nu}$ will be independently conserved, and the PPF parameterization can proceed as normal (we explain why the relation $\nabla_{\mu}U_{\nu}^\alpha = 0$ is necessary in §2.2.5).

In the case of non-universal couplings [10, 243] it may not be possible to find a frame in which the relations $\nabla_{\mu}U_{\nu}^\alpha = 0$ and $\nabla_{\mu}T_{\nu}^\alpha = 0$ hold individually. Our parameterization cannot be applied to these cases in its present formulation.
2.2.4 Perturbation Framework

The components of the perturbed Einstein tensor are:

\[-a^2 \delta G_0^0 = E_\Delta = 2 \nabla^2 \hat{\Phi} - 6 \mathcal{H} k \dot{\Gamma} - 3 \mathcal{H} (3 \mathcal{H}^2 - 3 \dot{\mathcal{H}}) (\dot{\nu} + 2 \epsilon)\]
\[-a^2 \delta G_i^0 = \nabla_i E_\Theta = 2 k \dot{\Gamma} + (3 \mathcal{H}^2 - \dot{\mathcal{H}}) (\dot{\nu} + 2 \epsilon)\]
\[a^2 \delta G_i^i = E_P = 6 k \dot{\Gamma} + 12 \mathcal{H} k \dot{\Gamma} - 2 \nabla^2 (\hat{\Phi} - \hat{\Psi}) - 6 (3 \mathcal{H}^2 - \dot{\mathcal{H}}) \hat{\Psi} + 3 (3 \mathcal{H}^3 + 3 \mathcal{H} \mathcal{H} - 3 \dot{\mathcal{H}}) (\dot{\nu} + 2 \epsilon)\]
\[a^2 \delta \tilde{G}_j^i = D_j^i E_\Sigma = D_j^i (\hat{\Phi} - \hat{\Psi})\]

where

\[\delta \tilde{G}_j^i = \delta G_j^i - \frac{\delta_i^j}{3} \delta G_k^k\]

Recall that \(D_j^i\) projects out the longitudinal, traceless, spatial part of \(\delta G_{\mu \nu}^\mu\); we have defined the quantities \(E_i\) in order to refer to the left-hand side of the Einstein equations with ease. We will also use analogous symbols to refer to the components of the \(\delta U_{\mu \nu}^\mu\) tensor, i.e.

\[U_\Delta = -a^2 \delta U_0^0\]
\[\tilde{\nabla}_i U_\Theta = -a^2 \delta U_i^0\]
\[U_P = a^2 \delta U_i^i\]
\[D_j^i U_\Sigma = a^2 (U_j^i - \frac{1}{3} \delta U_k^k \delta_j^i)\]

Having eliminated any matter perturbations from \(\delta U_{\mu \nu}^\mu\) via the method described in the previous subsection, we can expand \(\delta U_{\mu \nu}^\mu\) in terms of the gauge-invariant variables of eqs. (2.6), (2.8) and (2.10). We write down all terms permitted in a second-order theory. We will see in §2.2.5 that the equation(s) of motion (hereafter e.o.m.) for the new d.o.f. are obtained from the Bianchi identities, and require differentiation of \(\delta U_0^0\) and \(\delta U_i^0\); hence these components can only contain first-order time derivatives if the e.o.m.s are to remain at second order.

For simplicity we will consider the case where only one new scalar d.o.f. is present, but the extension to multiple d.o.f. is straightforward. The PPF parameterization of the modified field equations is then as follows:

\[E_\Delta = \kappa a^2 G \rho_M \delta_M + A_0 k^2 \hat{\Phi} + F_0 k^2 \hat{\Gamma} + \alpha_0 k^2 \dot{\hat{\chi}} + \alpha_1 k \dot{\hat{\chi}} + k^3 M_\Delta (\dot{\nu} + 2 \epsilon)\]
\[E_\Theta = \kappa a^2 G \rho_M (1 + \omega_M) \theta_M + B_0 k \dot{\hat{\Phi}} + I_0 k \dot{\hat{\Gamma}} + \beta_0 k \dot{\hat{\chi}} + \beta_1 \dot{\hat{\chi}} + k^2 M_\Theta (\dot{\nu} + 2 \epsilon)\]
\[ E_P = 3 \kappa a^2 G \rho M \Pi_M + C_0 k^2 \Phi + C_1 k \dot{\Phi} + J_0 k^2 \Gamma + J_1 k \dot{\Gamma} \]
\[ + \gamma_0 k^2 \ddot{\chi} + \gamma_1 k \dot{\chi} + \gamma_2 k^2 \chi + k^3 M_P (\dot{\nu} + 2\epsilon) \] (2.20)
\[ E_S = \kappa a^2 G \rho M (1 + \omega_M) \Sigma_M + D_0 \Phi + \frac{D_1}{k} \dot{\Phi} + K_0 \dot{\Gamma} + \frac{K_1}{k} \dot{\Gamma} + \epsilon_0 \ddot{\chi} + \frac{\epsilon_1}{k} \dot{\chi} + \frac{\epsilon_2}{k^2} \dddot{\chi} \] (2.21)

The coefficients \(A_0 - \epsilon_2\) appearing in these expressions are \emph{not} constants - they are functions of the cosmological background, i.e. functions of time and scale. These dependencies have been suppressed above for the sake of clarity. The factors of \(k\) (Fourier wavenumber) accompanying each term are such that \(A_0 - \epsilon_2\) are all dimensionless. A particular known theory will specify exact functional forms for these coefficients; they can be considered as the PPF equivalent of the ten PPN parameters. They are the ‘slots’ one maps a theory of modified gravity onto.

\(M_{\Delta}, M_{\Theta}\) and \(M_P\) are the gauge-invariance-fixing terms described in §2.2.1, and are similarly functions of background variables. However, the \(M_i\) differ from the coefficients \(A_0 - \epsilon_2\) in that the former are fixed by the zeroth-order field equations, whilst the latter are determined by the linearly perturbed field equations.

Let us demonstrate this explicitly for eq.(2.18). Under gauge transformations of the form:
\[ x^\mu \rightarrow x^\mu + \xi^\mu \]
\[ \xi^\mu = 1/a (\xi^0, \nabla^i \psi) \] (2.22)
eq(2.18) becomes:
\[ E'_{\Delta} - 6\mathcal{H}(\mathcal{H}^2 - \dot{\mathcal{H}}) \frac{\xi^0}{a} = \kappa a^2 \rho_M \left[ \delta'_{M} - 3\mathcal{H}(1 + \omega_M) \frac{\xi^0}{a} \right] + k^3 M_{\Delta} \left[ (\dot{\nu}' + 2\epsilon') + \frac{2\xi^0}{a} \right] \] (2.23)

where primed variables indicate those belonging to the transformed coordinate system. In order for the form of eq.(2.18) to be exactly preserved by the transformation the terms containing \(\xi^0\) must cancel, which determines \(M_{\Delta}\):
\[ M_{\Delta} = \frac{3\mathcal{H}}{2k^3} \left[ 2(\mathcal{H}^2 - \dot{\mathcal{H}}) - \kappa a^2 \rho_M (1 + \omega_M) \right] \]
\[ = \frac{3\mathcal{H}}{2k^3} a^2 (X + Y) \] (2.24)

where eq.(2.3) has been used in reaching the second equality. Analogously one finds:
\[ M_{\Theta} = \frac{1}{2k^2} a^2 (X + Y) \] \[ M_P = \frac{3}{2k^3} a^2 \dot{Y} \] (2.25)
No gauge-invariance-fixing term is needed in eq.(2.21) because all the terms there are individually gauge invariant, including the matter shear perturbation $\Sigma_M$. Note that the $M_i$ are unlikely to be useful discriminators between theories because their evolution with redshift should be broadly similar in theories that reproduce a $\Lambda$CDM-like background. It is in the coefficients $A_0 - \epsilon_2$ that the differences between theories will be manifested.

### 2.2.5 Constraint Equations

The parameterization laid out in eqs.(2.18)-(2.21) contains twenty-two coefficient functions, $A_0 - \epsilon_2$. We will see in the worked examples of §2.3 that these coefficients are often very simply related and not independent. However, the purpose of PPF is to obtain non-model-specific constraints from the data - does this mean that we must run a Markov Chain Monte Carlo analysis with twenty-two free functions? It would be justifiable to object that current and near-future data may not possess sufficient constraining power for such a task.

Fortunately this is not the case for the analyses most relevant to current research. Imposing some restrictions on the types of theories being considered allows one to derive constraint relations between the PPF coefficients, thereby immediately eliminating some freedom from the parameterization. We will present a set of seven constraints relations here; it is very likely that others exist, which we will leave for future investigation.

This set of seven constraint relations stems from the divergenceless nature of the Einstein tensor, $\nabla_\mu G^\mu_\nu = 0$ (a straightforward consequence of the Bianchi identities). We will assume that ordinary matter obeys its standard conservation law, $\nabla_\mu T^\mu_\nu = 0$. Therefore the U-tensor must also be divergenceless, $\nabla_\mu U^\mu_\nu = 0$; the result $\delta (\nabla_\mu U^\mu_\nu) = 0$ follows at the perturbative level. Note that this statement is not valid for models that involve a non-universal coupling between a quintessence field and matter species, see §2.2.3.3.

In an isotropic spacetime the expression $\delta (\nabla_\mu U^\mu_\nu) = 0$ has two independent components (for $\nu = 0, i$) and so yields two second-order equations that specify how the evolution of $\hat{\chi}$ is tied to the metric potentials. However, this situation is potentially problematic - how can we guarantee that the solutions of these two equations agree? In a known theory with a single new field there will be one equation of motion for $\hat{\chi}$. How do we reconcile this fact with the two evolution equations coming from $\delta (\nabla_\mu U^\mu_\nu) = 0$?

There are three possible ways that the problem could be resolved:

1. The true e.o.m. is given by the $\nu = 0$ component of the Bianchi identity.
The $\nu = i$ component reduces to a triviality because all of the coefficients of $\hat{\chi}$, $\hat{\Phi}$ and $\hat{\Gamma}$ in it vanish identically.

2. Instead the converse is true: the $\nu = i$ component becomes the e.o.m., and the $\nu = 0$ component reduces to a triviality.

3. The true e.o.m. corresponds to a combination of the $\nu = 0$, $i$ components.

In this case one set of solutions for $\{\hat{\chi}, \hat{\Phi}, \hat{\Gamma}\}$ would necessarily have to solve both equations.

We will categorize theories as ‘type 1’, ‘type 2’ etc. according to which of the three possibilities above occurs. All of the single-field theories we tackle in this chapter are type 1 theories, except for the special case of §2.3.7 as discussed below. This is what one might intuitively expect - we often think of the e.o.m. of a field as an expression of conservation of its energy, and it is the $\nu = 0$ component of $\nabla_\mu T^\mu_\nu = 0$ that reduces to the energy conservation statement of a classical fluid in Minkowski space.

One can also construct an argument as to why type 1 theories are the most natural occurrence by deriving the conservation law for $U^\mu_\nu$ at the level of the action. In this derivation the statement $\nabla_\mu U^\mu_0 = 0$ is obtained by considering translation along a timelike gauge vector, which corresponds to evolving the new fields along a worldline. The statement $\nabla_\mu U^\mu_i = 0$ arises from translation along the spacelike Killing vectors. As this corresponds to a symmetry of the FRW spacetime, it should not lead us to dynamical equations. See §2.6.2 for more details.

For a type 1 theory, then, the e.o.m. in terms of the parameterization of eqs. (2.18-2.21) is:

$$
\begin{align*}
[\alpha_1 + \mathcal{H}_k \gamma_2] \dddot{\hat{\chi}} + \left[\alpha_0 + \beta_1 + \frac{\dot{\alpha}_1}{k} + \mathcal{H}_k (\alpha_1 + \gamma_1)\right] k \dot{\hat{\chi}} + \left[\frac{\dot{\alpha}_0}{k} + \mathcal{H}_k (\alpha_0 + \gamma_0) + \beta_0\right] k^2 \hat{\chi} \\
+ \left[A_0 + \mathcal{H}_k C_1 - 3 \frac{a^2}{k^2} (X + Y)\right] k \dot{\hat{\Phi}} + \left[A_0 \frac{\dot{k}}{k} + \mathcal{H}_k (A_0 + C_0) + B_0\right] k^2 \dot{\hat{\Phi}} \\
+ \left[F_0 + \mathcal{H}_k J_1\right] k \dot{\hat{\Gamma}} + \left[\frac{\dot{F}_0}{k} + \mathcal{H}_k (F_0 + J_0) + I_0\right] k^2 \dot{\hat{\Gamma}} = 0
\end{align*}
$$

(2.26)

where $\mathcal{H}_k = \mathcal{H}/k$. Meanwhile the seven coefficients of $\{\hat{\Phi}, \dot{\hat{\Phi}}, \hat{\Gamma}, \dot{\hat{\Gamma}}, \hat{\chi}, \dot{\hat{\chi}}, \dddot{\hat{\chi}}\}$ in the second Bianchi component ($\nu = i$) must all vanish identically, leading to the
following set of constraint equations:

\[
\begin{align*}
\beta_1 - \frac{\gamma_2}{3} + \frac{2}{3} \epsilon_2 &= 0 \quad (2.27) \\
\beta_0 + \frac{1}{k} (\dot{\beta}_1 + 2 \mathcal{H} \beta_1) - \frac{1}{3} (\gamma_1 - 2 \epsilon_1) &= 0 \quad (2.28) \\
\frac{\dot{\beta}_0}{k} + 2 \mathcal{H} \beta_0 - \frac{1}{3} (\gamma_0 - 2 \epsilon_0) &= 0 \quad (2.29) \\
B_0 - \frac{1}{3} (C_1 - 2 D_1) + \frac{1}{3 \mathcal{H} k} \alpha^2 (X + Y) &= 0 \quad (2.30) \\
\dot{B}_0 \frac{1}{k} + 2 \mathcal{H} \beta_0 - \frac{1}{3} (C_0 - 2 D_0) &= 0 \quad (2.31) \\
I_0 - \frac{1}{3} (J_1 - 2 K_1) &= 0 \quad (2.32) \\
\dot{I}_0 \frac{1}{k} + 2 \mathcal{H} \beta_0 - \frac{1}{3} (J_0 - 2 K_0) - \frac{1}{3 \mathcal{H} k} \alpha^2 (X + Y) &= 0 \quad (2.33)
\end{align*}
\]

Thus if one is prepared to focus on type 1 theories - which covers a large part of current investigations into modified gravity - then seven free functions are immediately removed from the PPF parameterization.

Theories with more than one new d.o.f. cannot be classified as type 1/2/3. For a theory with precisely two new d.o.f. it is necessary to use both components of the conservation law for \(U_{\mu \nu}\) to obtain the required two evolution equations. However, if one of the evolution equations can be suitably inverted it may be possible to eliminate one of the two extra fields; this will modify the PPF coefficients of the original system. The new system can then be classified as type 1/2/3 if either/neither Bianchi component reduces to a triviality when the new PPF coefficients are used. An example of this is given in §2.3.7.

For theories with more than two new d.o.f. - for example DGP (§2.3.3) and Eddington-Born-Infeld gravity (§2.3.4) - extra relations between the new fields must be provided to close the system of equations. This will prevent us from being able to constrain regions of unknown theory space with more than two new fields. Fortunately, based on the relative scarcity of such theories in the literature - and simplicity arguments - this type of theory is not of primary interest at present.

### 2.3 Worked Examples

In this section we demonstrate how a variety of commonly-discussed theories of modified gravity can be mapped onto our parameterization. We will not discuss the motivation or phenomenology of each theory at length; our intention is to begin compiling a ‘dictionary’ of theories and their translation into PPF format.
We will begin with a simple example involving cosmological scalar fields; we progress to more complicated cases which demonstrate how theories involving timelike vector fields, Lorentz violation and brane scenarios can be encapsulated by the PPF parameterization. We then treat Horndeski theory, itself a powerful parameterization that subsumes a large portion of theory space. We conclude with the example of GR supplemented by an exotic fluid. Whilst conceptually simple, this final example has the unusual property that it can be transformed from a theory of two new fields to a single-field theory of either type 1 or type 2.

2.3.1 Scalar-Tensor Theory and $f(R)$ Gravity

A general scalar-tensor theory has an action of the following form:

$$S_{ST} = \frac{1}{2\kappa} \int \sqrt{-g} d^4x \left[ f(\phi) R - K(\phi) \nabla_\mu \phi \nabla^\mu \phi - 2V(\phi) \right] + S_M(\psi^a, g_{\mu\nu}) \quad (2.34)$$

where $\psi^a$ are matter fields. For convenience (in this section only) we have renormalized the scalar field $\phi \to \phi/M_P$ such that it is dimensionless.

In fact one of the functions $f(\phi)$ and $K(\phi)$ is redundant. Through a redefinition of the scalar field one can always write the above action in a Brans-Dicke-like form, with $f(\phi) = \phi$ and $K(\phi) = \omega(\phi)/\phi$ [125]. We will work with the action in this form.

As explained in §2.2.1 and §2.6.1, a Legendre transformation maps $f(R)$ gravity into to a scalar-tensor theory, with the following equivalences:

$$\phi \equiv \frac{df(R)}{dR} = f_R \quad \omega(\phi) = 0$$

Hence the expressions below can be straightforwardly adapted for use with $f(R)$ models. In fact both theories can be obtained as special cases of the Horndeski Lagrangian (see §2.3.6), but due to their prevalence we will treat them separately here.

The background effective energy density and effective pressure for scalar-tensor theory are (see eqs.(2.2) and (2.3)):

$$a^2 X = 3\dot{H}^2(1 - \phi) + \frac{1}{2} \omega(\phi) \frac{\dot{\phi}^2}{\phi} - 3\dot{\phi} + a^2 V(\phi) \quad (2.36)$$

$$a^2 Y = - (2\dot{\phi} + 3\dot{H})(1 - \phi) + \frac{1}{2} \omega(\phi) \frac{\dot{\phi}^2}{\phi} + \ddot{\phi} + 3\dot{\phi} - a^2 V(\phi) \quad (2.37)$$
The equation of motion for the scalar field is (where matter terms have been eliminated using the gravitational field equations):

\[
\omega(\phi) \left( \frac{\ddot{\phi}}{\phi} - \frac{1}{2} \frac{\dot{\phi}^2}{\phi^2} + 2 \mathcal{H} \frac{\dot{\phi}}{\phi} \right) + \frac{1}{2} \frac{d\omega}{d\phi} \frac{\dot{\phi}^2}{\phi} + a^2 \frac{dV(\phi)}{d\phi} - 3 \left( \ddot{\mathcal{H}} + \mathcal{H}^2 \right) = 0 \quad (2.38)
\]

The gauge-invariant combination for perturbations of the scalar field dictated by eqns.(2.10) and (2.11) is:

\[
\hat{\chi} = \delta \phi - \frac{\dot{\phi}}{6 \mathcal{H}} (\beta + k^2 \nu) \quad (2.39)
\]

To implement the PPF parameterization, we take the linearly perturbed gravitational field equations of scalar-tensor theory and regroup terms to form the gauge-invariant variables \( \hat{\Phi}, \hat{\Gamma} \) and \( \hat{\chi} \). This puts the equations into the format of eqs.(2.18-2.21), from which the PPF coefficient functions can be read off:

\[
\begin{align*}
A_0 &= -2(1 - \phi) + \frac{\dot{\phi}}{\mathcal{H}k^2} \left[ \omega \frac{\dot{\phi}}{\phi} \left( 2 \mathcal{H} + \frac{\dot{\mathcal{H}}}{\mathcal{H}} \right) + k^2 - 6 \mathcal{H} - a^2 \frac{dV(\phi)}{d\phi} \right] + 3 \frac{\dot{\phi}}{k^2} \\
B_0 &= -\frac{1}{\mathcal{H}k} \left[ \dot{\phi} + \dot{\phi} \left( \omega \frac{\dot{\phi}}{\phi} - \mathcal{H} - \frac{\dot{\mathcal{H}}}{\mathcal{H}} \right) \right] \\
C_0 &= 2(1 - \phi) - \frac{3\phi^{(3)}}{k^2 \mathcal{H}} - \frac{3\dot{\phi}}{k^2 \mathcal{H}^2} \left( \mathcal{H}^2 - 2 \dot{\mathcal{H}} \right) \\
&\quad - \frac{3\dot{\phi}}{k^2 \mathcal{H}} \left( 4(\dot{\mathcal{H}} + \mathcal{H}^2) + \frac{2}{3} k^2 - \frac{\omega}{\mathcal{H}} \frac{\dot{\phi}}{\phi} (\dot{\mathcal{H}} + 2\mathcal{H}^2) - \frac{\ddot{\mathcal{H}}}{\mathcal{H}} + \frac{2\mathcal{H}^2}{\mathcal{H}^2} - a^2 \frac{dV(\phi)}{d\phi} \right) \\
C_1 &= \frac{2}{k^2 \mathcal{H}} (1 - \phi) \left( k^2 + 3\mathcal{H}^2 - 3\ddot{\mathcal{H}} \right) - \frac{3\dot{\phi}}{k^2 \mathcal{H}^2} (\mathcal{H}^2 - \dot{\mathcal{H}}) \\
D_0 &= 1 - \phi - \frac{\dot{\phi}}{\mathcal{H}} \\
D_1 &= \frac{k}{\mathcal{H}} (1 - \phi) \\
F_0 &= \frac{6}{k} \left[ \dot{\phi} - \mathcal{H}(1 - \phi) \right] - \frac{\omega \dot{\phi}^2}{k \mathcal{H} \phi} \\
I_0 &= 2(1 - \phi) - \frac{\dot{\phi}}{\mathcal{H}} \\
J_0 &= \frac{2}{\mathcal{H}k} (1 - \phi)(3\ddot{\mathcal{H}} + 3\mathcal{H}^2 - k^2) - 6 \frac{\dot{\phi}}{k \mathcal{H}} - \frac{3\dot{\phi}}{k \mathcal{H}} \left( 2\mathcal{H} - \frac{\dot{\mathcal{H}}}{\mathcal{H}} + \frac{\omega \dot{\phi}}{\phi} \right) \\
J_1 &= 6(1 - \phi) - \frac{3\dot{\phi}}{\mathcal{H}} \\
K_0 &= -\frac{k}{\mathcal{H}} (1 - \phi)
\end{align*}
\]
2. THE PARAMETERIZED POST-FRIEDMANN FRAMEWORK
FOR THEORIES OF MODIFIED GRAVITY

\[ K_1 = 0 \]
\[ \alpha_0 = \frac{1}{k^2} \left[ \frac{1}{2} \frac{d\omega(\phi)}{d\phi} \frac{\dot{\phi}^2}{\phi} - \frac{1}{2} \omega(\phi) \frac{\dot{\phi}^2}{\phi^2} - k^2 - 3H^2 + a^2 \frac{dV(\phi)}{d\phi} \right] \]
\[ \alpha_1 = \frac{1}{k} \left[ \omega(\phi) \frac{\dot{\phi}}{\phi} - 3H \right] \]
\[ \beta_0 = \frac{1}{k} \left[ \omega(\phi) \frac{\dot{\phi}}{\phi} - H \right] \]
\[ \beta_1 = 1 \]
\[ \gamma_0 = \frac{3}{k^2} \left[ \frac{1}{2} \frac{d\omega(\phi)}{d\phi} \frac{\dot{\phi}^2}{\phi} - \frac{1}{2} \omega(\phi) \frac{\dot{\phi}^2}{\phi^2} + H^2 + 2H + \frac{2}{3} k^2 - a^2 \frac{dV(\phi)}{d\phi} \right] \]
\[ \gamma_1 = \frac{3}{k} \left[ H + \omega(\phi) \frac{\dot{\phi}}{\phi} \right] \]
\[ \gamma_2 = 3 \quad \epsilon_0 = 1 \quad \epsilon_1 = \epsilon_2 = 0 \] (2.40)

One can verify by direct substitution that these coefficients obey the constraints of eqs.(2.27 – 2.33). Hence scalar-tensor theories and \( f(R) \) gravity are type 1 theories according to the classification of §2.2.5.

For computations of observables in \( f(R) \) gravity it is useful to define a parameter \( Q \), which has the rough interpretation as the ratio of the Compton wavelength of the ‘scalaron’ field \( f_R \) to the wavelength of a Fourier mode [251]:

\[ Q = 3k^2 \frac{f_{RR}}{f_R} \approx \left( \frac{\lambda_C}{\lambda} \right)^2 \] (2.41)

Making use of the quasistatic approximation (see §2.4.3), the correspondence between our parameterization and \( Q \) is given by:

\[ Q \approx \frac{3}{2} \left[ 1 - \left( 1 + D_0 - \epsilon_0 \frac{B_0}{\beta_0} \right)^{-1} \right] \]
\[ \approx \frac{3}{2} \left[ D_0 - \epsilon_0 \frac{B_0}{\beta_0} \right] \] (2.42)

This result is most easily obtained by mapping both PPF and \( f(R) \) gravity onto the common \( \{ \mu(a,k), \gamma(a,k) \} \) parameterization discussed in §2.4.3.
2.3.2 Einstein-Aether Theory

Let us now give an example where the new d.o.f. do not arise from a scalar field. Modern Einstein-Aether theory was introduced by Jacobson & Mattingly [171] (though in fact the earliest incarnation dates back to Dirac [107, 108, 109, 110]) as a minimalistic model for a theory that violates Lorentz invariance. It achieves this by introducing a dynamical unit vector field, dubbed the ‘aether’, that picks out a preferred spacetime frame. The aether must be time-like at the unperturbed level in order to preserve invariance under spatial rotations, and this requirement is enforced by a Lagrange multiplier in the action:

\[
S_{EA} = \frac{1}{2\kappa^2} \int d^4x \sqrt{-g} \left[ R + M^2\mathcal{K} + \lambda(A^\alpha A_\alpha + 1) \right] + S_M[\psi^a, g_{\mu\nu}] \tag{2.43}
\]

where \( M \) has the dimensions of mass and the kinetic term of the aether field is specified by the constants \( c_i \):

\[
\mathcal{K} = M^{-2}\mathcal{K}^{\alpha\beta}_{\gamma\delta}\nabla_\alpha A^\gamma \nabla_\beta A^\sigma \\
\mathcal{K}^{\alpha\beta}_{\gamma\delta} = c_1 g^{\alpha\beta} g_{\gamma\delta} + c_2 \delta^\alpha_\gamma \delta^\beta_\delta + c_3 \delta^\alpha_\delta \delta^\beta_\gamma \tag{2.44}
\]

For convenience we define \( \alpha = c_1 + 3c_2 + c_3 \). The constraint equation from the Lagrange multiplier is \( g_{\mu\nu} A^\mu A^\nu = -1 \), which leads to \( A^\mu = (1/a, 0, 0, 0) \) in an FRW background.

In our parameterization of the zeroth-order field equations the effective energy density and pressure of the aether are:

\[
a^2 X = \frac{3}{2} \alpha \mathcal{K}^2 \\
a^2 Y = -\frac{1}{2} \alpha \left( 2\dot{\mathcal{K}} + \mathcal{K}^2 \right) \tag{2.45}
\]

The aether experiences perturbations around its zeroth-order direction. We consider only the spin-0 perturbations here, and write the perturbed aether as

\[
A = A^{(0)} + \delta A = \frac{1}{a} (1 + \Xi, \vec{\nabla} V) \tag{2.46}
\]

where the perturbed constraint from the Lagrange multiplier has enforced \( \delta A^0 = \Xi/a \).

The potential \( V \) specifying the perturbation of the spatial aether components represents a new degree of freedom. Following the prescription of §2.2.3.2, it

\[1\]In this thesis we treat only the linear Einstein-Aether theory. In the generalized Einstein-Aether theory \( \mathcal{K} \) is replaced by an arbitrary function \( F(\mathcal{K}) \). This will lead to more complicated PPF coefficients but is not intrinsically problematic.
has the transformation properties $G_4 = -1$, $G_3 = \mathcal{H}$ and $G_1 = G_2 = 0$, so its gauge-invariant partner is:

$$\hat{V} = V + \frac{1}{6\mathcal{H}}(\beta + k^2 \nu) - \epsilon$$

$$= V - \frac{\dot{\Phi}}{\mathcal{H}} + \frac{1}{2} b'$$

(2.47)

$\hat{V}$ has the dimension of a length, so a factor of $k$ is included when defining the dimensionless perturbation appearing in the parameterization of eqs.(2.18–2.21), i.e. $\hat{\chi} \equiv k\hat{V}$.

With the preparations completed, we can now put the linearized field equations of Einstein-Aether theory into the standard format of our parameterization. The PPF coefficients can then be read off:

$$A_0 = c_1 \left(1 - \frac{\ddot{\mathcal{H}}}{\mathcal{H}^2}\right) - \alpha$$

$$B_0 = \frac{k}{\mathcal{H}}(c_1 + c_2 + c_3)$$

$$C_0 = \alpha \left(2 - \frac{\mathcal{H}}{\mathcal{H}^2}\right)$$

$$C_1 = \frac{\alpha}{k^3 \mathcal{H}} \left(k^2 + 3\mathcal{H}^2 - 3\dot{\mathcal{H}}\right)$$

$$D_0 = -(c_1 + c_3) \left(2 - \frac{\mathcal{H}}{\mathcal{H}^2}\right)$$

$$D_1 = -\frac{k}{\mathcal{H}}(c_1 + c_3)$$

$$F_0 = \frac{k}{\mathcal{H}}(c_1 - 3\alpha \mathcal{H}^2)$$

$$I_0 = \alpha$$

$$J_0 = \frac{3\alpha}{k^3 \mathcal{H}} \left(3\mathcal{H}^2 + \dot{\mathcal{H}}\right)$$

$$J_1 = 3\alpha$$

$$K_0 = 0$$

$$K_1 = 0$$

$$\alpha_0 = \mathcal{H}_k (\alpha - c_1)$$

$$\alpha_1 = c_1$$

$$\beta_0 = (c_1 + c_2 + c_3)$$

$$\beta_1 = 0$$

$$\gamma_0 = 2\mathcal{H}_k \alpha$$

$$\gamma_1 = \alpha$$

$$\gamma_2 = 0$$

$$\epsilon_0 = -2\mathcal{H}_k (c_1 + c_3)$$

$$\epsilon_1 = -(c_1 + c_3)$$

$$\epsilon_2 = 0$$

(2.48)

These coefficients satisfy the constraints of eqs.(2.27–2.33), hence Einstein-Aether is a type 1 theory.

### 2.3.3 DGP

Dvali-Gabadadze-Porrati gravity (DGP) [116] is undoubtedly the most commonly-discussed member of the braneworld class of modified gravity theories. It is also
one of the most extensively tested: recent results [255] have placed tight constraints on the normal (non-accelerating) branch of the theory in addition to the existing constraints on the self-accelerating branch [126]. Although increasingly disfavoured, we will present its PPF correspondence here as a representative example of braneworld scenarios [214].

The theory considers our four-dimensional spacetime to be embedded in a five-dimensional bulk, with standard matter fields confined to the 4D brane. The ratio of the effective gravitational constants in the bulk and brane defines a crossover scale, \( r_c = \kappa_5/(2\kappa_4) \) (constrained to be of order the horizon scale by supernovae data), above which gravitational forces are sensitive to the additional dimension. Below the crossover scale the theory can be treated as effectively four-dimensional. However, this is not to say that DGP reduces to GR below the crossover scale; the theory allows a new scalar d.o.f. to propagate\(^1\), and enters a scalar-tensor-like regime before strong coupling of the scalar switches off departures from GR at the Vainshtein scale [300].

Neglecting brane tension, the full action of the theory is:

\[
S_{DGP} = \frac{1}{2\kappa_5} \int d^5 \tilde{x} \sqrt{-\tilde{g}} \left[ \tilde{R} - 2\Lambda_5 \right] \tag{2.49}
+ \int d^4 x \sqrt{-g} \left[ \frac{1}{2\kappa_4} (4) R + \mathcal{L}_M(\psi^a, g_{\mu\nu}) \right]
\]

\( \mathcal{L}_M \) is the Lagrangian of the brane-confined matter fields and \((5)\tilde{R}\) is the Ricci scalar of the 5D spacetime metric \((5)\tilde{g}_{\mu\nu}\). The effective energy density, pressure and equation of state that modify the Friedmann and Raychaudhuri equations are (see eqs.(2.2) and (2.3)):

\[
X = \frac{3\varepsilon}{ar_c} \mathcal{H} \tag{2.50}
\]
\[
Y = -\frac{\varepsilon}{ar_c^2} \left( \dot{\mathcal{H}} + 2\mathcal{H}^2 \right) \tag{2.51}
\]
\[
\omega_E = \frac{Y}{X} = -\frac{1}{3} \left( \frac{\dot{\mathcal{H}}}{\mathcal{H}^2} + 2 \right) \tag{2.52}
\]

\( \varepsilon = +1 \) corresponds to the self-accelerating solution branch, which is asymptotically de Sitter at late times but suffers from ghost pathologies [75, 139]. \( \varepsilon = -1 \) corresponds to the normal solution branch, on which a cosmological constant or dark energy is still required to achieve accelerated expansion.

\(^1\)The new scalar d.o.f. can arise from either a brane-bending mode (if the brane tension is negative), the helicity-0 mode of the graviton (if the brane tension is positive), or a superposition of the two (for zero brane tension) [139, 194].
The new fields in DGP arise from the projection of the electric part of the bulk Weyl tensor onto the brane. It is common to treat the components of the projected Weyl tensor $E_{\mu\nu}$ as an effective ‘Weyl fluid’ with a radiation-like equation of state, $\omega_W = 1/3$. The subdominance of the Weyl terms at late times means that the Weyl fluid is usually defined at the perturbative level, and its zeroth-order components are neglected. The effective Weyl fluid perturbations are:

\[
\begin{align*}
a^2 E_0^0 &= \kappa_4 a^2 \rho_W \delta_W \\
a^2 E_i^0 &= \nabla_i \left[ \kappa_4 a^2 \rho_W (1 + \omega_W) \theta_W \right] \\
a^2 E_j^i &= - \left( \frac{1}{3} \delta^i_j \kappa_4 a^2 \rho_W \delta_W + D^i_j \left[ \kappa_4 a^2 \rho_W (1 + \omega_W) \Sigma_W \right] \right)
\end{align*}
\]  

(2.53)

There are three new fields corresponding to the energy density perturbation, velocity potential and anisotropic stress of the Weyl fluid, so the two components of the four-dimensional $U$-conservation law do not provide us with sufficient equations of motion to solve the system. By considering perturbations of the full five-dimensional spacetime one can relate the three Weyl perturbations to a single d.o.f., the master variable [222]. However, solution of the e.o.m. for the master variable requires gradients perpendicular to the brane that are not encapsulated by the four-dimensional formalism of this thesis. Hence we must stick with three new fields, and thus DGP is not subject to the classification of §2.2.5.

The gauge-invariant Weyl fluid variables are:

\[
\begin{align*}
\hat{\delta}_W &= \delta_W + \frac{1 + \omega_W}{2} (\beta + k^2 \nu) \\
\hat{\theta}_W &= \theta_W - \frac{1}{6H}(\beta + k^2 \nu) \\
\hat{\Sigma}_W &= \Sigma_W
\end{align*}
\]  

(2.54)

The metric PPF coefficients for DGP are:

\[
\begin{align*}
A_0 &= - \frac{3}{r_c^2 X} \left( 1 + \frac{3 \kappa a^2 \rho_W (1 + \omega_W)}{2k^2} \right) \\
B_0 &= \frac{3}{r_c^2 X} \frac{\kappa a^2 \rho_W (1 + \omega_W)}{2kH} \\
C_0 &= \frac{3}{r_c^2 X} \left[ 4 + 3 \omega_E + \frac{3 \kappa a^2 \rho_W (1 + \omega_W)}{2k^2} (2 + 3 \omega_E) \right] \\
C_1 &= \frac{3}{r_c^2 X \tilde{k} k} (k^2 + 3 \tilde{k}^2 - 3 \tilde{t})
\end{align*}
\]
\( D_0 = -\frac{3}{r_c^2(X + 3Y)} \)
\( D_1 = -\frac{k}{9\mathcal{H} r_c^2(X + 3Y)} \)
\( F_0 = -\frac{9\mathcal{H} k}{r_c^2X} \)
\( I_0 = \frac{3}{r_c^2X} \)
\( J_0 = \frac{3}{r_c^2X} \frac{1}{k\mathcal{H}} \left( -k^2 + 3\mathcal{H}^2(4 + 3\omega_E) + 3\dot{\mathcal{H}} \right) \)
\( J_1 = \frac{9}{r_c^2X} \)
\( K_0 = \frac{k}{\mathcal{H} r_c^2(X + 3Y)} \frac{3}{k^2} \)
\( K_1 = 0 \) \hspace{1cm} (2.55)

To incorporate three new fields, in principle the template of eqs.(2.18-2.21) needs to be extended to include two more sets of terms similar to the \( \hat{\chi} \)'s, each with their own coefficient functions. However, it turns out that for DGP most of these possible extra terms are not needed, and the only non-zero \( \hat{\chi} \)-coefficients are:

\[
\begin{align*}
\alpha^0_\delta &= -\frac{3}{2Xr_c^2} \frac{\kappa a^2 \rho_W}{k^2} \\
\beta^k_0 &= \frac{3}{Xr_c^2} \frac{\kappa a^2 \rho_W(1 + \omega_W)}{2k^2} \\
\gamma^0_\delta &= \frac{3}{Xr_c^2} \frac{3\kappa a^2 \rho_W}{2k^2} (2 + 3\omega_E) \\
\epsilon^0_{\kappa^2} &= \frac{3}{r_c^2(X + 3Y)} \frac{\kappa a^2 \rho_W(1 + \omega_W)}{k^2} 
\end{align*}
\] \hspace{1cm} (2.56)

where the superscripts indicate the degree of freedom to which each coefficient belongs, and should not be confused with spacetime indices. Note that, as per the previous subsection, factors of \( k \) have been used where necessary to define dimensionless perturbations of the Weyl fluid.

The authors of [17] reported that a distinctive signature of DGP is the unusual scale-dependence of new terms in the modified Poisson equation. Specifically, they found that DGP led to terms linear in wavenumber \( k \) as the result of a brane-bending mode. PPF is formulated in four dimensions, so off-brane terms such as this that explicitly probe the additional dimension are not present in our coefficient functions. In §2.4.2 we will see that the four-dimensional theory shares the \( k \)-dependence of a scalar-tensor theory.

75
2.3.4 Eddington-Born-Infeld Gravity

Let us give another demonstration of how the PPF parameterization can handle theories beyond scalar-field-type models, this time by considering a bimetric theory. Bimetric theories have been the subject of recent intense interest in the context of Massive Gravity (see [151] for a review); we will instead focus on the related theory of Eddington-Born-Infeld (EBI) gravity, where the coupling between the two metrics is substantially simpler. EBI gravity possesses a second graviton, but unlike Massive Gravity the associated scalar mode propagates and could potentially be a ghost.

EBI gravity arose from the realization of [24] that a theory coupling GR to a Born-Infeld-type Lagrangian could be reformulated as an action for two coupled metrics (a closely-related theory based on a bimetric reformulation of Eddington’s original affine action [118] was explored in [21, 22, 74, 124, 276]). The bimetric re-writing of the EBI action is:

\[
S_{EBI} = \frac{1}{16\pi G} \int d^4x \left[ \sqrt{-g} (R - 2\Lambda) + \sqrt{-\tilde{q}} (\tilde{K} - 2\tilde{\lambda}) - \frac{\sqrt{-\tilde{q}}}{l^2} (\tilde{q}^{-1})^{\mu\nu} g_{\mu\nu} \right] + S_M[\psi^a, g_{\mu\nu}] \tag{2.57}
\]

Matter couples to the usual spacetime metric \(g_{\mu\nu}\) and \(\tilde{K}\) is the curvature of the auxiliary metric \(\tilde{q}_{\mu\nu}\). \(\tilde{\lambda}\) can be considered as a cosmological constant for the auxiliary metric and \(l\) is a length.

For the purpose of the PPF framework we require the cosmological perturbation theory laid out in [23]. First note that, since our conformal time coordinate is defined by the standard spacetime metric, the ‘scale factors’ of the temporal and spatial parts of the auxiliary metric are not constrained to be equal to one another nor the cosmological scale factor \(a(\eta)\). Instead we write the unperturbed auxiliary metric as:

\[
\tilde{q}_{00} = -Z(\eta)^2 \quad \tilde{q}_{ij} = R(\eta)^2 \gamma_{ij} \tag{2.58}
\]

The background-level effective energy density and effective pressure are:

\[
a^2 X = \kappa a^2 \rho_{EBI} = \frac{R^3}{l^2 Z a} \tag{2.59}
\]

\[
a^2 Y = \kappa a^2 P_{EBI} = -\frac{a R Z}{l^2} \tag{2.60}
\]

The scalar perturbations of the auxiliary metric are written in an analogous
manner to those of the spacetime metric in eq.(2.4):

\[
\begin{align*}
\ds^2 &= - Z(\eta)^2 (1 - 2 \tilde{\Xi}) d\eta^2 - 2 R(\eta)^2 (\tilde{\nabla}_i \tilde{\epsilon}) d\eta \, dx + R(\eta)^2 \left[ \left( 1 + \frac{1}{3} \tilde{\beta} \right) \tilde{\gamma}_{ij} + D_{ij} \tilde{\nu} \right] \, dx^i \, dx^j \\
&\quad \text{(2.61)}
\end{align*}
\]

However, we need to be aware that the auxiliary metric perturbations transform in a slightly different way to those of the spacetime metric, due to the possible inequality of \(Z(\eta)\) and \(R(\eta)\). We must use the same gauge transformation vector as for the spacetime metric, \(\xi^\mu = 1/a(\xi^0, \nabla^i \psi)\). The auxiliary perturbations then transform as (using primes to denote the transformed variables):

\[
\begin{align*}
\tilde{\Xi}' &= \tilde{\Xi} - \frac{1}{a} \left[ \frac{\dot{Z}}{Z} - \mathcal{H} \right] \xi^0 + \dot{\xi}^0 \\
\tilde{\epsilon}' &= \tilde{\epsilon} + \frac{1}{a} \left[ \frac{a^2 Z^2}{R^2} \xi^0 + \mathcal{H} \psi - \dot{\psi} \right] \\
\tilde{\beta}' &= \tilde{\beta} + \frac{1}{a} \left[ \frac{6 R}{R} \xi^0 - 2k^2 \psi \right] \\
\tilde{\nu}' &= \tilde{\nu} + \frac{2\psi}{a} \\
&\quad \text{(2.62)}
\end{align*}
\]

These transformations reduce to those of the ordinary spacetime metric perturbations for the case \(Z(\eta) = R(\eta) = a(\eta)\). The four scalar perturbations of the auxiliary metric are the new fields of EBI gravity. Using the now-familiar algorithm of §2.2.3, the corresponding gauge-invariant combinations are:

\[
\begin{align*}
\hat{\Xi} &= \tilde{\Xi} - \Xi - \frac{1}{6 \mathcal{H}} \left( \mathcal{H} - \frac{\dot{Z}}{Z} \right) (\beta + k^2 \nu) \\
\hat{\epsilon} &= \tilde{\epsilon} - \epsilon - \frac{1}{6 \mathcal{H}} \left( \frac{Z^2}{R^2} - 1 \right) (\beta + k^2 \nu) \\
\hat{\beta} &= \tilde{\beta} + k^2 \nu - \frac{\dot{R}}{3 \mathcal{H} R} (\beta + k^2 \nu) \\
\hat{\nu} &= \tilde{\nu} - \nu \\
&\quad \text{(2.63)}
\end{align*}
\]

We put the linearized field equations of the spacetime metric into the standard format of eqs.(2.18-2.21), with each of the gauge-invariant auxiliary perturbations \(\hat{\Xi}, \hat{\epsilon}, \hat{\beta}\) and \(\hat{\nu}\) acting as a new \(\hat{\chi}\)-type field. The PPF coefficients are (where the superscripts on the Greek coefficients indicate to which field they belong):

\[
A_0 = - \frac{R^3}{2 \mathcal{H} k^2 l Z a} \left( 3 \frac{\dot{R}}{R} - \frac{\dot{Z}}{Z} - 2 \mathcal{H} \right)
\]

77
2. THE PARAMETERIZED POST-FRIEDMANN FRAMEWORK FOR THEORIES OF MODIFIED GRAVITY

\[ B_0 = \frac{R^3}{k^2 l^2 Z a} \left( \frac{k}{2} \frac{Z^2}{R^2} - 1 \right) \]
\[ C_0 = \frac{3 Z R a}{\mathcal{H} k^2 l^2} \left( \frac{\dot{R}}{R} + \frac{\dot{Z}}{Z} - 2 \mathcal{H} \right) \]  \hspace{1cm} (2.64)

\[ \alpha^0_{\hat{\nu}} = \frac{R^3}{k^2 l^2 Z a} \]
\[ \alpha^0_{\beta} = \frac{R^3}{2k^2 l^2 Z a} \]
\[ \beta^k_0 = -\frac{R^3}{k^2 l^2 Z a} \]
\[ \gamma^0_{\hat{\nu}} = \frac{3 Z R a}{k^2 l^2} \]
\[ \gamma^0_{\beta} = -\frac{Z R a}{2k^2 l^2} \]  \hspace{1cm} (2.65)

Factors of \( k \) are included where necessary to define dimensionless perturbations to the auxiliary metric. The four additional fields present in EBI gravity means it is not subject to our type 1/2/3 classification. In addition to both components of the \( U \)-conservation law, two further relations are required to evolve the system; these are provided by eqs. (33) and (34) of [23].

2.3.5 Hořava-Lifschitz Theory

Hořava-Lifschitz gravity [161, 287] was proposed as a theory of quantum gravity which achieves power-counting renormalizability at ultra-violet energy scales, whilst potentially recovering GR at low energies. GR is rendered non-renormalizable by the scaling properties of the graviton propagator, but this situation can be remedied by introducing higher-order derivative terms in the gravitational action [185]. However, the price paid for this improved UV behaviour is the appearance of a ghostly (negative energy) degree of freedom. The origin of the ghost can be traced to the presence of higher-order time derivatives in the action.

Hořava’s theory sidesteps this issue by breaking Lorentz invariance and treating time and space unequally. The Hořava-Lifschitz action contains non-linear spatial curvature terms that introduce greater-than-second-order spatial derivatives without introducing higher time derivatives. Consequently the full diffeomorphism group of GR is broken, and only the smaller group of foliation-preserving diffeomorphisms is maintained. As is expected, the breaking of general covariance introduces a new scalar graviton mode; there has been considerable effort made to determine whether this mode propagates or not [49, 77, 135].

After using the ADM formalism to define a preferred time direction, two flavours of Hořava-Lifschitz gravity exist depending on whether the lapse function
is allowed to be a function of both space and time (the non-projectable theory) or time only (the projectable theory). In the non-projectable theory the action picks up extra terms that depend on the shift vector via:

\[ b_i = \partial_i \ln N(t, \vec{x}) \]  

(2.66)

where \( N \) is the lapse function and we avoid the standard ‘\( a_i \)’ notation to prevent confusion with the cosmological scale factor.

We will focus on the non-projectable extension of Hořava’s original action put forward by Blas, Pujolàs and Sibiryakov [50], which has dominated much of the recent discussion. The Blas et al. theory is motivated by the requirement that the new scalar mode possesses a healthy quadratic kinetic term, and also lifts the ‘detailed balance’ condition of Hořava’s original theory. The Blas et al. action is:

\[
S_{HL} = \frac{M_p^2}{2} \int dt d^3x \sqrt{-g} N [\mathcal{L}_K - V(g_{ij}, b_i)] + \int dt d^3x \sqrt{-g} N \mathcal{L}_M(N, N_i, g_{ij})
\]  

(2.67)

where

\[
\mathcal{L}_K = K_{ij}K^{ij} - \lambda K^2
\]

\[
K_{ij} = \frac{1}{2N} \left( \partial_i g_{jj} - \nabla_i N_j - \nabla_j N_i \right)
\]

\[
V(g_{ij}, b_i) = -R - c b_i b^i + M_p^{-2}(d_2 R^2 + d_3 R_{ij}R^{ij} + c_2 b_i \Delta b^i + c_3 R \nabla_i b^i + ...)
\]

\[
+ M_p^{-4}(d_4 R \Delta R + d_5 \nabla_i R_{jk} \nabla^i R^{jk} + c_4 b_i \Delta^2 b^i + c_5 \Delta R \nabla_i b^i + ...).
\]

(2.68)

Ellipses indicate terms not relevant for linear spin-0 perturbations, and \( c_i \) and \( d_i \) are constants. The recovery of GR in the limit \( \lambda = 1 \) is not clear-cut, as the new scalar mode becomes strongly coupled at low energies. Resolution of the issue depends on whether the scalar mode propagates, see the references above.

In a flat FRW spacetime the zeroth-order field equations of Hořava-Lifschitz gravity are identical to those of GR but with a rescaled gravitational constant, \( G_{eff} = 2G_0/(3\lambda - 1) \). In terms of our background-level parameterization this becomes:

\[
a^2 X = \kappa a^2 \rho_M \frac{3(1 - \lambda)}{3\lambda - 1} = 3(1 - \lambda) \frac{3\mathcal{H}^2}{2}
\]  

(2.69)

\[
a^2 Y = \kappa a^2 P_M \frac{3(1 - \lambda)}{3\lambda - 1} = -\frac{3}{2}(1 - \lambda)(2\dot{\mathcal{H}} + \mathcal{H}^2)
\]  

(2.70)

Moving on to parameterization of the linearized field equations, we need to be aware of the reduced diffeomorphism group of Hořava-Lifschitz gravity. The theory is invariant under the gauge transformations \( \eta \rightarrow \tilde{\eta}(\eta), x^i \rightarrow \tilde{x}^i(\eta, \tilde{x}) \), i.e.
we have lost the ability to perform space-dependent reparameterizations of the
time coordinate. However, by making use of the Stückelberg trick we can restore
full general covariance and remove the graviton scalar mode from the metric
sector of the theory; instead we can recast it as a extra field, thereby obtaining
something similar to a scalar-tensor theory.

The Stückelberg trick is implemented by promoting the temporal component
of the gauge transformation vector $\xi^\mu = 1/a(\xi^0, \nabla^i \psi)$ to a new field possessing
the necessary transformation properties to ensure gauge invariance of the whole
equation. Once these transformation properties are determined we can use our
standard procedure of §2.2.3.2 to construct the gauge-invariant variable for the
new Stückelberg field. Note that we are really using a quasi-Stückelberg trick, as
we are leaving the spatial diffeomorphisms unaltered.

Let the dimensionless Stückelberg perturbation be $\chi = k \xi^0$ ($\xi^0$ has the dimen-
sion of length). The accompanying gauge-invariant variable for the Stückelberg
field is:

$$\hat{\chi} = \chi + \frac{\beta + k^2 \nu}{6 \mathcal{H}_k}$$  \hspace{1cm} (2.71)

For convenience we also define the following higher-derivative spatial operators:

$$f_1 = -\left(\frac{2c_3 k^2}{M_P^2 a^2} + \frac{2c_5 k^4}{M_P^4 a^4}\right)$$  \hspace{1cm} (2.72)

$$f_2 = -\left(\frac{c + \frac{c_2 k^2}{M_P^2 a^2} + \frac{c_4 k^4}{M_P^4 a^4}}{2}\right)$$  \hspace{1cm} (2.73)

$$f_3 = -2\left(\frac{(8d_2 + 3d_3) k^2}{M_P^2 a^2} + \frac{(8d_4 - 3d_5) k^4}{M_P^4 a^4}\right)$$  \hspace{1cm} (2.74)

The PPF coefficient functions are found to be:

$$A_0 = 3(1 - \lambda) + f_2\left(1 - \frac{\mathcal{H}}{\mathcal{H}^2}\right) \quad B_0 = \frac{k}{\mathcal{H}}(1 - \lambda)$$

$$C_0 = 3(1 - \lambda)\left(2 - \frac{\mathcal{H}}{\mathcal{H}^2}\right) - f_1\left(1 - \frac{\mathcal{H}}{\mathcal{H}^2}\right) \quad C_1 = \frac{3}{\mathcal{H}k} (1 - \lambda) \left[3(\mathcal{H}^2 - \mathcal{H}) + k^2\right]$$

$$D_0 = -\frac{f_1}{2}\left(1 - \frac{\mathcal{H}}{\mathcal{H}^2}\right) \quad D_1 = 0$$

$$F_0 = 9\mathcal{H}_k(1 - \lambda) + \frac{k}{\mathcal{H}}f_2 \quad I_0 = 3(1 - \lambda)$$

$$J_0 = \frac{1}{\mathcal{H}k}\left[9(1 - \lambda)(\mathcal{H}^2 + \mathcal{H}) - k^2 f_1\right] \quad J_1 = 9(1 - \lambda)$$

$$K_0 = -\frac{k f_1}{\mathcal{H}^2} \quad K_1 = 0$$
\[ \alpha_0 = \mathcal{H}_k [3(\lambda - 1) + f_1 + f_2] \quad \alpha_1 = f_2 \]
\[ \beta_0 = 1 - \lambda \quad \beta_1 = 0 \]
\[ \gamma_0 = \mathcal{H}_k [6(1 - \lambda) - f_1 + f_3] \quad \gamma_1 = 3(1 - \lambda) - f_1 \]
\[ \epsilon_0 = \frac{1}{2} \mathcal{H}_k (f_3 - f_1) \quad \epsilon_1 = -\frac{f_1}{2} \]  

(2.75)

The Hořava-Lifschitz coefficients obey the constraint relations of eqs.(2.27-2.33), hence it is classified as a type 1 theory.

We see that Hořava-Lifschitz theory possesses a distinctive signature: nearly all of its PPF coefficients contain spatial derivative operators that are greater than second order. This feature is a result of the particular type of Lorentz violation of Hořava-Lifschitz theory, which allows time and space coordinates to behave differently under scaling transformations in the UV.

It is these kinds of distinctive signatures that could be a useful tool in guiding us towards (or eliminating) particular regions of theory space when confronted with data. In the case of Hořava-Lifschitz gravity, the challenge will be the detection of strongly scale-dependent components of the PPF functions which are likely to be subdominant to the ‘normal’ scale-free and \( k^2 \) terms (see §2.4.2).

2.3.6 Horndeski Theory

The last few years have witnessed a resurgence of interest in Horndeski theory [159], which had lain largely forgotten since 1974 until it was independently re-derived by Deffayet and collaborators [102]. Horndeski theory is the most general Lorentz-invariant extension of GR in four dimensions that can be constructed using a single additional scalar field, with the restriction that the equations of motion must remain second order in time derivatives. All theories that fit this description can be obtained via special choices of four arbitrary functions of the scalar field that appear in the Horndeski Lagrangian. A non-exhaustive list of theories that fall under the Horndeski umbrella is: Brans-Dicke and scalar-tensor gravity, \( f(R) \) gravity (in its scalar-tensor formulation), single-field quintessence and K-essence theories, single-field inflation models, the covariant Galileon, the Fab Four, Dirac-Born-Infeld theory, Kinetic Gravity Braiding, actions involving derivative couplings between a scalar field and the Einstein tensor and \( f(\mathcal{G}) \) theories, where \( \mathcal{G} \) is the Gauss-Bonnet term.
### 2. THE PARAMETERIZED POST-FRIEDMANN FRAMEWORK FOR THEORIES OF MODIFIED GRAVITY

Table 2.4: This table lists the choices for the four free functions of the Horndeski Lagrangian that reproduce some previously-studied theories of modified gravity. In the table above, if a function is left in general terms then the choice is arbitrary. For the covariant Galileon, $c_i$ are dimensionless constants and $m$ is a mass scale. For the fourth line, $\bar{\kappa}_i = \int \kappa_i d\mathcal{X}$ and expressions for the $\kappa_i$ that give rise to the ‘Fab Four’ theory are presented in [78]. In the penultimate line, $\varepsilon = +1$ for quintessence and $\varepsilon = -1$ for phantom scalar fields.

<table>
<thead>
<tr>
<th>Theory</th>
<th>$K(\phi, \mathcal{X})$</th>
<th>$G_3(\phi, \mathcal{X})$</th>
<th>$G_4(\phi, \mathcal{X})$</th>
<th>$G_5(\phi, \mathcal{X})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalar-tensor</td>
<td>$\frac{M_P \omega(\phi)}{\phi} - V(\phi)$</td>
<td>0</td>
<td>$\frac{M_P \phi}{2}$</td>
<td>0</td>
</tr>
<tr>
<td>$f(R)$ gravity</td>
<td>$-\frac{M_P^2}{2}(Rf_R - f(R))$</td>
<td>0</td>
<td>$\frac{M_P^2 f_R}{2}$</td>
<td>0</td>
</tr>
<tr>
<td>cov. Galileon</td>
<td>$-c_2 \mathcal{X}$</td>
<td>$\frac{c_3}{m^3} \mathcal{X}$</td>
<td>$\frac{M_P^2}{2} - \frac{c_4}{m^n} \mathcal{X}^2$</td>
<td>$\frac{3c_5}{m^3} \mathcal{X}^2$</td>
</tr>
<tr>
<td>Horndeski’s</td>
<td>$\kappa_9 + 4\mathcal{X}(\bar{\kappa}<em>{8,\phi} - 2\bar{\kappa}</em>{3,\phi\phi})$</td>
<td>$-2\left(6\bar{\kappa}<em>{1,\phi\phi} + 8\bar{\kappa}</em>{1,\phi} - \bar{\kappa}_8 + \mathcal{X}\bar{\kappa}<em>8 - 8\mathcal{X}\bar{\kappa}</em>{3,\phi}\right)$</td>
<td>$-4\left(\bar{\kappa}_{1,\phi} + \bar{\kappa}<em>3 - \mathcal{X}\bar{\kappa}</em>{3,\phi}\right)$</td>
<td>$-4\kappa_1$</td>
</tr>
<tr>
<td>original notation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K.G.B</td>
<td>$K(\phi, \mathcal{X})$</td>
<td>$G_3(\phi, \mathcal{X})$</td>
<td>$\frac{1}{2} M_P^2$</td>
<td>0</td>
</tr>
<tr>
<td>Quintessence &amp;</td>
<td>$\varepsilon \mathcal{X} - V(\phi)$</td>
<td>0</td>
<td>$\frac{1}{2} M_P^2$</td>
<td>0</td>
</tr>
<tr>
<td>phantom fields</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K-essence &amp;</td>
<td>$K(\phi, \mathcal{X})$</td>
<td>0</td>
<td>$\frac{1}{2} M_P^2$</td>
<td>0</td>
</tr>
<tr>
<td>K-inflation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Inevitably, the price to be paid for this powerful generality is that Horndeski theory is cumbersome to calculate with. Its action is [97, 134]:

\[ S_{HD} = \int d^4x \sqrt{-g} \left( \sum_{i=2}^{5} \mathcal{L}_i + \mathcal{L}_M (\psi^\alpha, g_{\mu\nu}) \right), \]  

(2.76)

where

\[
\begin{align*}
\mathcal{L}_2 &= K(\phi, X) \\
\mathcal{L}_3 &= -G_3(\phi, X) \Box \phi \\
\mathcal{L}_4 &= G_4(\phi, X) R + G_{4,X} \left[ (\Box \phi)^2 - (\nabla_\mu X_\nu) (\nabla^\mu X^\nu) \phi \right] \\
\mathcal{L}_5 &= G_5(\phi, X) G_{\mu\nu} (\nabla^\mu X^\nu) \\
&\quad - \frac{1}{6} G_{5,X} \left[ (\Box \phi)^2 - 3 (\Box \phi) (\nabla_\mu X_\nu) (\nabla^\mu X^\nu) \phi + 2 (\nabla^\mu X_\alpha) (\nabla^\alpha X_\beta) (\nabla^\beta X^\mu) \phi \right]
\end{align*}
\]

\( K \) and \( G_i \) (i=3,4,5) are four functions of a scalar field and its kinetic energy, \( X = -\partial^\mu \phi \partial_\mu \phi /2 \). GR is recovered by setting \( G_4 = M_P^2 /2 \) and \( G_3 = G_5 = K = 0 \). The cosmological perturbations of the Horndeski action were first derived in [134], but I will follow the notation of [97] by defining the following useful quantities:

\[
\begin{align*}
\mathcal{F}_T &= 2 \left[ G_4 - X \left( \frac{\dot{\phi}}{a^2} G_{5,X} - \frac{\mathcal{H} \dot{\phi}}{a^2} G_{5,\phi} \right) \right] \quad (2.81) \\
\mathcal{G}_T &= 2 \left[ G_4 - 2X G_{4,X} - X \left( \frac{\mathcal{H} \dot{\phi}}{a^2} G_{5,X} - G_{5,\phi} \right) \right] \quad (2.82) \\
\alpha \Theta &= -\dot{\phi} X G_{4,X} + 2X G_{4} - 8X G_{4,X} - 8X \mathcal{H} X G_{4,XX} + \dot{\phi} G_{4,\phi} + 2X \dot{\phi} G_{4,\phi X} \\
&\quad - \frac{3\mathcal{H}^2}{a^2} \dot{\phi} (5X G_{5,X} + 2X^2 G_{5,XX}) + 2\mathcal{H} X (3G_{5,\phi} + 2X G_{5,\phi X}) \quad (2.83) \\
\alpha^2 \Upsilon &= a^2 \{ 2X K_{X} + 2X^2 K_{XX} \} + 12X \phi X G_{3,X} + 6X \phi X^2 G_{3,XX} - 2a^2 \{ X G_{3,\phi} + X^2 G_{3,\phi X} \} \\
&\quad - 6X^2 G_{4} + 6 \left[ 3\mathcal{H}^2 (7X G_{4,X} + 16X^2 G_{4,XX} + 4X^3 G_{4,XXX}) \right] \\
&\quad - 3\mathcal{H} \phi (G_{4,\phi} + 5X G_{4,\phi X} + 2X^2 G_{4,\phi XX}) \quad (2.84) \\
&\quad + \frac{1}{a^2} \left\{ 30X \mathcal{H}^3 \phi X G_{5,X} + 26X^3 \phi X^2 G_{5,XX} + 4X^3 \phi X^3 G_{5,XXX} \right\} \quad (2.85) \\
&\quad - 6X^2 \mathcal{H} (6G_{5,\phi} + 9X G_{5,\phi X} + 2X^2 G_{5,\phi XX}) \quad (2.86)
\end{align*}
\]

where the relation \( X \partial_X \dot{\phi} = \dot{\phi} /2 \) has been used. Note that I have used conformal time, contrary to the authors of [97]. The gauge-invariant perturbation of the
2. THE PARAMETERIZED POST-FRIEDMANN FRAMEWORK FOR THEORIES OF MODIFIED GRAVITY

The parameterized scalar field is:

\[ \dot{\chi} = \frac{\delta \phi}{M_p} - \frac{\dot{\phi}}{M_p} \frac{1}{6 \dot{\chi}} (\beta + k^2 \nu) \]  \quad (2.87)

Using a tilde to denote division by the square of the reduced Planck mass (i.e. \( \tilde{G}_T = G_T / M_p^2 = \kappa G_T \)), the PPF coefficients for Horndeski theory are:

\[
\begin{align*}
A_0 &= -2 \left( 1 - a \frac{\Theta}{\dot{\chi}} \right) - \frac{\dot{\phi}}{\dot{\chi}} \frac{a^2}{k^2} \mu + \frac{2}{\dot{\chi}^2 k^2} \left( \ddot{\chi} - \frac{\dot{\chi}^2}{\chi} \right) (a^2 \ddot{\chi} + 3 \dot{\chi} a \dddot{\chi}) \\
B_0 &= \frac{1}{k \dot{\chi}} \left( \kappa a^2 \rho_M - 2(\dot{\chi}^2 - \frac{\dot{\chi}}{\chi}) \tilde{\Theta} a \right) \\
C_0 &= 2(1 - \tilde{g}_T) - 2 \frac{\tilde{g}_T}{\dot{\chi}} \left( 1 + 3 \frac{\dot{\chi}}{k^2} \right) - 6 \frac{\tilde{g}_T}{k^2} \left( 2 \frac{\dot{\chi}}{k} + \frac{\dot{\chi}}{\dot{\chi}} \right) - 3 \frac{\dot{\chi}}{k^2 \dot{\chi}} \kappa a^2 \rho_M \\
&\quad + 6 \frac{\dot{\chi}}{k^2} \left( 2 \tilde{g}_T^2 - \frac{\dot{\chi}}{H} \right) - 12 \frac{\dot{\chi}^2}{k^2 \dot{\phi}^2} \left( \tilde{g}_T - a \frac{\Theta}{\dot{\chi}} \right) + 6 \frac{\dot{\chi} a \dddot{\chi}}{k^2 \dot{\chi}} - 3 a^2 \dddot{\chi} \\
&\quad + \frac{3}{k^2} \left[ 2 \tilde{g}_T \left( \tilde{g}_T - \frac{a \dddot{\chi}}{\dot{\chi}} \right) + \dot{\phi} \left( 2 \dot{\chi} - 2 a \dddot{\chi} + 4 \frac{\dot{g}_T}{\dot{\chi}} \left( \dot{\chi} + \frac{\dot{\chi}}{\dot{\chi}} \right) - 8 a \dddot{\chi} + \frac{1}{k \dot{\chi}} \kappa a^2 \rho_M \right] \\
C_1 &= 2 \frac{k}{\dot{\chi}} (1 - \tilde{g}_T) + 6 \frac{\dot{\chi}}{k^2 \dot{\chi}} \left( 1 - \frac{\dot{\chi}}{\dot{\chi}} \right) \\
D_0 &= 1 - \tilde{g}_T - \frac{\dot{g}_T}{\dot{\chi}} \\
D_1 &= \frac{k}{\dot{\chi}} (1 - \tilde{g}_T) \\
F_0 &= -\frac{2}{k \dot{\chi}} (3 \dot{\chi} \dddot{\chi} + a^2 \dddot{\chi}) \\
H_0 &= 2 \left( 1 - \frac{\dot{\chi}}{\dot{\chi}} \right) \\
I_0 &= \frac{1}{k \dot{\chi}} \left[ -2k^2 (1 - \tilde{g}_T) + 3 \kappa a^2 \rho_M - 6 \frac{d}{d \eta}(a \dddot{\chi}) + 6 (\dot{\chi}^2 + \dddot{\chi}) - 6 \frac{\dot{\chi}}{\dot{\chi}} (2 \dot{\chi}^2 - \dddot{\chi}) \right] \\
J_0 &= 6 \left( 1 - \frac{\dot{\chi}}{\dot{\chi}} \right) \\
K_0 &= -\frac{k}{\dot{\chi}} (1 - \tilde{g}_T) \\
K_1 &= 0 \\
\alpha_0 &= M_p \left[ \frac{a^2}{k^2} \mu - \frac{2}{\dddot{\chi}} \left( \dddot{\chi} - \dot{\chi} \dddot{\chi} \right) \right] \\
\alpha_1 &= \frac{2 M_p}{k \dot{\phi}} \left[ a^2 \dddot{\chi} + 3 \dot{\chi} a \dddot{\chi} \right] \\
\beta_0 &= \frac{M_p}{k \dot{\phi}^2} \left[ -2 \dddot{\chi} (a \dddot{\chi} - \dot{\chi} \dddot{\chi}) - 2 \dot{\chi} a \dot{\chi} \dddot{\chi} + 2 a \dddot{\chi} a \dddot{\chi} \right] - \frac{M_p}{k \dot{\phi}} \frac{\kappa a^2 \rho_M}{k \dot{\phi}}
\[
\beta_1 = \frac{2M_P}{\dot{\phi}} \left[ \tilde{\Theta}a - 3\tilde{\mathcal{H}}\tilde{g}_T \right]
\]
\[
\gamma_0 = \frac{2M_P}{\dot{\phi}} \left( \dot{\tilde{g}}_T + 3\mathcal{H}\tilde{g}_T - 3\tilde{\mathcal{H}}\tilde{f}_T \right) + 3M_P \frac{a^2}{k^2} \tilde{\mathcal{V}}
\]
\[
\gamma_2 = \frac{6M_P}{\dot{\phi}} \left( \tilde{\Theta}a - 3\tilde{\mathcal{H}}\tilde{g}_T \right)
\]
\[
\gamma_1 = \frac{M_P}{k\dot{\phi}} \left[ -6\tilde{g}_T \left( \mathcal{H} + 2\dot{\kappa}^2 - 2\kappa \frac{\ddot{\phi}}{\dot{\phi}} \right) + 6 \frac{d}{d\eta} \left( a\tilde{\Theta} - 3a\tilde{\mathcal{H}}\tilde{g}_T \right) + 6a\tilde{\Theta} \left( 3\mathcal{H} - 2\frac{\ddot{\phi}}{\dot{\phi}} \right) \right]
\]
\[
-3\kappa a^2 \rho_M \right]
\]
\[
\epsilon_0 = \frac{M_P}{\dot{\phi}} \left[ \dot{\tilde{g}}_T + 3\mathcal{H}\tilde{g}_T - 3\tilde{\mathcal{H}}\tilde{f}_T \right]
\]
\[
\epsilon_1 = \epsilon_2 = 0 \quad (2.88)
\]

\[\mu \text{ and } \mathcal{V} \text{ are derivatives of the zeroth-order field equations with respect to the scalar field (00 and ii components respectively) - see [97] for the relevant expressions.}\]

Table 2.4 collects some ‘settings’ for the Horndeski Lagangian functions that reproduce theories of current interest. The application of the PPF formalism to Horndeski theory immediately brings a large realm of theory space within reach of our parameterization.

### 2.3.7 GR with a Dark Fluid

Our final example should really be classed as a dark energy model rather than a theory of modified gravity, though arguably the distinction is not important. We present the example of an adiabatic dark fluid characterized by a constant equation of state \(\omega_D\) and negligible anisotropic stress. We will use this as an example of how a theory with two additional fields can be recast as a single-field theory (and hence type 1/2/3) in some cases. The example also has relevance to theories that can be usefully written as an effective fluid at the level of the linearized gravitational field equations, e.g. quintessence and its progeny (whilst an effective fluid interpretation is possible for all theories, it is not always useful).

The zeroth-order modifications to the field equations are simply:
\[
a^2X = \kappa a^2 \rho_D \quad a^2Y = \kappa a^2 \omega_D \rho_D \quad (2.89)
\]
The two new fields are the fractional energy density perturbation and the velocity...
perturbation, $\delta_D$ and $\theta_D$, which appear in the $U$-tensor as:

\[
\begin{align*}
U_\Delta &= \kappa a^2 \rho_D \delta_D \\
U_\Theta &= \kappa a^2 \rho_D (1 + \omega_D) \theta_D \\
U_P &= 3 \kappa a^2 \rho_D \omega_D \delta_D \\
U_\Sigma &= 0
\end{align*}
\]

(2.90)

The fluid velocity is given by $v^i_D = \nabla^i \theta_D$. For an adiabatic, shear-free fluid the conservation and Euler equations are:

\[
\begin{align*}
\dot{\delta}_D &= -(1 + \omega_D) \left( k^2 \theta_D + \frac{1}{2} \dot{\beta} - k^2 \epsilon \right) \\
\dot{\theta}_D &= -\mathcal{H}(1 - 3 \omega_D) \theta_D + \frac{\omega_D}{1 + \omega_D} \delta_D - \Xi
\end{align*}
\]

(2.91, 2.92)

These correspond to the two components of the conservation equation $\nabla_\mu U^\mu = 0$. Eq.(2.91) can be used to eliminate $\theta_D$ in favour of $\delta_D$ from the gravitational field equations, resulting in a theory with a single new field. After forming the relevant g.i. combination

\[
\dot{\delta}_D = \delta_D + \frac{1}{2} (1 + \omega)(\beta + k^2 \nu)
\]

(2.93)

the PPF coefficients can be extracted (with all those not stated being zero):

\[
\begin{align*}
A_0 &= 9 \mathcal{H} k^2 \Omega_D (1 + \omega_D) \\
C_0 &= 27 \mathcal{H} k^2 \Omega_D (1 + \omega_D) \omega_D \\
\alpha_0^\delta &= 3 \mathcal{H} k^2 \Omega_D \\
\beta_1^\delta &= -3 \mathcal{H} k^2 \Omega_D \\
\gamma_0^\delta &= 9 \mathcal{H} k^2 \Omega_D \omega_D
\end{align*}
\]

(2.94)

where $\Omega_D$ is the ratio of the energy density of the dark fluid to the critical density $\rho_c = 3 \mathcal{H}^2 / (\kappa a^2)$, as per the usual definition. The superscripts indicate that these coefficients 'belong' to the $\delta_D$ perturbations, and should not be confused with spacetime indices. The e.o.m. is obtained by using eq.(2.91) to replace $\theta_D$ in eq.(2.92), leading to:

\[
\ddot{\delta}_D + \mathcal{H} \dot{\delta}_D (1 - 3 \omega_D) + k^2 \omega_D \left[ \delta_D + 3 (1 + \omega_D) \dot{\Phi} \right] + \frac{k^3}{\mathcal{H}} \omega_D \dot{\mathcal{H}} - \frac{k^2}{\mathcal{H}} \omega_D \dot{\Phi} = 0
\]

(2.95)

Using the PPF coefficients of the modified single-field system, we find that this is our first example of a type 2 theory. That is to say, the coefficients of each term in eq.(2.26) vanish identically.

Of course we can implement an analogous procedure to eliminate $\delta_D$ and treat $\theta_D$ as the single extra field. Note that this cannot be done for the special
case $\omega_D = 0$, for which we cannot invert eq.(2.92) to find $\delta_D(\theta_D, \dot{\theta}_D)$ (and using eq.(2.91) instead would lead to integral expressions that do not fit into our parameterization). The relevant gauge-invariant quantity in this case is:

$$\hat{\theta}_D = \theta_D - \frac{1}{6\mathcal{H}}(\beta + k^2\nu)$$

(2.96)

Repeating similar steps to before we obtain the PPF coefficients for the new theory, where now $\theta_D$ is the only additional field. Using tildes to indicate that these are not the same as eqs.(2.94), the results are:

$$\tilde{A}_0 = 33\mathcal{H}_k^2\Omega_D(1 + \omega_D)\left[3 - \frac{1}{\omega_D}\left(1 - \frac{\dot{\mathcal{H}}}{\mathcal{H}^2}\right)\right]$$

$$\tilde{B}_0 = -3\mathcal{H}_k\Omega_D(1 + \omega_D)$$

$$\tilde{C}_0 = 9\mathcal{H}_k^2\Omega_D(1 + \omega_D)\omega_D\left[3 - \frac{1}{\omega_D}\left(1 - \frac{\dot{\mathcal{H}}}{\mathcal{H}^2}\right)\right]$$

$$\tilde{F}_0 = -\frac{(1 + \omega_D)}{\omega_D}3\mathcal{H}_k\Omega_D$$

$$\tilde{J}_0 = -9\mathcal{H}_k\Omega_D(1 + \omega_D)$$

$$\tilde{\alpha}^{k\phi}_0 = 33\mathcal{H}_k^{2}\Omega_D(1 - 3\omega_D)(1 + \omega_D)$$

$$\tilde{\alpha}^{k\phi}_1 = 33\mathcal{H}_k^{2}\Omega_D(1 + \omega_D)$$

$$\tilde{\beta}^{k\phi}_0 = 33\mathcal{H}_k^{2}\Omega_D(1 + \omega_D)$$

$$\tilde{\gamma}^{k\phi}_0 = 9\mathcal{H}_k^{2}\Omega_D(1 - 3\omega_D)(1 + \omega_D)$$

$$\tilde{\gamma}^{k\phi}_1 = 9\mathcal{H}_k^{2}\Omega_D(1 + \omega_D)$$

(2.97)

This time the 0-component of the Bianchi identity provides the e.o.m.:

$$\ddot{\theta}_D + \mathcal{H}\dot{\theta}_D(1 - 3\omega_D) + \theta_D \left[\mathcal{H}(1 - 3\omega_D) + k^2\omega_D\right] + \dot{\Xi} + \frac{1}{2}(1 + \omega_D)(\ddot{\gamma} - 2k^2\dot{\epsilon}) = 0$$

(2.98)

The constraints of eqs.(2.27-2.33) are all satisfied. Hence in this formulation the dark fluid becomes a type 1 theory.

We have at last found a counter-example to the monopoly of type 1 theories. But is this an ‘artificial’ result of moulding a two-field theory into single-field format? We discuss the issue a little further in §2.6.2.
2.4 The PPF Coefficients

Different theories leave characteristic signatures in the set of PPF coefficient functions that in principle can be used to distinguish between them. In the first part of this section we discuss the issue of degeneracy between the functions of our framework. In the second part we consider how their scale-dependence should guide the implementation of PPF in numerical calculations.

It is generally argued that the lengthscales relevant to observables such as weak gravitational lensing and redshift-space distortions lie within the ‘quasistatic’ regime. We will briefly consider the behaviour of the PPF framework in this limit in §2.4.3; the quasistatic regime will be studied in detail in chapter 5. In the final part of this section we briefly connect our work to a recent parameterization of screening mechanisms [60], which are designed to suppress the effects of modified gravity on small scales.

2.4.1 Degeneracy

The PPF parameterization employs more free functions than some other similar frameworks in the literature, and it is difficult to predict the degeneracy structure of this set of functions \textit{a priori}. We expect to find that a subset of the PPF functions can be well-constrained, whilst another subset may be more difficult to pin down. We do not expect this to be problematic, as there should still be sufficient information in the well-constrained subset to distinguish between classes of theories.

Furthermore, based on the work of [38], we believe that sets of further relations between the PPF functions exist. These can be derived once the field content of a theory has been specified, and will remove further freedom from the parameterization. This will be in addition to the constraint relations of eqs.(2.27-2.33). We will pursue this point in a future work. In [32] we will consider in more concrete terms the prospects for discriminating between theories of modified gravity with specific current and future experiments.

The construction of the PPF parameterization necessarily restricts its use to lengthscales at which linear perturbation theory remains valid. However, the advance of errorbars of future experiments towards cosmic variance limits will allow tighter constraints to be obtained even just using the linear window of $k$-space. Furthermore, the advent of new high-redshift probes such as 21cm intensity mapping will give access to early times when a wider range of scales fell within the linear regime.
Table 2.5: The scale-dependence of the terms appearing in the modified Poisson equation (2.99) for the example theories of §2.3. Recall that \( \hat{\Phi} \propto k^{-1} \), see eq.(2.8).

Note that in some cases the scalar perturbation \( \chi \) includes a factor of \( k \) so that it is dimensionless, e.g. \( \chi = k \xi_0 \) in Hořava-Lifshitz gravity, \( \chi = k \theta \) for a fluid velocity potential. EBI and DGP gravity will contain multiple copies of terms 3 and 4 (eg. \( \hat{\chi}_1 = \hat{\delta}_W \), \( \hat{\chi}_2 = k \theta_W \) for DGP), but both terms have the same scale-dependence.

To make use of new data probing smaller scales one would need to use a prescription for non-linear corrections, such as [209]. Applying such a prescription requires calibration from N-body simulations of specific modified gravity models, and this undermines the model-independent approach that we are pursuing. It would also introduce further parameters to be marginalized over in a data analysis of the PPF formalism; given the possible degeneracies already present, we choose not to use non-linear-scale data (\( k \gtrsim 0.2 \) h Mpc\(^{-1} \) at \( z = 0 \)) at this time.

### 2.4.2 Signatures and Scale-Dependence

Naively, one might expect the well-constrained subset of PPF functions to be those which feature directly in the calculation of observable quantities. One rarely needs to use the full set of Einstein equations to calculate observables; usually only the Poisson equation, the slip relation \((\Phi - \Psi)\), and the evolution equations for \( \delta_M \) and \( \theta_M \) are required. The modified Fourier-space Poisson equation in the PPF parameterization is (taking eq.(2.18)+3\( \mathcal{H} \times \)eq.(2.19)):

\[
-2k^2 \hat{\Phi} = \kappa a^2 \rho_M \Delta_M + (A_0 + 3\mathcal{H}kB_0) k^2 \hat{\Phi} + (F_0 + 3\mathcal{H}kI_0) k^2 \hat{\Gamma} + (\alpha_0 + 3\mathcal{H}k\beta_0) k^2 \hat{\chi} + (\alpha_1 + 3\mathcal{H}k\beta_1) k \dot{\hat{\chi}}
\]  

(2.99)

where \( \Delta_M = \delta_M + 3\mathcal{H}(1 + \omega_M)\theta_M \) is a gauge-invariant density perturbation.

Table 2.5 indicates the scale-dependence of the non-standard terms of eq.(2.99)
2. THE PARAMETERIZED POST-FRIEDMANN FRAMEWORK FOR THEORIES OF MODIFIED GRAVITY

in each of the theories of §2.3. It is clear that modifications to the Poisson equation have the form of a power series in even powers of the Fourier wavenumber $k$. It is worth stressing that this form is exact, and not a Taylor series expansion of a more complicated function; an even power series in $k$ is the only possibility. In most cases only the scale-free and quadratic terms are present, in agreement with the findings of [17] (apart from DGP, as discussed in §2.3.3). This is no surprise - we have focused on theories that are second order in time derivatives, and Lorentz invariance then implies that they must be second order in spatial derivatives too. Odd powers of $k$ are forbidden by parity.

Horava-Lifshitz gravity is a notable exception - its explicit breaking of the symmetry between time and space coordinates allows higher-order spatial derivatives to enter the PPF coefficients. Interestingly, all the theories except Horava-Lifshitz gravity are found to have scale-independent slip relations (for brevity we have not decomposed the slip relation into its individual terms in Table 2.5). The coefficients $K_1$ and $\epsilon_2$ are zero in all cases.

Based on these observations there are three ways to implement the PPF coefficients in an Einstein-Boltzmann solver code such as the CAMB package [201]:

i) We can construct a sensible parameterization of the PPF coefficients by splitting them into purely time-dependent functions multiplying a spatial dependence. To be fully general (at least in terms of the theories treated in this chapter) we should expand each of the PPF functions appearing in the Poisson equation as:

$$A_0 = \sum_{n=0}^{4} \tilde{A}_0(z) k^{2(n-1)} \quad (2.100)$$

and similarly for $B_0, F_0$ etc. One must choose a motivated ansatz for the time-dependent function $\tilde{A}_0(z)$, such as a Taylor series in $\Omega_\Lambda$ or $(\mathcal{H}_0/\mathcal{H})^{-1}$. In reality, a form as general as eq. (2.100) is probably unnecessary for all of the coefficients.

ii) Alternatively one may employ a Principal Component Analysis (PCA), as advocated in [158, 252, 320, 321]. The PCA approach expands the functions to be constrained in a basis of orthogonal eigenmodes in $\{k, z\}$-space, using the eigenbasis that a given data set is most sensitive to. Informally speaking, PCA reveals ‘what the data knows’, rather than ‘what we would like to know’.

90
iii) A compromise between the two methods above is possible: expand the free functions as in eq.\((2.100)\), then perform a one-dimensional PCA for the unknown functions of redshift.

Method i) has the advantage of simplicity, and it utilizes the common features of modified gravity theories to jump directly to the most relevant regions of theory space. However, by choosing a form for functions such as \(\hat{A}_0(z)\) we are imposing our preconceptions on the behaviour of modified gravity, namely that it must emerge at late times \((z \lesssim 1)\) in order to reproduce the effects of an apparent dark energy.

Method ii) maintains a greater degree of agnosticism, avoiding the use of semi-arbitrary functional forms. However, it effectively throws away much of our prior knowledge about the structure of physically-reasonable theories.

We are most likely to pursue method iii), which represents a compromise between constraining known regions of the modified gravity landscape and exploring new unknown territory. For the sake of argument, let us suppose we use \(n = 1, 2\) for terms 1 and 3 in Table 2.5 and \(n = 1\) for terms 2 and 4. The Poisson equation is then described by six functions of redshift. Since \(K_1 = \epsilon_2 = 0\) for all theories (a result of the additional constraint equations we alluded to in §2.4.1), the slip function contains five functions of redshift (since all terms in it are scale-independent). This totals eleven functions of redshift only to constrain. The authors of [158] find that combining results from the Planck satellite with an LSST-like experiment yields 155 data points, which gives us fourteen data points per function. This should be sufficient to pin down a few eigenvalues for each of the eleven functions.

### 2.4.3 The Quasistatic Limit

On sufficiently subhorizon scales \((\mathcal{H}_k \ll 1)\) – but above the non-linearity scale – the time derivatives of metric perturbations are small relative to their spatial derivatives, and can be neglected (at least for the purposes of N-body simulations). This is known as the quasistatic regime [167, 281], and will be discussed in detail in chapter 5. It is believed that time derivatives of the additional fields in modified gravity theories can be similarly neglected in this regime: many scalar-field type models exhibit damped oscillatory solutions below a special lengthscale, which rapidly decay into insignificance. Only the averaged, stationary spatial profile then remains of relevance.

Let us consider type 1 theories, as these are the most prevalent. In the quasistatic limit eq.\((2.26)\), the e.o.m. for the new scalar d.o.f., reduces to a relation
between the perturbations of the extra field and the potential \( \Phi \):

\[
[\dot{\alpha}_0 + k\beta_0] \dot{\chi} + \left[ \dot{A}_0 + kB_0 \right] \dot{\Phi} = 0 \tag{2.101}
\]

Time derivatives of the PPF functions have been maintained because the quasistatic approximation only allows us to neglect time derivatives of perturbations.

To obtain the quasistatic limit of the Poisson equation we use eq.(2.101) to replace \( \dot{\chi} \) in eq.(2.99), and drop the \( \dot{\chi} \) and \( \dot{\Gamma} \) terms (recall that \( \dot{\Gamma} = 1/k (\dot{\Phi} + \mathcal{H}\dot{\Psi}) \), where \( \dot{\Phi} \) will be small and \( \dot{\Psi} \) is suppressed by \( \mathcal{H} \)). We apply a similar treatment to the slip relation, neglecting the anisotropic stress of matter. The results can be written in the form:

\[
-2k^2\dot{\Phi} = \kappa a^2 \mu(a, k) \rho_M \Delta_M \tag{2.102}
\]

\[
\frac{\dot{\Phi}}{\dot{\Psi}} = \gamma(a, k) \tag{2.103}
\]

where

\[
\mu(a, k) = \left[ 1 + \frac{A_0}{2} - \frac{\alpha_0}{2} \left( \frac{\dot{A}_0 + kB_0}{\hat{\alpha}_0 + k\beta_0} \right) \right]^{-1} \tag{2.104}
\]

\[
\gamma(a, k) = \left[ 1 - D_0 + \epsilon_0 \left( \frac{\dot{A}_0 + kB_0}{\hat{\alpha}_0 + k\beta_0} \right) \right]^{-1} \tag{2.105}
\]

We see that on quasistatic scales modifications to the Poisson and slip relations can be reparameterized in terms of two time- and scale-dependent functions, \( \mu(a, k) \) and \( \gamma(a, k) \). This form of parameterization has been explored in a large number of investigations \([40, 45, 90, 113, 167, 208, 252, 295]\), using a number of different ansatzes for the scale-dependence of \( \{\mu, \gamma\} \). From the discussion of the previous subsection we see that one does not really have the freedom to choose just any ansatz here; with a few exceptions, the physically-relevant choice is that of scale-free terms plus \( k^2 \) terms, such as used in \([45]\).

Deviations of \( \{\mu, \gamma\} \) from their GR values of \( \{1, 1\} \) are degenerate between modifications to the metric sector and new d.o.f. sector, indicated by the presence of Greek and Latin PPF coefficients in both. The advantage of maintaining the unapproximated PPF Poisson equation (eq.(2.99)) is that these coefficients are kept distinct, which is necessary if we hope to detect the signatures of particular theories discussed in the previous subsection. Keeping terms such as \( \dot{\chi} \) in the equations may be unfeasible or unnecessary for N-body simulations, but it should not pose a problem for Einstein-Boltzmann codes. Furthermore, lift-
ing the quasistatic approximation means that the PPF parameterization is valid all the way up to horizon scales, which are relevant for large-scale CMB modes contributing to the late-time Integrated Sachs-Wolfe effect and the lensing-ISW cross-correlation \[106, 162\].

2.4.4 Connection to Screening Parameterizations

Brax and collaborators have introduced a new parameterization \([58, 60]\) that is optimized for screening mechanisms, such as the chameleon \([183, 184]\), dilaton \([59]\) and symmetron mechanisms \([152, 153]\), see §1.3.5. In this parameterization a theory is described by the time-evolving mass of the scalar field, \(m(a)\), and its coupling to matter in the Einstein frame, \(\beta(a)\), where the Jordan and Einstein metrics are related by:

\[
g_{\mu\nu}^J = A^2(\phi) g_{\mu\nu}^E \tag{2.106}
\]

and the coupling function is:

\[
\beta(a) = \frac{d \ln A(\phi)}{d \phi} \tag{2.107}
\]

The parameterization is constructed explicitly in the Einstein frame; this means that when mapped to our (Jordan-frame) parameterization it mixes PPF coefficients from the metric and extra-field sectors. In the quasistatic regime the mapping is:

\[
\left[1 + \frac{A_0}{2} - \frac{\alpha_0}{2} \left(\frac{\dot{A}_0 + kB_0}{\dot{\alpha}_0 + k\beta_0}\right)\right]^{-1} = 1 - \frac{2\beta^2}{1 + m^2 \frac{a^2}{k^2}} \tag{2.108}
\]

\[
\left[1 - D_0 + \epsilon_0 \left(\frac{\dot{A}_0 + kB_0}{\dot{\alpha}_0 + k\beta_0}\right)\right]^{-1} = 1 + \frac{2\beta^2}{1 + m^2 \frac{a^2}{k^2}} \tag{2.109}
\]

Adding the two expressions above gives an additional constraint relation between the PPF coefficients in the quasistatic regime.

If we lift the quasistatic approximation then in principle there is sufficient information to express the twenty-two PPF functions in terms of \(\beta(a)\) and \(m(a)\). However, the expressions are not particularly elegant and we will not present them here on grounds of relevance: recent work \([304]\) indicates that the scalar fields involved in the chameleon, dilaton and symmetron screening mechanisms must have a Compton wavelength \(\lesssim 1\) Mpc in order to satisfy Solar System and Galactic constraints. Deviations from GR are then only expected in the
2. THE PARAMETERIZED POST-FRIEDMANN FRAMEWORK FOR THEORIES OF MODIFIED GRAVITY

non-linear regime, to which PPF does not apply.

2.5 Conclusions

We have presented a new framework, the Parameterized Post-Friedmann formalism (PPF), for conducting model-independent tests of modified gravity. The construction of the framework does not rely on ad-hoc modifications to GR, but is built by considering the limited number of ways that the linearized Einstein equations can be extended whilst maintaining the properties of a physically-relevant gravitational theory. An exact, analytic mapping exists between a theory of modified gravity that obeys the assumptions of Table 2.2 and the set of PPF coefficient functions.

There are four key reasons we believe that the PPF formalism represents a significant step forward in parameterized treatments of modified gravity:

1. New degrees of freedom, a common feature of many models, are explicitly parameterized for. Until recently [60] this has not generally been the case. This is not a minor issue; new degrees of freedom are at the heart of the phenomenology that might render a theory of modified gravity distinguishable from ΛCDM.

2. The parameterization is not restricted to models based solely on scalar fields. As experimental precision increases, the continued success of ΛCDM has forced model-builders to look to richer theories that additionally involve vector and tensor fields. PPF is able to handle the spin-0 perturbations of these theories as easily as canonical scalar fields.

3. Once cast into the format of PPF, a theory of modified gravity may leave a distinctive ‘calling card’ in the set of PPF functions. Such signatures may allow some theories to be ruled out, or at least indicate the general characteristics of theories that are consistent with the data, eg. one expects pure scalar field models to share similarities, whilst Lorentz-violating theories may share similar traits. It seems reasonable to suppose that higher-dimensional models may have their own unique features.

4. The parameterization is valid from horizon scales down to the scales at which non-linearities become important. This spans a wide range of observables targeted by the next generation of experiments [1, 3, 4, 258], including weak lensing (of both galaxies and the CMB), redshift-space distortions, pe-
culiar velocity surveys, the ISW effect and associated cross-correlations, and possibly galaxy clusters [295].

The PPF framework can be applied in two modes depending on the user’s interests. It can be used as a tool for tackling observations, a gateway that allows data constraints to be applied to a multitude of ‘known’ theories simultaneously. By ‘known’ theories we mean existing models for which we have the field equations in hand. Alternatively, one can use the framework as an exploratory tool for model-building. As mentioned in 3) above, PPF may be able to indicate the physical features that lead to tension with the data, thereby guiding theorists to the regions of theory space that are likely to prove most fruitful. We note that for theories involving more than two new fields the exploratory mode cannot be used, but the observational mode remains fully functional.

The cost of the increased capabilities of the PPF framework is the introduction of more ‘free’ functions than other parameterizations (though we have seen in this chapter that the structure of these functions is most definitely not arbitrary). In §2.2.5 we presented a number of constraints that can be used to reduce this freedom, and we believe that there exist further relations not presented here [32]. The situation is likely to be analogous to that of the Parameterized Post-Newtonian framework, where a set of ten new potentials and ten new parameters are necessary to capture nearly all possible distortions of the GR metric: it turns out that many of these new parameters are needed in only a handful of special cases, and are zero the rest of the time.

We have been ambitious. It is possible that the current and next-generation datasets may not be able to tease apart the degeneracies between contributions to the modified field equations to the extent required here. In this situation the PPF framework still functions as a null test for non-GR behaviour, even if it cannot reach its full potential as a discriminator of modified gravity theories. Then the most pertinent question to ask is: how much better do our experiments need to be to realize this discriminatory power? Tackling this question requires implementation of the PPF framework in a numerical code for computing cosmological observables. The first steps towards constraining the PPF framework using the growth of large-scale structure will be presented in chapter 5.
2. THE PARAMETERIZED POST-FRIEDMANN FRAMEWORK FOR THEORIES OF MODIFIED GRAVITY

2.6 Supporting Material for Chapter 2

2.6.1 Transformation of $f(R)$ Gravity into a Scalar-Tensor Theory

Consider the action:

$$S = \int d^4x \sqrt{-g} [f(\zeta) + f'(\zeta)(R - \zeta)] + S_M(\psi^a, g_{\mu\nu}) \quad (2.110)$$

where primes denote differentiation with respect to the new scalar $\zeta$. Variation of this action gives the field equation for $\zeta$ as:

$$f''(\zeta)(R - \zeta) = 0 \quad (2.111)$$

Hence $\zeta = R$ for all $f''(\zeta) \neq 0$. Substituting $R = \zeta$ into eq.(2.110) recovers the standard $f(R)$ action. Therefore eq.(2.110) is an equivalent action for $f(R)$ gravity, with the special case of $f''(\zeta) = 0$ corresponding to the Einstein-Hilbert action.

To recast the equivalent action as a scalar-tensor theory we make the definitions:

$$\phi = \frac{df(R)}{dR} = f_R \quad V(\phi) = \frac{1}{2} [\zeta \phi - f(\zeta)] \quad (2.112)$$

Then eq.(2.110) becomes:

$$S = \int d^4x \sqrt{-g} [\phi R - 2V(\phi)] + S_M(\psi^a, g_{\mu\nu}) \quad (2.113)$$

which has the form discussed in §2.3.1, with $\omega(\phi) = 0$ in this case.

Note that because the metric has not been transformed its couplings to the scalar field and matter are unaffected.

2.6.2 Is the Bianchi 0-Component Always the E.O.M.? A Plausibility Argument

In §2.2.5 we introduced a classification scheme for single-field theories, according to whether the spatial or temporal part of the perturbed conservation law $\delta(\nabla_\mu U^\mu) = 0$ reduced to a trivial relation (types 1 and 2 respectively), or neither component did (type 3). Of the seven examples considered in §2.3 we found that four were type 1 (Scalar-tensor, Einstein-Aether, Hořava-Lifshitz and Horndeski).
EBI and DGP (the effective 4D theory) were not subject to this classification, having more than two new fields. The adiabatic, shear-free dark fluid could be recast as either type 1 or type 2.

It seems, then, that type 1 theories dominate. Why is this? We have yet to find an example of a type 3 theory, and it could be argued that our example of a type 2 theory is the result of an artificial re-writing. We present here a plausibility argument why all single-field theories may be type 1, but do not make any claims that it is a fully rigorous or watertight proof.

Consider the following action:

$$S = \int \left[ \frac{R}{2\kappa} + \mathcal{L}_M + \frac{1}{\kappa} \mathcal{L}_U \right] \sqrt{-g} \, d^4x$$  \hspace{1cm} (2.114)

where $\mathcal{L}_U$ is the Lagrangian that gives rise to the $U$-tensor that appears in the gravitational field equations, eq.(2.1). We assume that any gravitational action can be written in the form of eq.(2.114) by suitable choice of $\mathcal{L}_U$, e.g. for $f(R)$ gravity we would have $\mathcal{L}_U = \frac{1}{2}[f(R) - R]$. The variation of the last term in eq.(2.114) is:

$$\delta S_U = \frac{1}{2\kappa} \int U^{\mu\nu} \delta g_{\mu\nu} \sqrt{-g} \, d^4x$$  \hspace{1cm} (2.115)

where

$$U^{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\delta (\sqrt{-g} \mathcal{L}_U)}{\delta g_{\mu\nu}}$$  \hspace{1cm} (2.116)

Now consider an infinitesimal co-ordinate transformation, $x^\mu \rightarrow x^\mu + \xi^\mu(x)$. The corresponding change in the metric is given by $\delta g_{\mu\nu} = -\nabla_\mu \xi_\nu - \nabla_\nu \xi_\mu$. Substituting this expression into eq.(2.115) and utilizing the symmetry between indices $\mu$ and $\nu$:

$$\delta S_U = -\int (\nabla_\mu \xi_\nu) U^{\mu\nu} \sqrt{-g} \, d^4x$$ \hspace{1cm} (2.117)

$$= -\int \nabla_\mu (\xi_\nu U^{\mu\nu}) \sqrt{-g} \, d^4x$$

$$+ \int \xi_\nu (\nabla_\mu U^{\mu\nu}) \sqrt{-g} \, d^4x$$  \hspace{1cm} (2.118)

The first term of eq.(2.118) can be converted into a boundary integral via the usual divergence theorem. This boundary term must equal zero, because it is an assumption of the variational procedure that the small variations $\xi_\nu$ vanish at spatial infinity.

Demanding stationarity of the action, the second term of eq.(2.118) yields the
expression:

$$\xi^\nu \nabla_\mu U^\mu_\nu = 0 \quad (2.119)$$

which is easily recognized as the conservation law for $U^{\mu\nu}$ contracted with the gauge transformation vector. We obtain the two components of the conservation law for $U^{\mu\nu}$ by considering two separate coordinate transformations, a temporal one ($x^0 \rightarrow x^0 + \xi^0$) and a spatial one ($x^i \rightarrow x^i + \xi^i$). The first corresponds to translation along a worldline, and the second to translation of spatial coordinates across a fixed-time hypersurface. A vector proportional to $\xi^0$ is not a Killing vector of the FRW spacetime, so it is not surprising that a coordinate translation along $\xi^0$ results in an evolution equation for the new degrees of freedom. In contrast, spatial vectors proportional to $\xi^i$ are Killing vectors of the FRW metric - so no dynamics can be obtained from a translation along $\xi^i$. Hence the terms of the $\nu = i$ equation must vanish identically.
Chapter 3

Ambiguities in Parameterizations of Modified Gravity

In the previous chapter I introduced the Parameterized Post-Friedmann formalism (PPF), a framework for unifying the linear cosmological perturbation theory of many modified gravity theories. However, PPF is not a unique way of describing generalized cosmological perturbations; for example, other authors have constructed parameterizations of perturbed gravitational Lagrangians, inspired by effective field theory techniques used in particle physics [38, 39, 53, 54, 138, 144], see the discussion in §1.3.2.

Other authors have chosen more phenomenological parameterizations, motivated by the ease with which they can be constrained using observational data [14, 17, 40, 44, 69, 90, 91, 113, 114, 157, 167, 252, 286, 318, 322]. In this chapter I compare one such phenomenological parameterization to PPF and show that, despite intuitive expectations, the form of the parameterization used can have a significant impact on the constraints one obtains on deviations from GR.

3.1 Introduction

The original Parameterized Post-Newtonian formalism (see previous chapter for references) describes deviations from the metric of GR using a set of ten potentials and their ten constant coefficients. This is appropriate for tests of gravity on timescales significantly shorter than the evolutionary timescale of the gravitational system in question (e.g. the Hulse-Taylor binary pulsar system has been observed for several decades, whilst its lifetime is of order $10^8$ years).

The situation for cosmological tests of gravity is very different. By combining local observations with high-redshift galaxy surveys we are attempting to describe
gravitational physics over a significant fraction of the age of the universe. Hence any parameterization of modified gravity must allow for both a) the expansion and evolution of the cosmological background and b) time-dependent deviations from GR. Similarly, we must allow for scale-dependent effects, as some theories of modified gravity incorporate a preferred length scale that marks a transition between regimes of behaviour (e.g. the Compton wavelength of a non-GR propagating degree of freedom [251]). One is therefore led to constrain functions of space and time rather than a set of constant parameters.

We will continue to focus on parameterizations of the gravitational field equations (as opposed to parameterizations of the action or the metric). There are two levels of structure that must be imposed:

**Level 1:** One must choose the way in which free functions are used to extend the field equations. For example, one could use free functions to rescale terms that exist in GR, or allow new terms to be present.

**Level 2:** When implementing the parameterized field equations in numerical codes one is forced to choose a sensible ansatz for the time- and scale-dependence of the free functions. Typically these ansatzes are motivated by the observation that the non-matter sector only dominates the energy density of the universe at late cosmological times. See [80] for an enumeration of ansatzes that have been used by other authors.

A choice often made at level 1 is to introduce two free functions directly into the modified Poisson equation and the ‘slip relation’ (the transverse, traceless component of the field equations), as these are the expressions relevant to observables such as weak lensing of galaxies and measures of structure growth [40, 69]. This form of parameterization was introduced briefly in §1.2.3.1. One free function acts as an effective rescaling of Newton’s gravitational constant, whilst the other is defined as the ratio of the two potentials that describe the perturbed metric in the conformal Newtonian gauge. We will refer to this approach as ‘phenomenologically-based’. It is not the case that particular choices of the two free functions are designed to reproduce the field equations of specific modified gravity theories (except in a handful of special cases). Rather, in the phenomenological approach, the free functions should be considered as indicator flags for non-GR behaviour.

A disadvantage of phenomenological-type parameterizations is their tendency to obscure which regions of theory space they correspond to, because there is no direct mapping between the free functions and the parameters of a specific
model. It is therefore difficult to translate constraints on the two free functions into constraints on a given theory of modified gravity.

The PPF parameterization implements a different philosophy. It directly specifies quantities needed to describe a $4 \times 4$ tensor of scalar modifications to the linearly perturbed Einstein equations, from which the modified Poisson equation and slip relation can be derived. Individual gravity theories can be exactly recovered through particular choices of the free functions, as shown in the examples of chapter 2. Hence constraints on the free functions of the parameterization can be directly mapped onto a specific theory. The disadvantage of the PPF approach is that significantly more than two free functions are needed for the most general case.

Arguments can be made in favour of both the phenomenologically-based and PPF approaches. For a large class of theories the two strategies become equivalent in the so-called ‘quasistatic regime’, that is, intermediate distance scales where the time derivatives of perturbations can be neglected in comparison with their spatial derivatives - see §2.4.3 and §5.2 for further details.

However, the quasistatic approximation is not applicable on scales comparable to the horizon size. Whilst such large scales cannot be probed directly with galaxy surveys, they are nonetheless important for accurate calculation of the Integrated Sachs-Wolfe (ISW) effect and large-scale matter power spectrum. These quantities are frequently computed using Einstein-Boltzmann solvers such as CAMB [201] or CLASS [51] which evolve perturbations through horizon crossing.

The purpose of this chapter is to illustrate that the choice made for the level-1 structure of a parameterization has a significant influence on the constraints obtained on deviations from GR – a caveat that is commonly forgotten. We will demonstrate the differences between two possible parameterizations by generating some of the simpler perturbative observables - the CMB power spectra, in particular the ISW effect, matter power spectra, and the growth function $f(z)$. In §3.4.2 we further show that the choice made for the level-2 structure has an equally important influence on the constraints obtained.

### 3.2 Parameterizations

In this section we describe the two parameterizations that we apply in this chapter. Our goal here is not to advocate either of them, but rather to highlight the fact that they lead to different results. Parameterization A described below is derived as a special case of the general PPF form detailed in chapter 2. However, for the purposes of this chapter we can ignore its origin and simply regard the
resulting field equations as another example of a phenomenological-type parameterization – one with different choices made for the level-1 structure.

### 3.2.1 Parameterization A

In [30, 31, 284] a parameterization was proposed by writing the modifications as an additional tensor to the Einstein equations of GR:

\[ G_{\mu\nu} = 8\pi G_0 a^2 T^M_{\mu\nu} + a^2 U_{\mu\nu} \]

(3.1)

where the stress-energy tensor \( T^M_{\mu\nu} \) contains all the standard cosmologically-relevant fluids and the tensor \( U_{\mu\nu} \) may contain metric, matter and additional field degrees of freedom coming from a modified theory of gravity. In writing (3.1) it is assumed that all known matter fields which are part of \( T^M_{\mu\nu} \) couple to the same metric \( g_{\mu\nu} \), and that \( G_{\mu\nu} \) is the Einstein tensor of that same metric. This ensures that the stress-energy tensor of matter obeys its usual conservation equations, and hence \( U_{\mu\nu} \) is separately conserved.

The formalism proceeds by parameterizing around the linearly perturbed version of equation (3.1). To enable a direct comparison with parameterization B below we will specialize to the case of purely metric theories, that is, those for which the action is constructed from curvature invariants (e.g. \( f(R) \) gravity) or non-local invariants (e.g. [104]), without new field content.

The most general perturbations of \( U_{\mu\nu} \) that are allowed in a second-order metric-only theory are as follows [30, 31, 128, 284]:

\[-a^2 \delta U^0_0 = k^2 A_0 \dot{\hat{\Phi}}\]
\[-a^2 \delta U^0_i = k B_0 \dot{\hat{\Phi}}\]
\[a^2 \delta U^i_i = k^2 C_0 \dot{\Phi} + k C_1 \ddot{\Phi}\]
\[a^2 \delta U^i_j = D_0 \ddot{\Phi} + D_1 \dot{\Phi}\]

(3.2)

where dots denote derivatives with respect to conformal time \( \eta \) and \( k \) is the Fourier wavenumber. Recall from §2.2.3 that the gauge-invariant metric perturbation \( \hat{\Phi} \) reduces to the curvature perturbation \( \Phi \) in the conformal Newtonian (CN) gauge. In our conventions the CN gauge is defined by:

\[ ds^2 = -a^2(1 + 2\Psi)d\eta^2 + a^2(1 - 2\Phi)d\vec{x}^2 \]

(3.3)

The coefficients \( A_0 \ldots D_1 \) appearing above are functions of background quantities, \( A_0 = A_0(k, \eta) \) etc., and the factors of \( k \) are chosen such that \( A_0 \ldots D_1 \) are
dimensionless. However, these functions are not all independent. Perturbations of the Bianchi identity \( U^\mu_{\nu\mu} = 0 \) yield a set of additional constraint equations, similar to those obtained in §2.2.5. These constraints can be used to reduce the six free functions in equations (3.2) down to just two. One of these free functions is defined to be:

\[
\frac{D_1}{k} = \frac{\tilde{g}}{\mathcal{F}}, \quad \text{where} \quad \tilde{g} = -\frac{1}{2} \left( A_0 + 3 \frac{\mathcal{K}}{k} D_0 \right) \tag{3.4}
\]

The modified Poisson equation can then be written (in Fourier space):

\[
-k^2 \Phi = 4\pi G_0 \frac{a^2}{1 - \tilde{g}} \rho \Delta \equiv 4\pi G_{\text{eff}} a^2 \rho \Delta \tag{3.5}
\]

with \( G_0 \) denoting the canonical value of Newton’s gravitational constant as measured by a Cavendish experiment on the Earth, and where the gauge-invariant comoving density perturbation \( \rho \Delta \) is a summation over all conventional cosmologically-relevant fluids. The combination \( G_{\text{eff}} = G_0 \left(1 - \tilde{g}\right)^{-1} \) plays the role of a modified Newton’s constant.

We choose the second free function to be \( D_0 \), which we will hereafter relabel as \( \zeta = \zeta(k, \eta) \) to distinguish it from its appearance in the more general format of equations (4.29). \( \zeta \) appears in the ‘slip’ relation between the potentials \( \hat{\Phi} \) and \( \hat{\Psi} \) (which are equal to one another in GR):

\[
\hat{\Phi} - \hat{\Psi} = 8\pi G_0 a^2 (\rho + P) \Sigma + \zeta \hat{\Phi} + \frac{\tilde{g}}{\mathcal{F}} \hat{\Phi} \tag{3.6}
\]

The anisotropic stress of conventional matter \( \Sigma \) is negligible after the radiation-dominated era.

The key point here is that in parameterization A the slip relation and the modified Newton’s constant are not independent, as the function \( \tilde{g} \) appears in both. This special case of the PPF formalism does not capture the behaviour of many modified gravity theories, because we have not allowed for additional degrees of freedom to appear [30, 31]. It is, however, directly comparable to a phenomenological-type parameterization. Hence we will regard eqns.(3.5) and (3.6) simply as possible alternatives to eqns.(3.9) and (3.10).

### 3.2.2 Parameterization B

If one steps back and looks at the structure of the evolution equations for cosmological perturbations, one finds that it is enough to work with only two of the four Einstein field equations: the Poisson equation and the slip equation. From
3. AMBIGUITIES IN PARAMETERIZATIONS OF MODIFIED GRAVITY

a practical point of view, the only observable modifications to gravity (at least at the perturbative level) will be modifications to the these equations, which we can write in the following form

\[-k^2\Phi - 4\pi G_0 a^2 \rho \Delta = F_1\]  
\[(\Psi - \Phi) + 8\pi G_0 a^2 (\rho + P) \Sigma = F_2\]

where $F_1$ and $F_2$ are arbitrary functions of space and time. A simple ansatz (from considering the dimensions of the equations) is that $F_1 = \alpha k^2 \Phi$ and $F_2 = -\zeta \Phi$. Reorganizing the equations we find that they can be rewritten as

\[-k^2 \Phi = 4\pi G_{\text{eff}} a^2 \rho \Delta\]  
\[\Psi = -8\pi G_0 a^2 (\rho + P) \Sigma + (1 - \zeta) \Phi\]

Parameterization B is defined in the CN gauge, so that $\Phi$ and $\Psi$ replace $\hat{\Phi}$ and $\hat{\Psi}$ in eqns.(3.5) and (3.6). This formulation is equivalent up to small corrections to that used in [40]; their anisotropic stress is related to ours by $k^2 \Sigma = \sigma$. This parameterization suggests that, once the anisotropic stress has become negligible, a theory of modified gravity could potentially modify one of equations (3.9) or (3.10) whilst leaving the other unchanged. This behaviour does not arise analytically from theories below third order [30]; but given enough freedom in the functional ansatz the numerical behaviour that occurs in any theory can be realized with this parameterization. (We note that a handful of parameterizations that are instead optimized for weak lensing analysis have also been suggested [90, 286, 322]).

To demonstrate the differences that the alternative parameterization types can generate, we will use the same ansatz for the free functions $G_{\text{eff}}$ and $\zeta$ in both parameterizations A and B. Since we are focusing on modifications to gravity associated with the non-matter sector, an appropriate parameterization is an expansion in $\Omega_\Lambda(z)$. We therefore use a polynomial expansion up to third order in $\Omega_\Lambda(z)$.

### 3.3 Spectra

We have modified a version of the Boltzmann code CAMB to use the new equations specified in §3.2. We use variants which include both parameterizations A and B. In this section we show and analyze CMB and matter power spectra generated by these models. We will show that the two variants, even when set up with equivalently parameterized functions, generate considerably different spectra.
Figure 3.1 shows CMB spectra from parameterization A with variations to the linear term in the $\Omega_A(z)$ polynomial. Because we have modified the late-time behaviour of gravity whilst leaving the era before dark energy domination unchanged, the small-scale CMB (which depends mainly on what happens at recombination) is unaffected, but the large-scale features caused by the ISW effect vary. It can be seen that strengthening gravity by a small amount decreases the amplitude of the ISW effect by counteracting the suppressive effect of $\Lambda$ on structure formation, whereas weakening $G_{\text{eff}}$ increases the ISW effect. This difference is significant because the ISW effect is only weakly present observationally, see §1.2.1.6.

The behaviour of the spectra for a simple linear variation in $G_{\text{eff}}$ shows the same general trends for model B as those for model A shown in Fig. 3.1. The most important difference between the theories, however, appears when the two free functions of the parameterization interact to allow cancellations. This is discussed in detail below.

Examining the plot of linear variations in $G_{\text{eff}}$ alone, one might be tempted to conclude that modifications to the strength of gravity, as measured by the $G_{\text{eff}}$ parameter, are rather well-constrained using just CMB data. We wish to emphasize in this chapter that this is not so. Firstly, the various ways one can construct a theory mean that direct interpretation of the parameter, except on small scales, is not unique. Secondly and probably more importantly, the linear variation alone masks a richer phenomenology, which we will discuss more in the next section.

Fig. 3.2 shows the growth function, $f(z) = \frac{1}{\Omega} \frac{\dot{\Delta}_M}{\Delta_M}$ scaled by $\sigma_8$, which is measured by, for example, redshift-space distortion experiments. Results from the recent WiggleZ survey [47] and other data [146, 241, 285, 294] are also shown. These results provide slightly weaker constraints than the ISW, but near-future experiments should reach sensitivity levels good enough for precision tests of late-time gravity. The behaviour of the quantity $f(z)\sigma_8(z)$ in modified gravity theories will be further explored in chapter 5 of this thesis.

We do not display the matter power spectra because they are very mildly affected by the modifications of gravity considered here, except on very large scales far beyond the reach of galaxy surveys.

Figs. 3.1 and 3.2 demonstrate that modifications of about 20\% to the effective Newton’s constant in the Poisson equation can have effects that are easily detectable at large scales in the CMB. But it is the combined effects of $G_{\text{eff}}$ and $\zeta$ that are of most interest. Fig. 3.3 demonstrates how this combination can cancel even quite extreme individual effects. The curves in this plot have
3. AMBIGUITIES IN PARAMETERIZATIONS OF MODIFIED GRAVITY

\[ \ell (\ell + 1) C_{\ell} / 2\pi \left[ \mu K^2 \right] \]

\[ G_{\text{eff}} = G_0 \]
\[ G_{\text{eff}} = G_0(1 + 0.1 \Omega_{\Lambda}) \]
\[ G_{\text{eff}} = G_0(1 - 0.1 \Omega_{\Lambda}) \]
\[ G_{\text{eff}} = G_0(1 + 0.2 \Omega_{\Lambda}) \]
\[ G_{\text{eff}} = G_0(1 - 0.2 \Omega_{\Lambda}) \]

Figure 3.1: CMB power spectra for variations of the linear term of the Taylor expansion in \( \Omega_{\Lambda}(z) \). The Taylor expansion is the ‘level 2’ ansatz (see \S 3.1) for the function \( G_{\text{eff}} \) that appears in the Poisson equation, as defined in eq.(3.5).

\[ 8(z) \cdot f(z) \]

\[ G_{\text{eff}} = G_0 \]
\[ G_{\text{eff}} = G_0(1 + 0.1 \Omega_{\Lambda}) \]
\[ G_{\text{eff}} = G_0(1 - 0.1 \Omega_{\Lambda}) \]
\[ G_{\text{eff}} = G_0(1 + 0.2 \Omega_{\Lambda}) \]
\[ G_{\text{eff}} = G_0(1 - 0.2 \Omega_{\Lambda}) \]

Figure 3.2: The growth function \( f(z)\sigma_8(z) \) for the same models shown in Fig. 3.1.
$G_{\text{eff}}/G_0 = 1 + 5.3 \Omega_\Lambda - 1.1 \Omega_\Lambda^2 + 1.3 \Omega_\Lambda^3$, and $\zeta = 4.8 \Omega_\Lambda - 1.0 \Omega_\Lambda^2 + 6.6 \Omega_\Lambda^3$. The other cosmological parameters are within normal ranges, though not identical to those in Fig. 3.1. There is a near-total cancellation of the ISW effect in model A, but when the same parameter values are used in model B the resulting spectrum is completely ruled out by current data. Of course there will exist alternative parameter combinations in model B that result in similar cancellation effects, but the fact that they are different from those in model A is what we want to highlight here.

We are not, of course, advocating these models as actual cosmologies – they simply demonstrate that a large region of parameter space can be consistent with CMB data. Since other upcoming data sets like lensing and redshift-space distortions also constrain (different) combinations of the potentials there should be equivalent degeneracies present there.

The importance of other data sets in breaking the degeneracy in the CMB is illustrated in Fig. 3.4. The figure shows the growth rate $f(z)\sigma_8(z)$ for the extreme models described above. The behavior of the extreme case is once again very different between the two parameterizations.

### 3.4 Constraints

#### 3.4.1 Parameter Estimation

Now that we have demonstrated that different parameterization approaches can yield considerably different results for power spectra, we can check how these differences go forward into parameter estimation. Note that at this stage we are not attempting to find the tightest possible constraints on these parameterizations using all available data. In particular we are not using data from weak lensing, galaxy-ISW correlations, or growth rates.

Rather, we intend to demonstrate that the different parameterization schemes applied to the same data can produce significantly different constraints, even on quantities we usually regard as having a well-defined physical meaning – such as the the effective Newton’s constant controlling the growth of structures under gravity. We hope to show that headline constraints on the two free functions of parameterized approaches should be taken with a degree of caution.

The plots in this section are generated from Monte-Carlo Markov Chains running parameterizations A (equations (3.5) and (3.6)) and B (equations (3.9) and (3.10)). In each case, as in the previous section, the same ansatz is applied to the free functions $G_{\text{eff}}(z)$ and $\zeta(z)$. The same constraining data is used for
Figure 3.3: A demonstration of how extreme changes to the effective gravitational constant can be permitted by the CMB when they can be counteracted by a significant gravitational slip (see text for form of extreme curves). The cancellation does not occur when these parameters are used in model B. This delicate difference underscores the need for a rigorous treatment of modified functions.
both models A and B: the 7-year WMAP CMB data [199], the SDSS DR7 matter power spectrum [260], a prior $H_0 = 73.8 \pm 2.4$ [261], the BBN constraint $\Omega_b h^2 = 0.022 \pm 0.002$ [147], and the Union2 Supernova Ia data [9]. With these data sets the strongest constraining power comes from the ISW effect (or rather lack thereof) in the large-scale CMB temperature power spectrum.

### 3.4.2 Insufficient Ansatzes

The simplest ambiguity we can consider when constraining the free functions can arise if we use what might be termed an incomplete parameterization – one where the free functions are not flexible enough to explore the available parameter space. One example of this would be our restriction here to functions only of $z$, with no scale-dependence.

Fig. 3.5 illustrates another example of an incomplete parameterization. In that figure we show how the constraints on the theory depend on the number of terms in the polynomial expansions of $G_{\text{eff}}(\Omega_\Lambda)$ and $\zeta(\Omega_\Lambda)$. The larger contours

---

1It is not generally correct to use published supernovae constraints directly when using modified gravitational physics, since the calibration factors applied to them are cosmology-dependent; it is only because we leave our background evolution unchanged from GR that is is possible here.
use cubic models for the evolution, and the much smaller ones use only a linear term in the expansion. Both curves are for parameterization A, and the contours are for 68% and 95% probability mass.

Of course we do not claim that the cubic model is general enough to describe the free functions sufficiently – it is the comparison that we highlight. Whenever contours are presented for the free functions of some parameterized model one should bear in mind that adding only a little more freedom can change the areas of the error ellipses by a large factor. Hence all constraints of this form should be taken with a pinch of salt, so to speak.

![Graph showing constraints on free functions](image)

Figure 3.5: Constraints on the free functions, evaluated at redshift zero, in a model restricted to linear evolution in $\Omega_\Lambda$ (filled contours) and one with cubic evolution (line contours). The apparent constraints on $G_{\text{eff}}(z = 0)$ and $\zeta(z = 0)$ are misleadingly small in the more restrictive parameterization.

### 3.4.3 Variant Parameterization

The second result of this section is shown in Fig. 3.6; it demonstrates the parameterization-dependence of the constraints, independent of the ‘level 2’ ansatz defined in §3.1. Most clear is that the free functions are significantly less correlated in model A. In model B there is a simple and straightforward impact on the term whose
derivative is the ISW source:

$$\Phi + \Psi \bigg|_B \propto (\zeta - 2)G_{\text{eff}} \frac{a^2}{k^2} \rho \Delta$$

(3.11)

which yields a multiplicative correlation, whereas model A does not have this simple relation:

$$\Phi + \Psi \bigg|_A \propto \mathcal{H}^{-1} (G_{\text{eff}} - G_0) \frac{a^2}{k^2} \rho \Delta$$

$$+ \left[ (\zeta - 3)G_{\text{eff}} + \mathcal{H}^{-1} \left( 1 - \frac{G_0}{G_{\text{eff}}} \right) \dot{G}_{\text{eff}} + G_0 \right] \frac{a^2}{k^2} \rho \Delta$$

(3.12)

The edge of the curve for parameterization A in Fig. 3.6 cuts off rather sharply just below $G_{\text{eff}} = 1$. This is due to the impact of $G_{\text{eff}}$ and $\zeta$ on the growth of matter density perturbations, and the restrictions that they should obey in order to reproduce the observed amount of structure in the universe today. During the matter-dominated epoch density perturbations on subhorizon scales grow as $\delta \propto a^{p_A}$, where in parameterization A [29]:

$$p_A = \frac{1}{2} \left( 1 - \frac{3}{2} \frac{G_{\text{eff}}}{G_0} \right)$$

$$+ \frac{1}{4} \sqrt{9 \left( \frac{G_{\text{eff}}}{G_0} \right)^2 + 12 \frac{G_{\text{eff}}}{G_0} (3 - 2\zeta) - 20}$$

(3.13)

In parameterization B the exponent is different:

$$p_B = -\frac{1}{4} + \frac{1}{4} \sqrt{1 + 24 \frac{G_{\text{eff}}}{G_0} (1 - \zeta)}$$

(3.14)

The derivation of eqs.(3.13) and (3.14) will be presented in chapter 4. One can see that in the GR limit ($\zeta = 0, G_{\text{eff}} = G_0$) we recover $p_A = p_B = 1$, in agreement with the standard result for an Einstein-de Sitter universe. Equation (3.13) implies that in a universe with $\zeta = 0$ matter perturbations fail to grow during the matter-dominated epoch for values of $G_{\text{eff}}/G_0 < 0.5$. Hence we should expect this region of parameter space to be disfavoured by the matter power spectrum. The analogous boundaries in parameterization B will be different, which may explain the different shapes of the fall-offs of the distributions in Fig. 3.6. However, we ought to remember that equations (3.13) and (3.14) apply only in a pure Einstein-de Sitter setting. The boundaries in parameter space that they imply may not be obeyed rigidly in the real universe, due to growth during the radiation- and $\Lambda$-dominated eras.
For small deviations from GR $p_a \approx p_b$. This is because in EdS universes $\dot{\Phi} = 0$ so the slip equations (3.6) and (3.10) are the same. This does not apply when $\Lambda$ becomes important.

It is in the tails of the distributions where the difference between the two parameterizations can be most stark – this will be particularly important when we are trying to decide whether GR is threatened by any evidence of deviations from the $\Lambda$CDM model. The histograms in Fig. 3.7 demonstrate this most clearly. In addition, the spectral index is much less constrained in model A than in either GR or model B, because in that model we can more freely carry $G_{\text{eff}}$ at low redshift; so the matter power spectrum amplitude is not as constraining.

3.5 Conclusions

The first, and uncontroversial, issue we have noted here is the significant effect that using insufficiently free functions has on the constraints one obtains on modifications to the slip and Poisson relations. We now argue that there are two reasonable ways to get around this problem. One is to embrace the constraints, but to make them correspond to some regime in theory space that we wish to
model. This is the approach that we are working towards when developing the PPF parameterization presented in chapter 2. The alternative is to evade the constraints as far as is possible by making the functions completely free in both $k$ and $\eta$. This numerically challenging approach is the one taken in [252], where the data itself is given the freedom to choose the parameterization using a Principal Component Analysis (PCA) method.

The second argument we have advanced in this chapter is as follows. Different parameterizations – even at the level of where the free functions are placed analytically – can lead to radically different effects on the gravitational potentials $\Phi$ and $\Psi$, and their effects on cosmological observables. We have focussed on the ISW effect, which gives one of the most important constraints on parameterized approaches. In that case significant changes to the ISW plateau can be absent due to cancellations in these theories, even in the case of very large modifications. The details of this cancellation depend on which terms are included in the extension to the Poisson and slip equations. Given that the goal of these parameterized frameworks is to model and detect realistic deviations from GR, it makes sense to consider modifications to the equations that reflect the reasonable regions of theory space as well as possible, as advocated in [31]. The differences highlighted in this chapter between the parameterizations underline the delicacy of this question.

Numerical approaches with very free functions can generate the same effect by numerically finding the same cancelling solutions, but models with an explicit
parameterization will not in general do the same unless they are chosen for that specific purpose.

In the analysis we have undertaken here the differences between the parameterizations are smaller than the error bars on either of them. This is a temporary situation – the data will soon improve to the point where the differences are significant. In the longer term as they improve further they will clearly pick out which part of theory space is correct regardless of which parameterization is used, but in the intermediate phase before then the choice will matter. Furthermore, we have focussed on the ISW in the CMB, but the arguments follow through to other cosmological observables such as weak lensing, redshift-space distortions and any other probe of the gravitational potentials.

It is often argued that the purpose of these methods is not to model alternative theories, but simply to detect any deviation from GR+ΛCDM in as simple a way as possible [157, 252]. But, as we have shown in this chapter, two different parameterizations can lead to different constraints – calling into question the interpretation of any individual detection if not properly put into the context of the class of theories that are being considered.

Furthermore, the best way to maximize the ability of a method to find a signal like this is to model as closely as possible its expected characteristics, much like matched filter approach used in signal analysis.

This motivates the use of a more sophisticated parameterization of modified gravity, such as the PPF formalism introduced in chapter 2. An approach which carefully includes as many phenomena as possible in a model offers the best chance of detecting any small deviations from GR in the upcoming era of high-precision cosmic structure data.
Chapter 4

The Growth of Parameterized Perturbations in an Einstein-de Sitter Universe

In this chapter I take the first steps towards linking generalized, parameterized frameworks of the kind introduced in chapter 2 to observable measures of cosmic structure growth. I do this by studying the growth of matter overdensities in a toy model Einstein-de Sitter (EdS) universe, using a simplified version of the PPF parameterization. The EdS scenario provides a safe testbed in which to improve our understanding of the effects of modified gravitational field equations before we can tackle a more realistic case. I also consider here the effects of the reduced parameterization on the Integrated Sachs-Wolfe effect and canonically-conserved superhorizon perturbations.

This chapter should be viewed as a preparatory investigation, leading to the more sophisticated study of chapter 5.

4.1 Introduction

Historically, our theories of structure formation have been developed using matter-dominated cosmological models [55, 198, 219, 236, 237]. Though not appropriate for realistic calculations, a toy EdS cosmology remains an immensely useful testing ground for new theories; here, we can gain a qualitative understanding of perturbation evolution whilst our knowledge of the background expansion remains on a firm footing. Therefore, maintaining the principle of caution advertised above, we will consider the effects of parameterized gravity on an EdS universe (which contains pressureless matter only). This will be an excellent approximation to
the matter epoch of the real universe, and has the added advantage that analytic solutions are achievable in some cases due to the simple properties of pressureless matter.

Despite the wide variety of modified gravity theories present in the literature, a survey of field equations reveals some common features – this is the motivation behind the phenomenological parameterization examined in chapter 3. In many theories the gravitational constant appearing in the Poisson equation acquires a time-dependence and/or scale-dependence [225], and the Newtonian gravitational potential $\Psi$ and curvature perturbation $\Phi$ (in the conformal Newtonian gauge) are not equal as they are in GR. We can ask what typical effects these generic features might have on observables: for example, in addition to affecting the growth rate of structure, distortion of gravitational potentials will imprint secondary anisotropies on the CMB. However, we will assume that modifications to GR must be negligible at very early times, to avoid significantly impacting the primary CMB and the sensitive reaction rates of Big Bang Nucleosynthesis [34, 67, 259]; see [57] for consideration of scalar-field models that are non-negligible at the time of recombination.

This chapter investigates the evolution of cosmological perturbations in the parameterized framework implemented in [30, 128, 284, 325], for an EdS background. Note that this is a reduced version of the PPF framework of chapter 2. §4.2 introduces the necessary formalism in the context of theories that are constructed purely from metric quantities. In §4.3 we use this framework to calculate the evolution of density perturbations on intermediate and large scales. Two other quantities of interest are also calculated: the growth rate $f(z)$, and the Integrated Sachs-Wolfe effect that is induced in such gravity theories (note that this is not the same as the late-time Integrated Sachs-Wolfe effect that occurs in a $\Lambda$-dominated era, as we are considering a matter-only universe here).

In §4.4 we take the first steps towards implementing a similar treatment of gravitational theories that introduce additional scalar degrees of freedom. New degrees of freedom cause a significant increase in complexity; we will deal with this by implementing a parameterized effective fluid approach known as ‘Generalized Dark Matter’ (GDM) [167]. The approach treats all scalar degrees of freedom as effective fluids; we will focus on the case where the effective pressure of the new exotic fluid is negligible (this is necessary to maintain an EdS background). See [39] for a similar approach to modified gravity.

The restriction on the effective equation of state excludes some classes of theories from our analysis, such as $f(R)$ gravity, but is applicable in other cases; we drop the restriction again in §4.5. In §4.5 we consider a perturbation that is
conserved on super-horizon scales in GR, and ask under what conditions this fact remains true in parameterized modified gravity. The conclusions of this work are presented in §4.6.

The purpose of this chapter is to develop an understanding of parameterized gravity, not to pursue accurate calculations for the real universe. Hence the plots and trial parameter values used here are intended to be illustrative rather than realistic.

4.2 Parameterization of Metric-Only Gravity Theories

For convenience we here remind the reader of the two parameterizations introduced in chapter 3, which we labelled as A and B.

A common parameterization choice is to introduce a free function that describes any time- or scale-variation of the gravitational constant in the Poisson equation [40, 45, 90, 158, 167, 252]:

\[-2k^2\Phi = \kappa a^2 \mu(a, k) \rho \Delta\]  

(4.1)

where \(\kappa = 8\pi G_0\), \(G_0\) is the canonical value of Newton’s gravitational constant, \(\Delta = \delta + 3H\theta\) is a comoving gauge-invariant density perturbation and \(\theta\) is the velocity potential of a fluid defined by \(v_i = \nabla_i \theta\). Often a second free function is used to describe the ratio of the two conformal Newtonian potentials, \(\gamma(a, k) = \Phi/\Psi\).

By redefining \(\zeta = (1 - 1/\gamma)\) this can be rewritten:

\[\Phi - \Psi = \zeta(a, k)\Phi\]  

(4.2)

The arguments of the ‘modified gravity functions’ (MGFs) \(\mu\) and \(\zeta\) will be suppressed hereafter. This form of the slip relation corresponds to parameterization B introduced in §3.2. In [30] it was argued that a parameterization of this type is only applicable in the quasistatic regime, and that in the case of purely metric theories it implicitly corresponds to higher-derivative theories. This is true if one assumes that eqn.(4.2) is an exact ‘template’ for the slip relation of a modified gravity theory, which means that \(\zeta\) can only be a function of homogeneous background quantities. If instead one is prepared to let \(\zeta\) be a function of non-homogeneous environmental variables and initial conditions then the slip relation may not necessarily imply a higher-derivative theory; it becomes impossible to ascertain the derivative order of the theories being parameterized without further
4. THE GROWTH OF PARAMETERIZED PERTURBATIONS IN AN EINSTEIN-DE SITTER UNIVERSE

information [253].

In [30] an alternative format was suggested, in which the order of derivatives in the field equations is made explicit. In this alternative parameterization a metric theory with an unmodified background (ie. one governed by the usual Friedmann equations) incurs extra constraint equations (see Table 4.1 in §4.7.2), which force the slip relation to be [128, 284]:

\[ \Phi - \Psi = \zeta \Phi + \left( \mu - 1 \right) \frac{H}{\sqrt{\mu}} \phi \] (4.3)

This is the slip relation belonging to parameterization A of the previous chapter. The key feature to note here is that in parameterization A the MGF \( \mu \) appears in both the Poisson and slip equations, unlike parameterization B.

However, the purpose of this chapter is not to discuss the subtleties of parameterization choice at length; we wish to keep our results as general as possible. We can treat both of parameterizations A and B simultaneously by adopting the following Poisson and slip equations:

\[ -2k^2 \phi = \kappa a^2 \mu P \rho_M \Delta_M \] (4.4)

\[ \phi - \Psi = \zeta \phi + \frac{\mu_s - 1}{\sqrt{\mu_s}} \phi \] (4.5)

To recover parameterization B we set \( \mu_s = 1 \) but keep \( \mu P \) general. To recover parameterization A we set \( \mu_s = \mu P \). Note that \( \mu_s \) is related to the function \( \tilde{g} \) in eqs.(3.4), (3.5) and [30] by \( \mu_s = (1 - \tilde{g})^{-1} \). Throughout this section we will sometimes leave results expressed in terms of the three functions \( \mu P, \mu_s \) and \( \zeta \).

We wish to emphasize from the outset that there are only really two independent functions, and all expressions should be evaluated in either the A-type or B-type instance. We write our expressions in this general format because it can be instructive to see whether the modified terms have their origin in the Poisson equation (indicated by the presence of \( \mu P \)) or the slip relation (indicated by \( \mu_s \) and \( \zeta \)).

What theories map onto equations (4.4) and (4.5)? The answer is ‘very few’, which is a cause for concern given that the above forms are often used to obtain constraints on modified gravity from current data. To map exactly onto these parameterizations a theory must stem from an action that is constructed only from curvature invariants and leads to fields equations that contain at most second- or third-order time derivatives (for parameterizations A and B respectively). By the immense power of Lovelock’s theorem [211, 212] such a theory can only differ from GR if it introduces either nonlocality or spacetimes of dimension greater
than four, see §1.3.1. This is a very restricted class of theories, though examples do exist [63, 100, 104].

The limitations described above can be relaxed if one is satisfied with an approximate correspondence between theories and parameterization, rather than an exact one. For example, it is frequently assumed that the perturbed Einstein equations will retain the form of eqns.(4.1) and (4.2) in theories with additional scalar degrees of freedom. This cannot be exactly true; any new scalar coupled to gravity will modify the zeroth-order Einstein equations in some way, and we expect to see perturbations of the new field appearing in the linearized Einstein equations, see eqs.(2.18-2.21). However, the form of eqns.(4.1) and (4.2) is retained within a limited range of distance scales for some theories, as will be discussed in chapter 5 [14, 30, 193, 251, 270].

To avoid such approximations we will proceed by taking eqns.(4.4) and (4.5) at ‘face value’, i.e. assuming that there are no new scalar degrees of freedom hidden behind them. This is the assumption that is implicitly being made if equations such as (4.1) and (4.2) are implemented in an Einstein-Boltzmann solver [112, 157, 158, 319] and used to generate ISW and matter power spectra. In §4.4 we will introduce the afore-mentioned GDM extension to the parameterization of eqns.(4.4) and (4.5), to account for additional scalars explicitly.

### 4.3 Density Perturbations

In this section we will consider how the growth of cold dark matter (CDM) density perturbations in an EdS universe is influenced by the MGFs. This is a model for the matter-dominated epoch of the real universe.

We will assume that any time-variation of the MGFs during the matter-dominated epoch must be very small in order to prevent them from evolving to a region of parameter space that would cause conflict with observations. We will therefore take the time derivatives of $\mu_P, \mu_s$ and $\zeta$ to be negligible in comparison to the rate of evolution of other variables (note that this approximation is not generally made in the full PPF parameterization).

Although we are neglecting their time-dependence, $\mu_P$ and $\zeta$ may still contain scale-dependence. They are dimensionless functions, but scale-dependence can appear as a ratio to the Hubble scale or to some special scale that arises in a given theory, i.e. $k/k_*$. An example of such a privileged scale is the Compton wavelength of the scalaron in $f(R)$ gravity [288].

The fluid conservation equations for CDM energy density and momentum
perturbations are given by:

\[ \dot{\delta}_M = -k^2 \theta_M + 3\dot{\Phi} \quad (4.6) \]

\[ \dot{\theta}_M = -3\mathcal{H}\theta_M + \Psi \quad (4.7) \]

where \( \theta_M \) is the velocity potential. Differentiating eqn.(4.6) and combining with eqn.(4.7) leads to a second-order equation for \( \delta_M \):

\[ \ddot{\delta}_M + \mathcal{H}\dot{\delta}_M - 3\ddot{\Phi} - 3\mathcal{H}\dot{\Phi} + k^2\Psi = 0 \quad (4.8) \]

This is the same as found in GR. However, differences from GR arise when we use a non-trivial slip relation to eliminate \( \Psi \). Substituting eqn.(4.5) into eqn.(4.8):

\[ \ddot{\delta}_M + \mathcal{H}\dot{\delta}_M - 3\ddot{\Phi} - 3\mathcal{H}\dot{\Phi} \left[ 1 + \frac{k^2}{3\mathcal{H}^2} \left( \frac{\mu_s - 1}{\mu_s} \right) \right] + k^2(1 - \zeta)\Phi = 0 \quad (4.9) \]

Before studying the behaviour of this equation it is useful to delineate a hierarchy of distance scales:

1. Non-linear scales at which clusters and galaxies form.
2. Quasistatic scales on which time derivatives of perturbations can be neglected in comparison to their spatial derivatives.
3. Larger scales on which the above approximation is no longer valid, but are still well within the horizon.
4. Scales that are greater than our observable horizon.

We will consider the solutions of eqn.(4.9) in regions 3 and 4. In fact it is possible to derive a single equation for \( \Phi \) that is valid in both regions 3 and 4, and then use the Poisson equation to relate its solutions to \( \delta_M \) (we thank C. Skordis for pointing this out). We will use an equivalent method that is simpler but a little less elegant.

### 4.3.1 Subhorizon Scales

In region 3 described above we can approximate \( \mathcal{H}/k \ll 1 \) and \( \Delta_M \approx \delta_M \) since \( |\theta_M| \sim |v_M|/|k| \) is small. We use derivatives of eqn.(4.4) to eliminate \( \Phi \) from eqn.(4.9):

\[ \ddot{\delta}_M + \mathcal{H}\dot{\delta}_M \left[ 1 + \frac{3}{2} \mu_P \frac{(\mu_s - 1)}{\mu_s} \right] - \frac{3}{2} \mathcal{H}^2 \mu_P \delta_M \left[ 1 - \zeta + \frac{(\mu_s - 1)}{\mu_s} \right] = 0 \quad (4.10) \]
where we have used the result $\dot{H} = -\frac{1}{2} \mathcal{H}^2$ in an EdS universe. We write the solutions of this equation in the form

$$\delta_M = N^+ a^{n^+} + N^- a^{n^-}$$

(4.11)

where $N^+$ and $N^-$ are constants and the power-law indices are:

$$n^\pm = -\frac{1}{2} \left( 1 + 3\mu_P \left( \frac{\mu_s - 1}{\mu_s} \right) \right) \pm \frac{1}{2} \left[ 9\frac{\mu_P^2}{\mu_s^2} (\mu_s - 1)^2 - 30 \frac{\mu_P}{\mu_s} + 6 \mu_P (9 - 4\zeta) + 1 \right]^{\frac{1}{2}}$$

(4.12)

In parameterization A this reduces to

$$n^A_\pm = \left( 1 - \frac{3}{2} \mu \right) \pm \frac{1}{2} \sqrt{9\mu^2 + 12\mu(3 - 2\zeta) - 20}$$

(4.13)

whereas in parameterization B it becomes

$$n^B_\pm = -\frac{1}{2} \pm \frac{1}{2} \sqrt{1 + 24\mu_P(1 - \zeta)}$$

(4.14)

It can be verified that in the limit $\mu_s = \mu_P = 1$ and $\zeta = 0$ eqn.(4.12) recovers the GR result $\delta_M \propto a$. We can see immediately that in both parameterizations there is a term $-24\mu_P\zeta$ that leads to degeneracy between the effects of the individual MGFs. Note that $\zeta$ only appears within this degenerate combination, so it cannot significantly impact growth if $\mu_P$ is small. We note that if a $\mu$-like MGF is implemented in the Poisson equation for $\Psi$ instead of $\Phi$ then this degeneracy does not arise in the parameterization B case [157, 158, 251].

When $n^\pm$ are imaginary the solutions for $\delta_M$ are damped oscillations. However, since our calculation has neglected the effects of baryons or radiation this oscillatory behaviour simply indicates unphysical solutions rather than anything meaningful. To have at least one growing mode we need $n^+$ to be positive, for which the relevant condition is:

$$\mu_s(2 - \zeta) > 1$$

(4.15)

We expect that an approximate version of this bound should be obeyed in the real universe, in order to reproduce the observed matter power spectrum – see §4.3.3. However, in the real universe the hard bound of eqn.(4.15) will be softened by contributions to growth from the radiation and $\Lambda$-dominated eras. Note that there is no restriction that prevents $\zeta$ from adopting negative values.
4. THE GROWTH OF PARAMETERIZED PERTURBATIONS IN AN EINSTEIN-DE SITTER UNIVERSE

We wish to understand the physical mechanisms through which the MGFs are exerting their influence on subhorizon scales. We can get a feel for this by thinking about eqn.(4.10) in the context of a simple mechanical system. The last term on the righthand side represents a time-dependent forcing that drives the collapse of density perturbations. The $\dot{\delta}_M$ term is analogous to a frictional force, which in familiar physical situations always acts to oppose motion; its magnitude decreases with time due to the factor of $H$. Since we are parameterizing around a EdS background the evolution of $H$ is unaffected by the MGFs, and hence cannot be contributing to deviations from GR.

The overall magnitude of the driving term is controlled by $\mu_P$. This intuitively makes sense – if we increase the gravitational coupling strength then we expect structures to collapse faster. Less intuitive is the appearance of $\mu_s$ and $\zeta$ in the driving term, which have the ability to change its sign. We will assume throughout that $\mu_P, \mu_s > 0$ always to maintain agreement with our physical notion of attractive gravity, but note that there is no such restriction on $\zeta$. The condition for the driving force to maintain the same direction as it has in GR is exactly eqn.(4.15). It is interesting to see that in parameterization A $\mu_s$ and $\zeta$ can have counteracting effects on the driving term. Qualitatively, a negative $\zeta$-value with large magnitude enables one to weaken $\mu_P$ considerably whilst maintaining growth during a matter-dominated epoch. If parameterization B is adopted this effect does not exist because the modification to the Poisson equation has no influence on the sign of the driving term.

The friction term has a somewhat simpler behaviour, as it is unaffected by $\zeta$. In parameterization B the friction is unchanged from GR, but in parameterization A $\mu = \mu_P = \mu_s$ has the power to enhance or suppress friction effects. A value of $\mu > 1$ acts to increase friction, but simultaneously strengthens the driving term that drives perturbations to collapse. One expects that some degree of cancellation between these two effects may be possible, even without the extra freedom provided by $\zeta$.

4.3.2 Superhorizon Scales

On large scales we ought not to make the approximation $\Delta_M \approx \delta_M$, as the magnitude of the velocity potential is not negligible. We will adopt a different strategy, using the linearized ‘$\delta C^0_0$’ equation to solve for $\Phi$ as a proxy for $\delta_M$.

We first adopt a phenomenological approach (in the spirit of parameterization B), and assume that Newton’s constant is identically modified in all perturbed
field equations. We then have (neglecting a small term proportional to $k^2$):

$$-\mathcal{H}(\dot{\Phi} + \dot{\mathcal{H}}\Psi) = \frac{\kappa a^2}{6} \rho_M \mu_P \delta_M = \frac{3\mathcal{H}^2}{2} \mu_P \delta_M$$  \hspace{1cm} (4.16)

Differentiating, and using that $\delta_M \approx 3\Phi$ on large scales (from eqn.(4.6)):

$$\ddot{\Phi} + \frac{\mathcal{H}}{2} (1 + 3\mu_P) \dot{\Phi} + \mathcal{H} \dot{\Psi} = 0$$  \hspace{1cm} (4.17)

Using the slip relation (eq.(4.5)):

$$\ddot{\Phi} + \dot{\Phi} \frac{\mathcal{H}}{2} [3 + 3\mu_P - 2\zeta] = 0$$  \hspace{1cm} (4.18)

The power-law solutions are $\Phi \propto a^{p^{\pm}}$, with $p^+ = 0$ and $p^- = (2\zeta - 2 - 3\mu_P)$. On these scales $\delta_M$ follows the behaviour of $\Phi$ up to a constant offset, which can be set to zero by initial conditions. So the dominant mode – constant potential outside the horizon and $\delta_M$ frozen – is the same as in GR, see §1.1.4.2. However, the decaying mode is affected by the MGFs. For $\zeta = 0$, increasing $\mu_P$ will result in faster decay. This seems somewhat counterintuitive, since one would usually associate an increase in gravitational strength with reduced decay of overdensities. However, since we are working on superhorizon scales gauge issues may invalidate our physical notions of gravitational growth and decay.

In parameterization A the linearized Friedmann equation is

$$-6\mathcal{H}(\dot{\Phi} + \dot{\mathcal{H}}\Psi) = \kappa a^2 \rho_M \delta_M + A_0 k^2 \Phi$$  \hspace{1cm} (4.19)

where

$$A_0 = -2 \left( \frac{\mu_P - 1}{\mu_P} \right) \left( 1 + \frac{\mathcal{H}^2}{Q} \right) + 2\zeta \frac{\mathcal{H}^2}{Q}$$

with $Q = \mathcal{H}^2 + \frac{k^2}{3} - \dot{\mathcal{H}}$

The derivation of the above expressions is given in §4.7.2.

Note that $\mu_P$ does not feature explicitly in eqn.(4.19). Newton’s constant is not modified directly – instead one considers all possible additional terms that could appear in the linearized Einstein equations, which can be determined up to a dimensionless function of background quantities. In the case of a second-order metric-only theory the only terms that can be added to the linearized Friedmann and ‘$\delta G^0_i$’ equations are proportional to $\Phi$, which can be absorbed into an effective Newton’s constant.
4. THE GROWTH OF PARAMETERIZED PERTURBATIONS IN AN EINSTEIN-DE SITTER UNIVERSE

Figure 4.1: Joint constraints on the slip parameter $\zeta$ and $\mu_P$ at $z = 0$, for parameterization A (black lines) and B (filled green). In both cases 68% and 95% contours are shown.

On very large scales $\lim_{k \to 0} (k^2 A_0) = 0$, so we can neglect the extra term on the RHS of eq.(4.19). Repeating the steps we took for the parameterization B-like case we obtain $p^+ = 0$, $p^- = \mu_P (2\zeta - 5)$. The values of $p^-$ in the two parameterizations converge as one tends to the GR limit, as of course they must.

However, the decaying mode is not hugely interesting as it is unobservable (unless there are some very radical modifications to GR involved). The important result here is that the potential remains constant on superhorizon scales, as usual.

4.3.3 Connection to Constraints

To what extent are our results for an idealized EdS model borne out in the real universe? Fig. 4.1 was shown in §3.4.3, but we reproduce it here for convenience. It shows the joint constraints on the MGFs $\mu_P$ and $\zeta$ obtained using the following data sets: the 7-year WMAP CMB data [199], the SDSS DR7 matter power spectrum [260], a prior $H_0 = 73.8 \pm 2.4$ [261], the BBN constraint $\Omega_b h^2 = 0.022 \pm 0.002$ [147], and the Union2 Supernova Ia data [9]. The differences between the two sets of contours were discussed in detail in the previous chapter. Here our main interest is the extent to which the constraints reflect the analytic solutions of the previous two subsections.
Given the restriction that $\mu_\ast$ must be positive, eqn.(4.15) tells us to expect that $\zeta < 2$ on subhorizon scales in parameterization A. For $\zeta = 0$ we must have $\mu_P = \mu_\ast > 1/2$ for density perturbations to grow during the matter era, which corresponds to the approximate location of the near-vertical black contour in Fig. 4.1. This contour is not the result of any artificially-imposed boundary, but delineates a very sharp fall-off in the likelihood distribution for $\mu_P$ (see Fig. 3.7). In contrast, in parameterization B any $\mu_P > 0$ permits growing modes for $\zeta = 0$, giving rise to the more gradual fall-off shown by the shaded contours.

However, eqn.(4.15) also implies that in parameterization B $\zeta < 1$ is necessary for growth of CDM density perturbations on subhorizon scales (since $\mu_\ast = 1$ in parameterization B), which is contradicted by Fig. 4.1. This is not too surprising – we expect the simple bounds implied by our EdS example to be blurred by the complexities of a realistic cosmological model. If models in the region of parameter space $\zeta > 1$ experience sufficient growth during radiation- and $\Lambda$-dominated epochs, or on scales outside the validity of eqn.(4.14), then they will not be excluded by an MCMC analysis.

The degeneracy between $\mu_P$ and $\zeta$ is visible in both sets of contours, but it is more pronounced in parameterization B. This is because in parameterization A the quadratic term in eqn.(4.13) makes it more difficult to compensate for the effects of a large $\mu_P$ with a small $\zeta$, or vice-versa.

### 4.3.4 Other Growth Observables

#### Integrated Sachs-Wolfe Effect

A cosmological model governed by GR will not experience an Integrated Sachs-Wolfe (ISW) effect [265] during a matter-dominated epoch. The kernel of interest for the ISW effect is $\dot{\Phi} + \dot{\Psi}$ (in the conformal Newtonian gauge), which in GR is equal to $2\dot{\Phi}$. From the standard Poisson equation one has, on scales well below the horizon:

$$k^2 \Phi \propto a^2 \rho_M \delta_M$$  \hspace{1cm} (4.20)

Energy-momentum conservation gives $\rho_M \propto a^{-3}$, whilst in GR $\delta_M \propto a$, leaving $\Phi$ with zero time-dependence. The situation only changes for $z \lesssim 0.5$ when $\Lambda$ begins to suppress the growth rate of $\delta_M$ [87].

In modified gravity this behaviour is affected in two ways: a non-trivial slip relation will cause the ISW kernel to differ from $2\dot{\Phi}$, and the scaling of $\delta_M$ with $a$ will be altered. Then generically we expect a non-zero ISW effect, even during a matter-dominated phase of the universe [136, 317]. We will calculate the con-
4. THE GROWTH OF PARAMETERIZED PERTURBATIONS IN
AN EINSTEIN-DE SITTER UNIVERSE

tribution to the CMB temperature power spectrum of this effect for the metric
theories discussed in §4.2.

We begin by using the slip relation (eq.(4.5)) to express the ISW kernel purely
in terms of \( \Phi \):

\[
\dot{\phi} + \dot{\psi} = \frac{1}{2} \frac{\dot{\phi}}{3 - 2\zeta + \frac{1}{\mu_s}} - \dot{\phi} \left( \frac{\mu_s - 1}{3\mu_s} \right) \quad (4.21)
\]

Using the Poisson equation to connect \( \Phi \) and \( \delta_M \), we have (discarding the decaying
mode):

\[
\Phi = M(k)\eta^{(n^+ - 2)}
\]

where

\[
M(k) = -\frac{k\delta M_0(k)\mu_P\rho_{M_0}}{2k^2} \quad (4.22)
\]

\( n^+ \) is given by eqn.(4.12) and density perturbations are normalized by their
present values i.e. \( \delta_M(k, \eta) = \delta_{M_0}(k)\eta^{n^+} \). Generally one expects the growth
rate of density perturbations to be scale-dependent in a modified gravity theory,
whereas in GR all linear subhorizon modes grow at the same rate. However, we
are modelling a region of theory space close to GR and hence we will assume neg-
ligible variation of the growth rate over the range of \( k \) relevant to observations of
the ISW plateau.

So the ISW kernel is:

\[
\dot{\phi} + \dot{\psi} = \frac{M(k)}{2} (n^+ - 2)\eta^{(n^+ - 3)} \left[ 6 - n^+ - 2\zeta + \frac{1}{\mu_s}(n^+ - 2) \right] \quad (4.23)
\]

Note that in the GR limit \( n^+ = 2 \) the above expression vanishes as expected.

Next we need to compute the power spectrum of this modified gravity-induced
ISW effect. The expression to be evaluated is (recall §1.2.1.4):

\[
C_l = \frac{2}{\pi} \int_0^\infty dk k^2 P(k) \left| \frac{\Theta_l(k)}{\delta_{M_0}(k)} \right|^2 \quad (4.24)
\]

The temperature perturbation observed today, \( \Theta_l(k, \eta_0) \), consists of three parts:

\[
\Theta(k, \eta_0) = \text{monopole term} + \text{dipole term} \quad (4.25)
\]

\[
+ \int_0^{\eta_0} d\eta e^{-\tau} \left[ \dot{\phi}(k, \eta) + \dot{\psi}(k, \eta) \right] j_l[k(\eta_0 - \eta)]
\]

Let us focus on the dominant contribution to the ISW power spectrum which
comes from the \( (\dot{\phi} + \dot{\psi})^2 \) term (the square comes from eq.(4.24)), and ignore the
cross terms with the monopole and dipole. The cross-terms should only yield
small corrections because the monopole and dipole terms are evaluated at the
time of last scattering, and hence affect different $l$-values from the subsequent ISW
effect. We will take the visibility function $e^{-\tau}$ to be a Heaviside step function at
recombination. Of course, if we are considering a truly EdS universe then there is
no recombination event or time of last scattering, but this detail is irrelevant – we
are only interested in modelling the real universe well after recombination. In the
real universe recombination occurs sufficiently early that for practical purposes
we can take $\eta_* \approx 0$.

The derivation of §1.1.4.2 relates the power spectrum of density fluctuations
today to the power spectrum of the primordial potential. The only modification
to this that occurs in our theory is a factor of $\mu_P^2$, which arises when we relate
density fluctuations to $\Phi$ via the Poisson equation. However, this is cancelled
by a factor of $\mu_P^2$ in $M(k)$ defined in eq. (4.22). The scales of interest to us are
sufficiently large that we can set the transfer function $T(k) \sim 1$. Then, for a
Harrison-Zel’dovich spectrum we find

$$C_l^{I\text{SW,sq}} = \frac{9\pi}{4} (n^+ - 2)^2 \left( 6 - n^+ - 2 \zeta + \frac{1}{\mu_*(n^+ - 2)} \right)^2 \delta_H^2$$

$$\times \int_0^\infty \frac{dk}{k} \left[ \int_0^{\eta_0} \eta^{(n^+-3)} j_l[k(\eta_0 - \eta)] d\eta \right]^2$$

(4.26)

Recall from §1.1.4.2 that $\delta_H$ is the amplitude of primordial perturbations at
horizon-crossing during a single-field slow-roll inflation scenario. The superscript
‘sq’ on $C_l$ reminds us that we have only evaluated the ISW-squared term and not
the cross-terms.

A plot of this power spectrum for several combinations of $\mu$ and $\zeta$ in param-
eterization A is shown in Fig. 4.2. The normalization of the y-axis is arbitrary
because we have not attempted an accurate calculation of the $C_l$s; we are more
interested in how the shape and amplitude of the power spectrum is affected by
different parameters. Note that our subhorizon solution for $\Phi$ is likely to become
invalid for the largest scales, hence the region $l \gtrsim 10$ in Fig. 4.2 is not fully
accurate.

Provided that $\mu_P$ is not unusually small, $\zeta$ can have a significant effect on the
amplitude of the power spectrum. For example, compare the curves with $\{\mu_P, \zeta\}$
equal to $\{1, -0.25\}$ and $\{2, 0\}$ – a small change in $\zeta$ dominates over a large change
in $\mu_P$. The models $\{2.0, -0.25\}$ and $\{3.0, 0.0\}$ have the same value of $n^+$, and
hence the same spectral shape. It may seem a little surprising that a model with
parameters $\{1.0, -0.25\}$ predicts a larger ISW effect than one with $\{2.0, -0.25\}$
for $l > 6$; this is because a larger value of $n^+$ shifts the dominant contribution to
4. THE GROWTH OF PARAMETERIZED PERTURBATIONS IN AN EINSTEIN-DE SITTER UNIVERSE

Figure 4.2: Low-$l$ power spectrum of the ISW effect induced by a theory of modified gravity constructed in parameterization A (see §4.2 for details). The $y$-axis scaling is arbitrary. The region $l \lesssim 10$ is expected to be subject to corrections.

the integral in eqn.(4.26) to later times, therefore shifting the corresponding ISW effects towards larger scales.

In general, however, it is difficult to cleanly disentangle the effects of $\mu_P$ and $\zeta$ on the ISW power spectrum, because they appear in a degenerate combination inside $n^+$. 

Growth Rate

With the results of §4.3.1 in hand we can trivially compute the growth rate $f(z)$ defined in eq.(1.66). If we assume that we are not dealing with very radical departures from GR then any scale-dependence is likely to be small over a restricted range of $k$, and can be neglected. On the scales of region 3 (see the beginning of this section) where $\Delta_M \sim \delta_M$, the growth function is simply $f(z) = \frac{n^+}{2}$, where $n^+$ is given by eqn.(4.12).
4.4 Theories With Additional Degrees of Freedom

As discussed in §1.3.2, many gravity theories introduce degrees of freedom (d.o.f.) other than perturbations of the metric. For model-independent constraints to be genuinely feasible we need to construct a parameterization that is able to accommodate such d.o.f.

In §4.2 the parameterization-A-based approach was to add to each linearized Einstein equation all possible terms that could appear in the context of a metric-only second-order theory, which amounted to terms in $\Phi$ and $\dot{\Phi}$ multiplied by some function of background quantities (see §4.7.2 for the explicit forms of these). When we extend this framework to include extra d.o.f. two new types of terms appear. Firstly we must allow for perturbations of the d.o.f themselves – for example, in a Brans-Dicke theory [56] one expects the perturbations of the scalar field $\delta \phi$, $\delta \dot{\phi}$ and $\delta \ddot{\phi}$. Such new scalars are awkward to work with directly without knowing their underlying equations of motion. Therefore we will adopt an alternative approach based on a scheme by Hu [167] and treat this first type of additional term as perturbations of an effective fluid. A similar approach was adopted in [39, 190, 220].

In a fully general case one should allow the effective fluid to have a time-varying equation of state, non-adiabatic perturbations and significant anisotropic stress. This would render our simple EdS model invalid by modifying the evolution of the cosmological background. To proceed with our analytic treatment we will sacrifice some generality by setting the equation of state parameter of the effective fluid, $\omega_{\text{eff}}$, to zero. It then contributes to the zeroth-order Friedmann equation and supports density and velocity perturbations, but has negligible pressure and anisotropic stress.

This approximation is not unreasonable; for example, in Hořava-Lifshitz gravity (§2.3.5) a non-local Hamiltonian constraint gives rise to an integration constant which may be interpreted as a pressureless fluid. The additions to the Einstein equations in linear Einstein-Aether (§2.3.2) theory also behave as a pressureless fluid during an EdS phase, as can be seen using eqns.(15) and (16) of [324]. The effective fluid of Eddington-Born-Infeld gravity (§2.3.4) can experience a CDM-like phase too. However, making this restriction does exclude some important classes of theories, such as $f(R)$ gravity (except for special choices of $f(R)$).

The second type of new term that arises when extra d.o.f. are included is a combination of metric potentials that was introduced (in a gauge-invariant form)
in §2.2.3, as an ‘ingredient’ of the PPF parameterization:

$$\Gamma = \frac{1}{k} \left( \dot{\Phi} + \mathcal{H} \Psi \right)$$ (4.27)

If the background equations are fixed to be EdS with no new fields, as in §4.3, then the gauge-invariant version of $\Gamma$ cannot appear in the linearized field equations because it introduces higher-order time derivatives of the scale factor. However, when new d.o.f. are added these unwanted time derivatives can be cancelled by perturbation of the new scalar. This somewhat technical point was demonstrated explicitly for scalar-tensor theories in appendix A of [30].

Concretely, then, we write the Einstein equation in the form:

$$G_{\mu \nu} = 8\pi G_0 a^2 T_{\mu \nu} + a^2 U_{\mu \nu}$$ (4.28)

where the tensor $U_{\mu \nu}$ contains all the non-standard terms arising from a theory of modified gravity. At the perturbed level the components of $\delta U_{\mu \nu}$ are then given by (in Fourier space):

$$U_\Delta = A_0 k^2 \Phi + F_0 k^2 \Gamma + \kappa a^2 \rho_E \delta E$$

$$U_\Theta = B_0 k \Phi + I_0 k \Gamma + \kappa a^2 \rho_E \theta_E$$

$$U_P = C_0 k^2 \Phi + C_1 k \dot{\Phi} + J_0 k^2 \Gamma + J_1 k \dot{\Gamma}$$

$$U_\Sigma = D_0 \Phi + \frac{D_1}{k} \dot{\Phi} + K_0 \Gamma + \frac{K_1}{k} \dot{\Gamma}$$ (4.29)

where we recall from §2.2.4:

$$U_\Delta = -a^2 \delta U_0^0$$

$$\vec{\nabla}_i U_\Theta = -a^2 \delta U_i^0$$

$$U_P = a^2 \delta U_i^i$$

$$D_{ij} U_\Sigma = a^2 (\delta U_j^i - \frac{1}{3} \delta U_k^k \delta_j^i)$$ (4.30)

Note that these components of $\delta U_{\mu \nu}$ are somewhat different from those used in the full PPF formalism, eqs.(2.18-2.21). The coefficients $A_0, \ldots, K_1$ above are functions of background quantities, dependencies which we will suppress to avoid cluttered expressions. A subscript $E$ denotes the quantities relating to the effective fluid. The expressions above contain fewer terms than the corresponding ones in [30]; the additional metric terms present in that paper are those needed to form a gauge-invariant combination with perturbations of the new scalar. Here we have kept those terms folded into the effective fluid.

The system to be solved then comprises of the six variables $\{\Phi, \Gamma, \delta_E, \theta_E, \delta_M, \theta_M\}$ and six dynamical equations: the two spatial components of the linearized Ein-
stein equations and two fluid conservation equations for each of CDM and the effective fluid. Note that because $\delta U_{\mu\nu}$ contains additional metric perturbations the conservation equations of the effective fluid will contain some non-standard terms. The full system of equations is displayed in §4.7.1.

$\Gamma$ can be eliminated from the two spatial Einstein equations to give a second-order equation in terms of $\Phi$:

$$
\ddot{\Phi} + \dot{\Phi} \left[ \frac{\dot{\alpha}}{\alpha} - \frac{\dot{W}_1}{W_1} + k \frac{W_2}{\alpha} + \frac{k Z_1}{\alpha} \right] + \frac{k}{\alpha} \Phi \left[ \ddot{W}_2 - \frac{\dot{W}_1 W_2}{W_1} + k \frac{W_2 Z_1}{\alpha} - k \frac{W_1 Z_2}{\alpha} \right] = 0
$$

(4.31)

where $\mathcal{H}_k = \mathcal{H}/k$ and

$$
\alpha = \left( D_1 - \frac{1}{\mathcal{H}_k} \right) (J_1 - 6) + K_1 \left( 9\mathcal{H}_k - C_1 + \frac{2}{\mathcal{H}_k} \right)
$$

$$
W_1 = K_1 \left( 3\mathcal{H}_k - J_0 - \frac{2}{\mathcal{H}_k} \right) + (J_1 - 6) \left( K_0 + \frac{1}{\mathcal{H}_k} \right)
$$

$$
W_2 = (J_1 - 6)(D_0 - 1) + K_1 (2 - C_0)
$$

$$
Z_1 = \left( \frac{1}{\mathcal{H}_k} - D_1 \right) \left( 3\mathcal{H}_k - J_0 - \frac{2}{\mathcal{H}_k} \right) + \left( \frac{1}{\mathcal{H}_k} + K_0 \right) \left( 9\mathcal{H}_k - C_1 + \frac{2}{\mathcal{H}_k} \right)
$$

$$
Z_2 = (2 - C_0) \left( \frac{1}{\mathcal{H}_k} - D_1 \right) + (D_0 - 1) \left( 9\mathcal{H}_k - C_1 + \frac{2}{\mathcal{H}_k} \right)
$$

We have assumed that time derivatives of the MGFs are negligible, consistent with our treatment of metric-only theories. If we set the MGFs to zero the third term vanishes and we recover the GR result $\Phi = \text{constant}$. Modifications to the Poisson equation mean that the solution for $\Phi$ is not as easily translated into a solution for $\delta \mathcal{M}$ as it is in GR. One route is to substitute the solution for $\Phi$ into either of eqns.(4.56) or (4.57) and solve for $\Gamma$, then use both of these solutions in eqn.(4.8) – see §4.7.3 for an explicit calculation.

We wish to consider eqn.(4.31) on superhorizon and subhorizon scales, as we did in §4.3. However, there is some uncertainty involved in taking this limit without knowing specifically the functional forms hiding behind the MGFs. We will assume that any scale-dependence appears relative to some preferred scale of a given theory, i.e. as a function of $(k/k_x)$.
4. THE GROWTH OF PARAMETERIZED PERTURBATIONS IN AN EINSTEIN-DE SITTER UNIVERSE

4.4.1 Subhorizon Scales

Under the assumptions stated above, and retaining only the dominant terms when expanded in powers of \((k/H)\), on subhorizon scales eqn.(4.31) reduces to:

\[
\ddot{\Phi} + \frac{(2D_1 - C_1 + 2K_0 - J_0)}{(6 - J_1 + 2K_1)} k \dot{\Phi} + \frac{(2D_0 - C_0)}{(6 - J_1 + 2K_1)} k^2 \Phi = 0 \tag{4.32}
\]

With the change of variable \(x = k\eta\) the above equation can be rewritten (using primes to denote derivatives with respect to \(x\)):

\[
\Phi'' + \frac{\beta}{\lambda} \Phi' + \frac{(2D_0 - C_0)}{\lambda} \Phi = 0 \tag{4.33}
\]

where \(\beta = (2D_1 - C_1 + 2K_0 - J_0), \quad \lambda = 6 - J_1 + 2K_1\).

Without taking specific forms of the MGFs we cannot say whether the coefficients of the second and third terms above are positive or negative, only that sufficiently large modifications to gravity have the power to flip their signs (although the magnitude of such non-GR terms is expected to be small in the domain of validity of this equation). However, one might expect \((2D_0 - C_0)/\lambda > 0\) and \(\beta/\lambda > 0\) so that changes in \(\Phi\) damp out (and therefore return to a GR-like situation) rather than grow.

Eqn.(4.33) has the form of a simple mechanical oscillator, and will display the usual phenomena of ringing or over-damping depending on the values of the coefficients. Specifically, its behaviour will depend on the value of \(\beta^2/\lambda - 8D_0 + 4C_0\), with negative values of this quantity leading to damped oscillations in \(\Phi(x)\) and positive values leading to exponentially growing and decaying solutions. Let us apply the single boundary condition that the potential is constant on superhorizon scales; we will see shortly that this is likely to remain true in theories with extra d.o.f. We can then determine the subhorizon solution up to an overall constant:

\[
\Phi(x) \propto e^{m_+ x} - \frac{m_+}{m_-} e^{(m_+ - m_-) x} e^{m_- x} \tag{4.34}
\]

where \(m_{\pm} = -\frac{\beta}{2\lambda} \pm \sqrt{\left(\frac{\beta}{\lambda}\right)^2 - \frac{4(2D_0 - C_0)}{\lambda}} \tag{4.35}\)

At first sight this exponential solution might seem to be a cause for concern, as we would normally expect the growth of \(\Phi\) or \(\delta_M\) on small scales to follow power-law behaviour. The unusual solution above has arisen because the modifications in eqns.(4.29) introduce factors of \(k\) that dominate over the usual GR terms.
on small scales. In the toy model considered here we took the MGFs to be dimensionless functions of order one; without such assumptions it is difficult to make any general statements about the growth of perturbations, because we do not know how to assess the relative importance of two factors such as $J_0$ and $3k/\mathcal{H}$. Of course, if we know the specific functional forms hidden behind the MGFs then such assumptions are not necessary, but then we would not be pursuing a model-independent approach.

A more realistic situation would be to take the MGFs to be much smaller in magnitude than the GR terms. However, our intention here is to assess qualitatively the effects that parameterized systems of modifications to gravity have on the growth of structure - a task which becomes difficult when the parameters are taken to be vanishingly small. We will therefore maintain the simple assumptions described above, remembering that in a more realistic model the effects described here would only be manifest as small distortions of a predominantly $\Lambda$CDM+GR-controlled universe.

Alternatively, we could turn this problem on its head. One reason theories with additional d.o.f. are difficult to work with is because we lose the ability to derive a full hierarchy of constraint equations between the MGFs, as we did in the purely metric case (these arise from the Bianchi identities – see §4.7.2. If we specialize to either a ‘Type 1’ or a ‘Type 2’ theory as defined in §2.2.5 then we can recover a partial set of constraints; but only half the number obtainable for metric-only theories). Can we use the growth of $\Phi(x)$ to infer some replacement constraints?

To clarify, we wish to restore power-law behaviour for $\Phi(x)$. By expanding eqn.(4.31) in powers of $k/\mathcal{H}$ we can find the conditions necessary to remove the dominant terms that are causing the exponential solution of eqn.(4.35). We find these to be:

$$2D_0 - C_0 = 0 \quad 2D_1 - C_1 = 0 \quad 2K_0 - J_0 = 0 \quad (4.36)$$

If the above conditions are satisfied on small scales then eqn.(4.31) reduces to:

$$\ddot{\Phi} + \mathcal{H}\dot{\Phi} \left[ 1 - D_0 + \frac{12}{\lambda} \right] + \frac{\mathcal{H}^2}{2} \Phi \left( 1 - D_0 \right) \left[ \frac{6}{\lambda} - 1 \right] = 0 \quad (4.37)$$

The solutions are then power laws in $a$ (or $\eta$), as desired:

$$\Phi(a) = Pa^\nu + Qa^{\mu}$$
4. THE GROWTH OF PARAMETERIZED PERTURBATIONS IN AN EINSTEIN-DE SITTER UNIVERSE

\[ q^\pm = D_0 - \frac{12}{\lambda} - \frac{1}{2} \pm \left( D_0^2 - 3D_0 - \frac{12D_0}{\lambda} + \frac{144}{\lambda^2} + \frac{9}{4} \right)^{\frac{1}{2}} \] (4.38)

In §4.7.3 we convert this solution for \( \Phi(a) \) into a solution for \( \delta_M \) and \( \delta_E \). Here we shall simply state the results:

\[ \delta_M(a) = Pk^2 \frac{(D_0 - 1)}{(q^+ + 2)(q^+ + 3)} a^{q^+ + 1} \] (4.39)

\[ \delta_E(a) = \frac{PK^2}{6\Omega_E} \left[ 6(1 - D_0)(1 - \Omega_E) - \frac{1}{\mu_P} \right] a^{q^+ + 1} \] (4.40)

In the GR limit \( \Omega_E \to 0, D_0 \to 0 \) (which sends \( q^+ \to 0 \)) we recover that \( \delta_M \) scales with \( a \).

It is interesting to see that the conditions in eqn.(4.36) are satisfied in parameterization A of §4.2 in the limit \( k \to \infty \) (the last condition trivially so). The same is true in parameterization B, where \( D_1 = 0 \). If one wished to implement this gravitational framework into a realistic cosmological model, without necessarily taking the absolute magnitude of the MGFs to be very small, then these are the constraints that must be satisfied to give a reasonable degree of structure formation. Of course, this does not prevent the model from being ruled out by other observables such as the ISW effect. Having restored power-law growth, the effects of this theory on the ISW effect and growth function are expected to be qualitatively similar to those in §4.3.4.

4.4.2 Superhorizon Scales

On superhorizon scales eqn.(4.31) simplifies to:

\[ \ddot{\Phi} + k\dot{\Phi} \left( \frac{K_0 - D_1}{3} \right) - \frac{k^2\Phi}{18K_1} \left[ (J_1 - 12)(D_0 - 1) + K_1(2 - C_0) \right] = 0 \] (4.41)

This can be reduced to an oscillator equation in \( k\eta \), in complete analogy to the subhorizon case. However, in the superhorizon limit \( k \to \infty \) these oscillations become infinitely slow in \( k\eta \); effectively, we have that \( \Phi \) is a constant. This matches the GR and metric-only cases.

4.5 Conserved Superhorizon Perturbations?

Any relativistic theory of gravity in which energy-momentum is covariantly conserved allows the definition of a perturbation, \( Z \), that is conserved on superhorizon
scales in the absence of non-adiabatic perturbations [44, 72, 303]. In the literature \( \mathcal{Z} \) is more commonly denoted as \( \zeta \), but the choice of notation for one of the MGFs in this chapter and previous work prevents us from reusing that letter. \( \mathcal{Z} \) is identified with the curvature perturbation on uniform-expansion hypersurfaces in a homogeneous and isotropic spacetime, which in GR coincide with hypersurfaces of constant energy density. For zeroth-order Einstein equations of the form:

\[
\mathcal{H}^2 = \frac{a^2}{3} f_0 \\
\dot{\mathcal{H}} - \mathcal{H}^2 = -\frac{a^2}{2} g_0
\]  

the conserved perturbation is [72]:

\[
\mathcal{Z} = -\Phi - \frac{\mathcal{H}}{f_0} \delta f
\]  

It would be interesting to know how \( \mathcal{Z} \) behaves in parameterized gravity theories. Since the new non-GR terms do not really originate from perturbations to a fluid, it is not immediately obvious whether they will be equivalent to adiabatic or non-adiabatic pressure perturbations. Therefore the conservation of \( \mathcal{Z} \) does not necessarily follow.

In this section we are able to lift the restriction of working in a strictly EdS universe. Instead we allow the cosmological expansion rate to be modified, by using the extended Friedmann and Raychaudhuri equations (eqs.(4.42), (4.47) and (4.48)).

Using the linearly perturbed versions of eqns.(4.42), one can derive an equation for the evolution of the metric potentials [72]:

\[
\ddot{\Phi} + \frac{3 \mathcal{H} \dot{\mathcal{H}} - \dot{\mathcal{H}} - \mathcal{H} \dot{\mathcal{H}}}{\mathcal{H} - \mathcal{H}^2} \dot{\Phi} + \mathcal{H} \dot{\Psi} + \frac{2 \dot{\mathcal{H}}^2 - \mathcal{H}^2 \dot{\mathcal{H}}}{\mathcal{H} - \mathcal{H}^2} \Psi = \frac{a^2}{2} \delta g_{\text{nad}}
\]  

where the perturbation \( \delta g \) has been decomposed into parts equivalent to adiabatic and non-adiabatic pressure perturbations:

\[
\delta g = \frac{g_0}{f_0} \delta f + \delta g_{\text{nad}}
\]  

The time derivative of \( \mathcal{Z} \) is related to the quantity on the LHS of eqn.(4.44). Hence the rate of change of \( \mathcal{Z} \) is found to be:

\[
\dot{\mathcal{Z}} = \frac{\mathcal{H}}{\mathcal{H} - \mathcal{H}^2} \frac{a^2}{2} \delta g_{\text{nad}}
\]  

135
Comparing eqns. (4.28) and (4.42) and defining $U_{00} = X$, $U_{ii} = Y$ we can read off:

$$f_0 = \kappa \rho_M + X$$  \hspace{1cm} (4.47)

$$g_0 = \kappa (\rho_M + P_M) + X + Y$$  \hspace{1cm} (4.48)

Rearranging eq.(4.45) and substituting in our expressions for $f_0$ and $g_0$, one finds:

$$\delta g_{nad} = \frac{\kappa \dot{\rho}_M}{\kappa \dot{\rho}_M + \dot{X}} \left( c_s^2(M) - c_s^2(X) \right) \Gamma_{\rho MX} + \kappa \delta P_{nad} + \delta Y_{nad}$$  \hspace{1cm} (4.49)

where

$$\Gamma_{\rho MX} = \frac{\delta \rho_M}{\dot{\rho}_M} - \frac{\delta X}{\dot{X}}$$

$$c_s^2(M) = \frac{\dot{P}_M}{\dot{\rho}_M}$$

$$c_s^2(X) = \frac{\dot{X}}{\dot{X}}$$

The perturbations $\delta P_M$ and $\delta Y$ have been decomposed in a manner analogous to eq.(4.45). $\Gamma_{\rho MX}$ represents a possible entropy perturbation between CDM and the modified sector, which can be non-zero even if each component does not support entropy perturbations by itself. Perturbations of the background quantities $X$ and $Y$ correspond to components of the tensor $U_{\mu\nu}$: $a^2 \delta X = -a^2 \delta U_0^0$ and $a^2 \delta Y = a^2 \delta U_i^i/3$. Then the effective non-adiabatic pressure perturbation of the modified sector is:

$$\delta Y_{nad} = \delta Y - c_s^2(X) \delta X = \frac{\delta U_i^i}{3} + c_s^2(X) \delta U_0^0$$  \hspace{1cm} (4.50)

In this chapter we are neglecting irrelevant (at late times) contributions from radiation and neutrinos, so $\delta P_M = \delta P_{nad} = c_s^2(M) = 0$.

### 4.5.1 Metric-Only Theories

For a purely metric theory with a $\Lambda$CDM-like background (recall we have lifted the EdS restriction of previous sections) $X$ and $Y$ are zero or equivalent to a cosmological constant (i.e. $X + Y = 0$), so there can be no entropy perturbations between CDM and the modifications (since $c_s^2(X) = 0$). The only possible source for non-conservation of $Z$ would be from non-adiabatic perturbations within the
modified sector itself, $\delta Y_{\text{nad}}$. Eqn. (4.46) becomes:

\[
\dot{Z} = \frac{H}{2(\dot{H} - \dot{H}^2)} \frac{a^2 \delta U^i_i}{3} = \frac{2H}{6(\dot{H} - \dot{H}^2)} \left( k^2 C_0 \Phi + k C_1 \dot{\Phi} \right)
\]  

(4.51)

Non-adiabatic perturbations within the modified sector would amount to fluctuations about $\omega_E = -1$, which could lead to a situation equivalent to a phantom field. Whilst a phantom $\omega_E$ is permitted by current data [192, 224, 242], explicit phantom scalar field models are plagued by severe difficulties because they lead to an unstable vacuum state [82] and favour an anisotropic universe [132]. However, it is known that theories such as scalar-tensor gravity, $f(R)$ gravity and some Lorentz-violating models can cause phases of an effective $\omega_E < -1$ without introducing a phantom field per se [5, 11, 195, 203, 218].

Fortunately the question of whether non-adiabatic perturbations lead to $\omega_E < -1$ turns out to be a moot point for the theories considered in §4.2 and §4.3. This is because the effective pressure perturbation in eqn. (4.50) vanishes on very large scales. The functions $C_0$ and $C_1$ are related to $\mu_P$ and $\zeta$ by a set of constraint equations, displayed in §4.7.2. There we also show that under the assumptions made in this chapter $\lim_{k \to 0} (k^2 C_0) = 0$ and $\lim_{k \to 0} (k C_1) = 0$ in both of the parameterizations described in §4.2. These results agree with the conclusions of [251].

4.5.2 Theories with Additional Degrees of Freedom

For theories with extra scalar degrees of freedom the situation is different. Firstly, the final term in $\delta Y_{\text{nad}}$ does not vanish, as the modifications to the background equations lead to $c_s^2(X) \neq 0$. Secondly, as has been mentioned, we cannot derive a full system of constraint equations on the functions $A_0 \ldots K_1$ like those in Table 4.1, because the new scalar now acts as a source in the Bianchi identities (see §4.7.2 for details). Thirdly, we can now have entropy perturbations between matter and the modified sector, as the coefficient of $\Gamma^{\rho M X}$ in eqn. (4.49) no longer vanishes.

Maintaining the approach of §4.4, we keep the metric components of $\delta U_{\mu\nu}$ distinct, but treat the perturbations of the new d.o.f. as an effective fluid. We previously set the equation of state parameter of this effective fluid to zero in order to preserve the EdS nature of our toy model; let us restore the general case
for the present. The full expression for $\dot{Z}$ then becomes:

$$
\dot{Z} = \frac{\mathcal{H}}{2(\mathcal{H} - 3\dot{\mathcal{H}})} \left[ \frac{\kappa \rho_M \dot{X}}{\kappa \rho_M + X} c_s^{2(X)} \left( \frac{a^2 \delta_M}{3 \mathcal{H}} + \frac{\kappa a^2 \rho_E \delta_E}{X} \right) + \kappa a^2 \rho_E \left( \Pi_E - c_s^{2(X)} \delta_E \right) \\
+ k^2 \Phi \left( \frac{1}{3} C_0 - A_0 c_s^{2(X)} \frac{\dot{X}}{\kappa \rho_M + X} \right) + k^2 \Gamma \left( \frac{1}{3} J_0 - F_0 c_s^{2(X)} \frac{\dot{X}}{\kappa \rho_M + X} \right) \\
+ \frac{1}{3} k \left( C_1 \dot{\Phi} + J_1 \dot{\Gamma} \right) \right]
$$

(4.52)

If the MGFs do not contain inverse powers of $k$ then the second and third lines above vanish on very large scales, leaving only perturbations to the effective fluid (although a similar assumption was used in §4.4 to make an analytic solution achievable, it does not necessarily have to hold true for all theories). The first set of round brackets (on the first line of eq.(4.52)) represents entropy perturbations between CDM and the modified sector, and the second term is akin to non-adiabatic perturbations within the modified sector itself. However, note that $c_s^{2(X)}$ is not necessarily equal to the sound speed of the effective fluid, since $X$ and $Y$ may contain terms constructed from the metric. This feature distinguishes a dark energy model like quintessence from a modified gravity model such as a scalar-tensor gravity, as per our definition in §1.3.2. In the latter the scalar field is non-trivially coupled to the metric and hence $X$ and $Y$ contain more terms than just the energy density and pressure of a scalar field. A quintessence-like case can be recovered from the above expression by setting the MGFs to zero and $X = \kappa \rho_E$, $Y = \kappa P_E$, which gives:

$$
\dot{Z} = -\frac{\mathcal{H}}{2(\mathcal{H} - 3\dot{\mathcal{H}})} \frac{a^2 \rho_E (1 + \omega_E) c_s^{2(E)} \Gamma_{M,E}}{1 + (1 + \omega_E) \Pi_E / \Pi_M} + \kappa \delta P_{E,\text{nad}}
$$

(4.53)

where

$$
\Gamma_{M,E} = \delta_M - \frac{\delta_E}{1 + \omega_E}
$$

(4.54)

If the quintessence field supports only adiabatic perturbations ($\delta P_{E,\text{nad}} = 0$) then $\Gamma_{M,E}$ can be set to zero through choice of adiabatic initial conditions, leaving $Z$ conserved.

As a second example, consider the theory treated in §4.4 in which the effective
fluid was assumed to be pressureless. Eqn.(4.52) reduces to:

$$\dot{Z} = -\frac{\mathcal{H}}{(\dot{\mathcal{H}} - \dot{\mathcal{H}}^2)} \frac{\kappa a^2}{2} \frac{\dot{X}}{\kappa \dot{\rho}_M + \dot{X}} c_s^2(\chi) \left[ \rho_M \delta_M + \rho_E \delta_E \right]$$  \hspace{1cm} (4.55)

The square brackets can be set to zero at through a choice of isocurvature initial conditions, so \( \dot{Z} \) remains conserved if \( \omega_E \) is constant. However, as one wishes the effects of modified gravity to become apparent at late times in the universe an evolving \( \omega_E \) would be more desirable, for which \( \dot{Z} \) would not be conserved.

### 4.6 Conclusions

Whilst there has been much effort made to constrain simple parameterizations of modified gravity with data [40, 90, 322], there has been relatively little investigation into the corresponding theoretical description: how does the parameterized system of perturbation equations evolve? (Although see [128]; general scalar field-type models are treated in [57, 95, 299]).

We have attempted to answer this question within the simplified setting of an Einstein-de Sitter universe, for two classes of theories: 1) those for which the degrees of freedom are the metric and matter perturbations, and 2) theories which explicitly introduce additional scalar degrees of freedom. In both cases we have found that the curvature perturbation remains constant on the very largest scales (well above our observable horizon).

On subhorizon scales we found that in the metric-only case perturbations grow in familiar power-law fashion, but the exponents are modified from their General Relativistic values. The two modified gravity functions \( \mu_P \) (controlling the effective Newton’s constant) and \( \zeta \) (controlling the dominant component of the gravitational slip) are partially degenerate in their effects, making them difficult to disentangle. The impact of \( \zeta \) becomes subdominant if the effective Newton’s constant is weakened. Modifications to the evolution of \( \Phi \) lead to an induced ISW effect and modification to the growth rate, \( f(z) \).

Theories with additional scalar degrees of freedom are considerably harder to study, as the number of undetermined functions is larger. In §4.4 we considered phases during which the new terms in the Einstein equations behave as an effective fluid with a negligible equation of state. We found that the modified terms dominate the evolution equations, leading to damped oscillatory or exponential behaviour. From a study of these equations we have found three relations (eqns.(4.36)) between the modified terms that, if satisfied, will restore
power-law behaviour. This analysis does not apply to all modified gravity theories (when $\omega_E$ is not small), but it is relevant to (for example) EdS regimes of linear Einstein-Aether theory, EBI gravity and Hořava-Lifshitz gravity.

Constraint equations such as this are desirable because the number of free functions required to parameterize common gravitational theories increases rapidly when new degrees of freedom are included. This proliferation could reduce our ability to constrain such parameterized frameworks satisfactorily. One can reduce this freedom somewhat by treating the new scalar degrees of freedom as an effective fluid; indeed this (or an equivalent approach) is necessary if the parameterization is to capture theories which introduce more than two new scalar degrees of freedom. The metric perturbations are also a considerable source of freedom, as displayed in eqns. (4.29). This splitting of the modifications into metric parts and effective fluid parts is the key to distinguishing between closely related models of dark energy and modified gravity, such as quintessence and scalar-tensor theories.

Under the assumptions made in this chapter the metric terms become irrelevant on ultra-large scales. In a metric-only theory this leaves the superhorizon perturbation $\mathcal{Z}$ (more commonly denoted by by $\zeta$) conserved, but in class 2) theories the possible non-conservation of $\mathcal{Z}$ depends on the equation of state of the effective fluid.

The EdS model considered for most of this chapter is useful in obtaining a qualitative understanding of how parameterized gravity might effect the matter-dominated phase of our universe. To obtain more quantitative predictions this must be embedded in a more complex cosmological model.

4.7 Supporting Material for Chapter 4

4.7.1 System of Equations for Theories with Extra D.o.F.

This section displays the system of six equations that is solved in §4.4. These are: the two spatial components of the Einstein equations (longitudinal and transverse traceless), the fluid conservation equations for cold dark matter (the two components of $\delta (\nabla_\mu T^\mu_\nu) = 0$) and the two Bianchi identities $\delta (\nabla_\mu U^\mu_\nu) = 0$. The non-standard terms containing MGFs arise from the parameterization laid out in eqns. (4.28) and (4.29), which treats the additional scalar degrees of freedom as a pressureless fluid (denoted by a subscript $E$). This reduces to parameterization A used in §4.3 for purely metric theories.

The variable $\Psi$ in these equations can be eliminated in favour of $\Gamma$ using the
definition $\Gamma = (\dot{\Phi} + \mathcal{H} \Psi)/k$.

The two spatial components of the linearized Einstein equations are:

$$6k \dot{\Gamma} + 12\mathcal{H}k \Gamma + 2k^2 (\Phi - \Psi) + 6(\dot{\mathcal{H}} - \mathcal{H}^2) \Psi = C_0 k^2 \Phi + C_1 k \dot{\Phi} + J_0 k^2 \Gamma + J_1 k \dot{\Gamma}$$

(4.56)

$$\Phi - \Psi = D_0 \Phi + \frac{D_1}{k} \dot{\Phi} + K_0 \Gamma + \frac{K_1}{k} \dot{\Gamma}$$

(4.57)

The conservation equations for matter:

$$\delta_M = -k^2 \theta_M + 3\dot{\Phi}$$

(4.58)

$$\dot{\theta}_M = -3 \mathcal{H} \theta_M + \Psi$$

(4.59)

The conservation equations for the new effective fluid:

$$\kappa a^2 \rho_E \delta_E = \kappa a^2 \rho_E \left( -k^2 \theta_E + 3\dot{\Phi} \right) - k^2 \Phi \left( \mathcal{H} A_0 + k B_0 + \mathcal{H} C_0 \right) + k \dot{\Phi} \left( k A_0 + \mathcal{H} C_1 \right)$$

$$+ k^2 \Gamma \left( \mathcal{H} F_0 + k I_0 + \mathcal{H} J_0 \right) + k \dot{\Gamma} \left( k F_0 + \mathcal{H} J_1 \right)$$

(4.60)

$$\kappa a^2 \rho_E \dot{\theta}_E = \kappa a^2 \rho_E \left( -\mathcal{H} \theta_E + \Psi \right) - k^2 \Phi \left( \frac{2}{k} \mathcal{H} B_0 - \frac{1}{3} C_0 + \frac{2}{3} D_0 \right) - k \dot{\Phi} \left( B_0 - \frac{1}{3} C_1 + \frac{2}{3} D_1 \right)$$

$$- k^2 \Gamma \left( \frac{2}{k} I_0 - \frac{1}{3} J_0 + \frac{2}{3} K_0 \right) - k \dot{\Gamma} \left( I_0 - \frac{1}{3} J_1 + \frac{2}{3} K_1 \right)$$

(4.61)

### 4.7.2 Conservation of $Z$ in Metric-Only Theories

In this section we demonstrate that any effective non-adiabatic perturbations that might prevent $Z$ from being conserved vanish on large scales for the metric-only theories considered in §4.2 and §4.3. First we must introduce some more detail about the parameterization underlying eqns.(4.4) and (4.5).

Consider a modified Einstein equation of the form (4.28). In [30] we demonstrated that in parameterization A (in which $\mu_s = \mu_P$) a theory containing up to second-order time derivatives corresponds to the following forms for the components of $\delta U_{\mu\nu}$ (see eqns.(4.30) for definitions of quantities on the left-hand side):

$$U_\Delta = A_0 k^2 \dot{\Phi}$$

$$U_\Theta = B_0 k \dot{\Phi}$$

$$U_P = C_0 k^2 \dot{\Phi} + C_1 k \ddot{\Phi}$$

$$U_\Sigma = D_0 \dot{\Phi} + \frac{D_1}{k} \ddot{\Phi}$$

(4.62)

where the coefficients $A_0, \ldots D_1$ are functions of background quantities, depen-
4. THE GROWTH OF PARAMETERIZED PERTURBATIONS IN AN EINSTEIN-DE SITTER UNIVERSE

<table>
<thead>
<tr>
<th>Constraint equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ( \dot{A}_0 + 3\mathcal{H}A_0 + kB_0 + \mathcal{H}C_0 = 0 )</td>
</tr>
<tr>
<td>2 ( A_0 + \mathcal{H}_k C_1 = 0 )</td>
</tr>
<tr>
<td>3 ( \dot{B}_0 + 2\mathcal{H}B_0 - \frac{1}{3}kC_0 + \frac{2}{3}kD_0 = 0 )</td>
</tr>
<tr>
<td>4 ( B_0 - \frac{1}{3}C_1 + \frac{2}{3}D_1 = 0 )</td>
</tr>
</tbody>
</table>

Table 4.1: Constraint equations for a metric theory in parameterization A, specified by eqns.(4.62).

dencies which we will suppress to avoid cluttered expressions.

The two components of \( \delta (\nabla_\mu U^\mu_\nu) = 0 \) result in equations containing \( \Phi \) and its derivatives. To avoid contradicting the solution for \( \Phi \) dictated by the Einstein equations these expressions must vanish identically. Setting the coefficients of each metric perturbation to zero leads to the four constraint equations listed in Table 4.1. This enables us to reduce the six free functions in eqns.(4.62) to just two. Note that terms in \( \hat{\Psi} \) are not permitted to appear in \( \delta U^0_0 \) and \( \delta U^0_i \), because they contain second-order time derivatives, which would lead to third-order Bianchi identities. Constraint equations similar to those in Table 4.1 prevent \( \hat{\Psi} \) from featuring in \( \delta U^i_i \) and \( \delta U^i_j \) if it is not present in \( \delta U^0_0 \) and \( \delta U^0_i \).

We will choose one of our free functions to be \( D_0 \) – this corresponds to \( \zeta \) in eqn.(4.5). The second free function will be the combination that appears when we form the Poisson equation:

\[
\mu_P = \frac{1}{1 + \frac{1}{2}(A_0 + 3\mathcal{H}_k B_0)}
\]  

(4.63)

where \( \mathcal{H}_k = \mathcal{H}/k \). In terms of these two MGFs the other coefficient functions are:

\[
A_0 = -2 \left( \frac{\mu_P - 1}{\mu_P} \right) \left( 1 + \frac{3\mathcal{H}^2}{Q} \right) + 2\zeta \frac{\mathcal{H}^2}{Q}
\]  

(4.64)

\[
B_0 = \frac{2\mathcal{H}k \left( \frac{\mu_P - 1}{\mu_P} \right) - \zeta}{3Q}
\]  

(4.65)

\[
C_0 = \frac{2}{Q} \left( \frac{\mu_P - 1}{\mu_P} - \zeta \right) \left( \dot{\mathcal{H}} - \mathcal{H} \frac{\dot{Q}}{Q} + 2\mathcal{H}^2 \right) + 2\zeta
\]

\( = \frac{4}{3} \left( \frac{\mu_P - 1}{\mu_P} - \zeta \right) \left( \frac{\mathcal{H}^2}{Q} + 2\zeta \right) \]  

(4.66)

\[
\]
Table 4.2: Table of the constraint equations for a metric theory in parameterization B, specified by eqns.(4.69).

<table>
<thead>
<tr>
<th>Constraint equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 [ \dot{A}_0 + \mathcal{H}A_0 + kB_0 + \mathcal{H}C_0 = 0 ]</td>
</tr>
<tr>
<td>2 [ \dot{A}_1 + \mathcal{H}A_1 + kB_0 + \mathcal{H}C_1 = 0 ]</td>
</tr>
<tr>
<td>3 [ kA_1 + \mathcal{H}C_2 = 0 ]</td>
</tr>
<tr>
<td>4 [ \dot{B}_0 + 2\mathcal{H}B_0 - \frac{1}{3}kC_0 + \frac{2}{3}D_0 = 0 ]</td>
</tr>
<tr>
<td>5 [ \dot{B}_1 + 2\mathcal{H}B_1 + kB_0 - \frac{1}{3}kC_1 = 0 ]</td>
</tr>
<tr>
<td>6 [ B_1 - \frac{1}{3}C_2 = 0 ]</td>
</tr>
</tbody>
</table>

Table 4.2: Table of the constraint equations for a metric theory in parameterization B, specified by eqns.(4.69).

\[ C_1 = \frac{2}{\mathcal{H}_k} \left( \frac{\mu_P - 1}{\mu_P} \right) \left( 1 + \frac{\mathcal{H}^2}{Q} \right) - 2\mathcal{H}k \zeta \]  

\[ D_1 = \frac{1}{\mathcal{H}_k} \left( \frac{\mu_P - 1}{\mu_P} \right) \]  

where \( Q = \mathcal{H}^2 + \frac{k^2}{3} - \mathcal{H} \)

The second expression for \( C_0 \) applies only in an EdS universe. From eqn.(4.51) we see that the quantity of interest for evaluating \( \dot{\zeta} \) is \( \delta U^i_i \). In fact we only need to know \( C_0 \), since we showed in §4.3.2 that the potential is constant on large scales. Under the assumption that \( \mu_P \) and \( \zeta \) do not contain inverse powers of \( k \), eqns.(4.62) and (4.68) imply that \( \lim_{k \to 0} \delta U^i_i = 0 \). Hence \( \zeta \) is conserved on superhorizon scales in parameterization A.

We can repeat this calculation for parameterization B, in which \( \mu_s = 1 \) but \( \mu_P \) remains a free function. This corresponds to a \( \delta U_{\mu\nu} \) tensor of the following form (see [30]):

\[ U_\Delta = A_0 k^2 \dot{\Phi} + A_1 k \ddot{\Phi} \]
\[ U_\phi = B_0 k \dot{\Phi} + B_1 \ddot{\Phi} \]
\[ U_P = C_0 k^2 \dot{\Phi} + C_1 k \ddot{\Phi} + C_2 \dddot{\Phi} \]
\[ U_\Sigma = D_0 \dot{\Phi} \]  

(4.69)

The constraint equations for this theory are listed in Table 4.2. In terms of the MGFs \( \mu_P \) and \( \zeta \) the coefficients are:

\[ A_0 = -2 \left( \frac{\mu_P - 1}{\mu_P} \right) \left( 1 + \frac{\mathcal{H}^2}{Q} \right) + 2\mathcal{H}^2 \zeta \]
4. THE GROWTH OF PARAMETERIZED PERTURBATIONS IN AN EINSTEIN-DE SITTER UNIVERSE

\[ A_i = -\frac{6\mathcal{H}_k}{1 + 3\mathcal{H}_k^2} \left( \frac{\mu_P - 1}{\mu_P} \right) \]

\[ B_0 = \frac{2\mathcal{H}k \left( \frac{\mu_P - 1}{\mu_P} - \zeta \right)}{3Q} \]

\[ B_1 = \frac{2}{(1 + 3\mathcal{H}_k^2)} \left( \frac{\mu_P - 1}{\mu_P} \right) \]

\[ C_0 = \frac{2}{Q} \left( \frac{\mu_P - 1}{\mu_P} - \zeta \right) \left( \mathcal{H} - \mathcal{H} \frac{\dot{Q}}{Q} + 2\mathcal{H}^2 \right) + 2\zeta \]

\[ = \frac{2}{Q} \left( \frac{\mu_P - 1}{\mu_P} - \zeta \right) \left( \frac{3}{2} \mathcal{H}^2 - \mathcal{H} \frac{\dot{Q}}{Q} \right) + 2\zeta \]

\[ C_1 = \frac{2\mathcal{H}k}{Q} \left( \frac{\mu_P - 1}{\mu_P} - \zeta \right) + 4\mathcal{H}k \left( \frac{\mu_P - 1}{\mu_P} \right) \frac{\left( Q + \frac{2}{3} \mathcal{H} \right)}{\left( Q + \mathcal{H} \right)^2} \]

\[ = \frac{2\mathcal{H}k}{Q} \left( \frac{\mu_P - 1}{\mu_P} - \zeta \right) + 4\mathcal{H}k \left( \frac{\mu_P - 1}{\mu_P} \right) \frac{\left( Q - \frac{1}{3} \mathcal{H}^2 \right)}{\left( Q - \frac{1}{2} \mathcal{H}^2 \right)^2} \]

\[ C_2 = \frac{6}{(1 + 3\mathcal{H}_k^2)} \left( \frac{\mu_P - 1}{\mu_P} \right) \]

As we found for the parameterization A case, \( \lim_{k \to 0} \delta U_i = 0 \). Hence \( Z \) is conserved in parameterization B also.

4.7.3 \( \delta_M \) and \( \delta_E \) in Theories with Extra D.o.F.

We wish to connect the subhorizon solution for \( \Phi \) in eqn.(4.38) to the density perturbations of CDM and the effective fluid. First we cast eqns.(4.56) and (4.57) in dimensionless format by using the substitution \( x = k\eta \) (and correspondingly \( \mathcal{H} = 2/x, \dot{\mathcal{H}} = -2/x^2 \)). We also apply the conditions necessary for power-law growth, given in eqns.(4.36). The spatial components of the Einstein equations become (where primes denote derivatives with respect to \( x \)):

\[ \Gamma' \left( 6 - J_1 \right) + \Gamma \left( \frac{6}{x} - x - 2K_0 \right) + \Phi' \left( \frac{18}{x} + x - 2D_1 \right) + 2\Phi(1 - D_0) = 0 \]

\[ \Gamma' K_1 + \Gamma \left( K_0 + \frac{x}{2} \right) + \Phi' \left( D_1 - \frac{x}{2} \right) - \Phi(1 - D_0) = 0 \]

Eliminating \( \Gamma' \) from these:

\[ \Phi' \left[ \frac{18}{x} + \frac{\lambda}{K_1} \left( \frac{x}{2} - 1 \right) \right] + \Phi \frac{\lambda}{K_1} (1 - D_0) + \Gamma \left[ \frac{6}{x} - \frac{\lambda}{K_1} \left( \frac{x}{2} + 1 \right) \right] = 0 \]
where $\lambda = 6 - J_1 + 2K_1$. Keeping only the growing mode, we write the solution for $\Phi$ as:

$$\Phi(x) = R(k)x^q \quad (4.72)$$

We have absorbed the $k$-dependence into the prefactor, $R(k) = P/k^q$ where $P$ is a constant (see eqn.(4.38)), and dropped the superscript $+$ on $q$ to avoid cluttered expressions.

Taking the subhorizon limit $x \rightarrow \infty$, eqn.(4.71) gives us the following solution for $\Gamma$:

$$\Gamma(x) = 2R(k) \left( \frac{q}{2} + 1 - D_0 \right) x^{q-1} \quad (4.73)$$

Next we recast the fluid conservation equations for CDM in dimensionless format and combine them in a manner analogous to eqns.(4.6)-(4.9), yielding:

$$\delta''_M + \frac{2}{x}\delta'_M = 3\Phi'' + \Phi' \left( \frac{6}{x} + \frac{x}{2} \right) - \frac{x}{2} \Gamma \quad (4.74)$$

This has the solution:

$$\delta_M(x) = c_1 + \frac{c_2}{x} + R(k)x^q \left[ \frac{(D_0 - 1)x^2}{(q + 2)(q + 3)} + 3 \right] \quad (4.75)$$

where $c_1$ and $c_2$ are integration constants. This solution has the desired behaviour that when we set $D_0 = 0$ (which sets $q = 0$, see eqn.(4.38)) we recover that $\delta_M$ grows as $\delta_M \propto x^2 \propto a$, as occurs in GR-controlled matter-dominated epoch.

Finally we use the (dimensionless) Poisson equation to relate the solutions for $\Phi(x)$ and $\delta_M(x)$ to $\delta_E(x)$:

$$-2\Phi = \frac{12\mu_P}{x^2} (\Omega_M\delta_M + \Omega_E\delta_E) + \Gamma \mu_P \left( F_0 + \frac{6}{x}J_0 \right) \quad (4.76)$$

In our toy EdS universe, containing only CDM and the effective fluid, $\Omega_M$ and $\Omega_E$ are constants. Retaining just the dominant growing modes for both $\delta_M(x)$ and $\delta_E(x)$, the above equation gives us the result:

$$\delta_E(x) = \frac{R(k)}{6\Omega_E} \left[ 6(1 - D_0)(1 - \Omega_E) \frac{1}{(q + 2)(q + 3)} \right] x^{q+2} \quad (4.77)$$

As expected, $\delta_M(x)$ and $\delta_E(x)$ grow at the same rate. Note that, as in GR, their evolution differs from that of $\Phi(x)$ by a single factor of $a$. 

145
4. THE GROWTH OF PARAMETERIZED PERTURBATIONS IN AN EINSTEIN-DE SITTER UNIVERSE
Chapter 5

Corrections to the Growth Rate in Modified Gravity

The growth rate of the large-scale structure of the universe, \( f(z) \), has been advocated as the observable par excellence for testing gravity on cosmological scales. In this chapter I will discuss the dependence of \( f(z) \) and its density-weighted version, \( f(z)\sigma_8(z) \), on deviations from General Relativity. By considering the PPF parameterization in the quasistatic regime, I show that key observables can be expressed in the form of an integral over a ‘source’ term weighted by a ‘response kernel’. I argue that this approach is a remarkably efficient and accurate method for assessing the constraining power of current and future galaxy surveys.

5.1 Introduction

The growth rate of large-scale structure is acutely sensitive to the nature of gravitational collapse. It has been argued that an accurate measure of \( f(a) \), which we recall from §1.2.2.1 is defined as:

\[
f(a) \equiv \frac{d\ln \Delta(a)}{d\ln a}
\]

(5.1)

where \( \Delta(a) \) is the amplitude of the growing mode of matter density perturbations, can be used to constrain deviations from GR.

The method of choice for measuring \( f(a) \) (or equivalently, \( f(z) \)) is through redshift space distortions [93, 146, 179, 235]. The two-point correlation function of galaxies in redshift space is both anisotropic and scale-dependent, due to two competing effects: on small scales, the virialized motions of galaxies dominate over the Hubble flow, resulting in the elongation of the contours of the correlation...
function along the line of sight - the fingers of god effect. On larger scales, gravitational infall leads to a squashing of the contours that is detectable on scales of $10 - 30 \, h^{-1} \, \text{Mpc}$, see Fig. 1.4. There has been substantial progress in modelling this effect, both analytically [273] and numerically [172, 173], and a number of systematic effects (non-linearity, the role of bias) have been studied. On still larger scales it is also possible to measure $f(z)$ with direct measurements of large-scale flows [168, 215].

The current observational status of $f(z)$ is promising and intriguing. The surveys of [46, 47, 99, 267, 297] have measured the growth rate from $z = 0.2$ to $z = 1.3$ on a range of scales, with errors of approximately $10 - 20\%$. These measurements have provided decisive evidence for ruling out some extreme theories of modified gravity [255]. We will show in this chapter that most theories give rise to more subtle signatures that still lie within current experimental error bars; however, this situation should change with the next generation of galaxy surveys (see §5.7).

The growth rate is a particularly attractive observable from a theoretical point of view. For a start, we expect to measure it predominantly on scales where linear cosmological perturbation theory is valid. There is a battery of well-seasoned techniques associated with linear perturbation theory, and as we have shown in earlier chapters of this thesis, it is possible to adapt these for use with widely-applicable parameterizations of modified gravity theories. Extending growth rate calculations to the mildly non-linear regime is possible [172, 209] but still in its infancy; furthermore, the reliance on theory-specific N-body simulations prevents one from making general statements about the effects of modified gravity on these scales.

A key advantage of $f(z)$ is that the range of scales probed is well inside the cosmological horizon, where the quasistatic approximation can be applied (see §5.2). This allows the gravitational field equations and the equations of motion for density perturbations to be simplified.

In this chapter we wish to explore the power of the growth rate as a probe of gravity. To do so, we will first introduce the quasistatic approximation mentioned above. In §5.3 we derive the evolution equation for $f(z)$. We show how, in the quasistatic regime, it depends on: a) a specific combination of gravitational parameters which we refer to as $\xi(z,k)$, and b) the background expansion history, parameterized through an equation of state, $\omega(z)$. In §5.4 we propose a simple and efficient method for linking deviations in the direct observable $f(z)\sigma_8(z)$ to $\xi(z,k)$ and $\omega(z)$. In §5.5 we discuss the degeneracy between $\xi(z,k)$ and $\omega(z)$ and how it may affect inferences made from the growth rate. Our technique can
be applied to a wide variety of gravity theories; §5.6 demonstrates this explicitly by deriving $\xi(z,k)$ and $\omega(z)$ from the Parameterized Post-Friedmann formalism of chapter 2 and calculating a suite of examples, including Horndeski theory. In §5.7 we use our method to give a broad picture of how effective the growth rate actually is in constraining deviations from General Relativity. In §5.8 we discuss our findings and how our method can be extended to other cosmological observables.

5.2 The Quasistatic Approximation

As recently highlighted by Silvestri et al. [281], there are two separate components to the quasistatic approximation; in GR one necessarily implies the other, but in alternative gravity theories this is not generally the case. The two components are:

1. The approximation that the time derivatives of scalar perturbations are negligible relative to their spatial derivatives. This includes both metric potentials and any new perturbations not present in GR (eg. $\delta\phi$ for theories involving a new scalar field $\phi$).

2. The consideration of only significantly subhorizon scales. When working in Fourier space this implies that terms containing factors of $H/k$ can be safely neglected.

Both points warrant brief further discussion. The unimportance of the time derivatives of metric potentials ($\dot{\Phi}, \dot{\Psi}, \ddot{\Phi}$ etc. in the conformal Newtonian gauge) is justified for the late-time universe because for viable cosmological models these potentials evolve on timescales comparable to a Hubble time. Implementation of the second point above then ensures that they are safely negligible.

The extension of this approximation to non-GR perturbations is less obvious. The equations of motion for $\Phi$ and $\Psi$ are, of course, the Einstein field equations for an FRW metric - hence they are naturally linked to the Hubble scale. Non-GR perturbations will be governed by additional equations of motion which could potentially involve radically different scales or source terms.

Instead the justification for neglecting these terms comes predominantly from N-body simulations for specific gravity theories, e.g. $f(R)$ gravity [173, 202], DGP [272] and Galileon theories [313]. When computed during these simulations, the time derivatives in question have typically been found to be 5 – 6 orders of magnitude smaller than the corresponding spatial derivatives [272]. These results
have recently been called into question [210] in the case of gravity theories that implement the symmetron screening mechanism [61, 152, 153], but this represents a specialized scenario.

Likewise, the implementation of the second point above becomes a little more subtle for theories introducing new preferred lengthscales. Non-GR perturbations with a sound speed less than unity can only affect scales up to their Compton wavelength. Use of the quasistatic approximation then imposes the stronger requirement $c_s k \gg \mathcal{H}$ (where $c_s$ is the effective sound speed of the perturbations), that is, the consideration of scales well below this second, smaller ‘horizon’.

A key goal of this chapter is study the effects of modified gravity in a way which is as theory-independent as possible. To this end, we will assume the absence of any new key length scale in point 2). We will also assume the N-body justification of point 1) to hold true for the majority of gravity theories, at least on the lengthscales of concern in this chapter (on which viable screening mechanisms are not operational). Doubtless this corresponds to the exclusion of some theories from our treatment, but we do not expect the restriction to be severe.

5.3 The Linear Growth Rate in Modified Gravity

We now show how modifications to the gravitational field equations will affect the evolution of the growth rate of density perturbations, as defined in eq.(5.1) [64]. Consider the pressureless matter component of the universe. Small inhomogeneities in the energy density, $\delta_M$, are defined through $\rho_M = \bar{\rho}_M(1 + \delta_M)$, where $\bar{\rho}_M$ is the mean energy density. In the conformal Newtonian gauge the evolution equations for the velocity potential $\theta$ (where the velocity perturbation is $v_i = \nabla_i \theta$) and the gauge-invariant density contrast $\Delta = \delta + 3\mathcal{H}/k(1 + \omega)\theta$ are:

$$\dot{\Delta}_M = 3(\dot{\Phi} + \Phi \dot{\Psi}) - \theta_M \left[ k^2 + 3(\mathcal{H}^2 - \dot{\mathcal{H}}) \right]$$  

(5.2)

$$\dot{\theta}_M = -\mathcal{H}\theta_M + \Psi$$  

(5.3)

Eq.(5.2) is derived by combining eq.(5.3) with eq.(1.31), using $\omega_M = 0$.

We will hereafter sometimes suppress the arguments of functions for ease of expression. We warn the reader that $\Omega_M$ should always be interpreted as a time-dependent quantity; we will use $\Omega_M^0$ for its value today.

The quasistatic regime concerns significantly sub-horizon scales, where $k \gg \mathcal{H}$
At these scales the $k^2$ term dominates eq.(5.2), so:

$$\dot{\Delta}_M \approx -k^2 \theta_M$$  \hspace{1cm} (5.4)$$

Differentiating this with respect to conformal time and using eq.(5.3) we obtain:

$$\ddot{\Delta}_M + \mathcal{H} \dot{\Delta}_M + k^2 \Psi \approx 0$$  \hspace{1cm} (5.5)$$

In the quasistatic limit the Poisson equation and the relationship between the two metric potentials can be written in the now-familiar form (see §1.2.3.1, §2.4.3, §3.2 and §4.2):

$$2 \nabla^2 \Phi = \kappa a^2 \mu(a,k) \bar{\rho}_M \Delta_M$$  \hspace{1cm} (5.6)$$

$$\frac{\Phi}{\Psi} = \gamma(a,k)$$  \hspace{1cm} (5.7)$$

where we have defined two time- and scale-dependent functions, $\mu(a,k)$ and $\gamma(a,k)$. We combine eqs.(5.6) and (5.7) to eliminate $\Phi$, and then use the resulting expression to eliminate $\Psi$ from eq.(5.5). Finally we use the Friedmann equation to express $\kappa a^2 \bar{\rho}_M = 3 \mathcal{H}^2 \Omega_M$. The result is:

$$\ddot{\Delta}_M + \mathcal{H} \dot{\Delta}_M - \frac{3}{2} \mathcal{H}^2 \Omega_M \xi \Delta_M = 0$$  \hspace{1cm} (5.8)$$

where we have defined $\xi \equiv \mu/\gamma$. The quantity $\xi(a,k)$ will appear frequently throughout this chapter; it is equal to 1 in GR. For convenience we will rewrite the equation above using $x = \ln a$ as the independent variable. The derivatives of $\Delta$ change according to:

$$\dot{\Delta} = \mathcal{H} \Delta'$$

$$\ddot{\Delta} = \mathcal{H}^2 \Delta'' + \mathcal{H} \dot{\mathcal{H}} \Delta'$$  \hspace{1cm} (5.9)$$

We use the above expressions in eq.(5.8) and divide through by $\mathcal{H}^2$ to obtain:

$$\Delta_M'' + \left(1 + \frac{\mathcal{H}'}{\mathcal{H}}\right) \Delta_M' - \frac{3}{2} \Omega_M \xi \Delta_M = 0$$  \hspace{1cm} (5.10)$$

Primes denote derivatives with respect to $x$. It is helpful to convert this second-order equation for $\Delta_M$ into a first-order equation for the growth rate. We employ the usual definition:

$$f = \frac{d \ln \Delta_M}{d \ln a} = \frac{\Delta_M'}{\Delta_M}$$  \hspace{1cm} (5.11)$$
and the consequential result $\Delta''/\Delta = f'' + f^2$. Substituting these expressions into eq.(5.10), we find that the evolution equation for $f$ is:

$$f'(x, k) + q(x) f(x) + f^2(x) = \frac{3}{2} \Omega_M(x) \xi(x, k)$$

(5.12)

where

$$q(x) = \frac{1}{2} [1 - 3\omega(x)(1 - \Omega_M(x))]$$

(5.13)

There are a couple of important points to be made about eq.(5.12). First of all, because $\xi$ can generally be a function of scale, we must allow for a possible scale-dependence of the growth rate; this is a common property of modified gravity theories which distinguish them from GR. Second, note that we have introduced a free function, $\omega(x)$, acting as an effective equation of state of the non-matter sector. The unperturbed expansion history of any dark energy or modified gravity theory can be written in the form of the usual GR Friedmann equation with a new fluid component, through a suitable choice of $\omega(x)$ [196, 284].

Whilst the growth rate is of prime importance, in practice one actually measures the density-weighted or observable growth rate, $f\sigma_8$, where $\sigma_8$ is the rms of mass fluctuations in spheres of $8 \, h^{-1}\text{Mpc}$ [238]. $\sigma_8$ evolves with the same growth factor $D$ as the matter overdensity, i.e.:

$$\frac{\sigma_R(z)}{\sigma_R(z = 0)} \simeq D(z, k = \frac{2\pi}{R}) = \frac{\Delta_M(z, k = \frac{2\pi}{R})}{\Delta_M(0, k = \frac{2\pi}{R})}$$

(5.14)

where $R = 8h^{-1}\text{Mpc}$. This will prove useful in §5.4.2.

### 5.4 The Linear Response Approach

We can assume that any viable theory of modified gravity must result in observables that match a General Relativistic model to a high degree of accuracy. We then ask the question: what small deviations from GR are still permissible within the error-bars of current and near-future experiments? We will answer this question by considering linear perturbations about the ‘background’ solution of $\Lambda$CDM+GR, $\omega(x) = -1$.

It is important to note that our approach should not be confused with standard cosmological perturbation theory of an FRW universe. We are already working within the context of spacetime linear perturbation theory. Here we are perturbing around $\Lambda$CDM+GR in solution space by assuming that the functions $\xi$ and $\omega$ source small deviations $\delta f$ from the $\Lambda$CDM+GR growth rate of $\Lambda$CDM (eq.(5.12)). We will see shortly that this an excellent approximation to the full
non-linear response.

For simplicity we will first investigate the impact of these small deviations on the (unobservable) growth rate $f$, before extending our treatment to the observable $f\sigma_8$ in §5.4.2.

### 5.4.1 The Response Function of the Growth Rate $f(z)$

We begin by decomposing $f$ into a zeroth-order part and a perturbation. As stated above, the zeroth-order solution is that of GR and hence is scale-independent, but the perturbation may not be:

$$f(x,k) = f_{GR}(x) + \delta f(x,k)$$ (5.15)

Likewise we perturb $\xi$, $\omega$ and $\Omega_M$ about their GR+$\Lambda$CDM values by writing

$$\begin{align*}
\xi &= 1 + \delta \xi(x,k) \\
\omega &= -1 + \beta(x) \\
\Omega_M &= \Omega_M^{(0)} + \delta \Omega_M
\end{align*}$$ (5.16)

where $\Omega_M^{(0)} = \rho_M/\rho_{M+\Lambda}$, and $\rho_{\Lambda}$ is the energy density of the non-matter sector in the zeroth-order $\Lambda$CDM solution, i.e. it evolves as a perfect fluid with equation of state $\omega = -1$. Equating first-order parts of eq.(5.12) and suppressing some arguments for clarity:

$$\begin{align*}
\delta f' + q_{\text{GR}}(x) \delta f + 2f_{\text{GR}} \delta f &= \frac{3}{2} \Omega_M^{(0)}(x) \delta \xi(x,k) + \frac{3}{2} (1 + f_{\text{GR}}) \delta \Omega_M \\
&\quad + \frac{3}{2} (1 - \Omega_M^{(0)}) f_{\text{GR}} \beta
\end{align*}$$ (5.17)

where $q_{\text{GR}}(x) = \frac{1}{2} [4 - 3 \Omega_M^{(0)}(x)]$. In Appendix 5.9.1 we show that:

$$\delta \Omega_M = 3 \Omega_M^{(0)} (1 - \Omega_M^{(0)}) u(x)$$ (5.18)

where $u(x) = \int_0^x \beta(x') dx'$ such that $u(0) = 0$. Using eq.(5.18) in eq.(5.17) we obtain:

$$\begin{align*}
\delta f' + q_{\text{GR}}(x) \delta f + 2f_{\text{GR}} \delta f &= \frac{3}{2} \Omega_M^{(0)}(x) \delta \xi(x,k) + \frac{3}{2} (1 - \Omega_M^{(0)}) \left[ (1 + f_{\text{GR}}) 3 \Omega_M^{(0)} u + f_{\text{GR}} \beta \right]
\end{align*}$$ (5.19)
5. CORRECTIONS TO THE GROWTH RATE IN MODIFIED GRAVITY

It is more useful to work with the fractional deviation from the ΛCDM+GR prediction, which we define as:

$$\eta(x, k) = \frac{\delta f(x, k)}{f_{GR}(x)} = \frac{f(x, k)}{f_{GR}(x)} - 1$$ (5.20)

$\eta(x, k)$ should not be confused with the conformal time coordinate; we will largely use $x = \ln a$ for the time coordinate from here on anyway. Re-expressing eq.(5.19) in terms of this new variable (and using the zeroth-order part of eq.(5.12)):

$$\eta' + \eta \left[ f_{GR} + \frac{3}{2} f_{GR} \right] = \frac{3}{2} \Omega_M^{(0)} f_{GR} \delta S$$ (5.21)

$$\delta S = \Delta \xi + \frac{(1 - \Omega_M^{(0)})}{\Omega_M^{(0)}} \left[ 3 \Omega_M^{(0)} (1 + f_{GR}) u + f_{GR} \beta \right]$$ (5.22)

This simple first-order equation can be solved using an integrating factor, leading to the solution:

$$\eta(x, k) = \frac{3}{2} \int_{-\infty}^{x} \frac{\Omega_M^{(0)}(\tilde{x})}{f_{GR}(\tilde{x})} \delta S(\tilde{x}, k) K(x, \tilde{x}) d\tilde{x}$$ (5.23)

where:

$$K(x, \tilde{x}) = \exp \left\{ - \int_{\tilde{x}}^{x} d\tilde{x} \left[ f_{GR}(\tilde{x}) + \frac{3}{2} \Omega_M(\tilde{x}) \right] \right\}$$ (5.24)

We see that the solution for $\eta$ takes the form of an integral over a ‘source’ term $\delta S$ and a ‘kernel’ $K(x, \tilde{x})$. Crucially, note that the kernel only depends on the ΛCDM+GR background. This means that the kernel is theory-independent and simple to calculate. The entire theory-dependence of the modified growth rate is encoded in the source term $\delta S(x, k)$ which at each moment in time (or $x$) is a degenerate combination of $\delta \xi$, $\beta$ and $u$ – we will explore this degeneracy in §5.5.

As one might expect, we need to know the background solution we are expanding about (due to the factors of $\Omega_M^{(0)}$ and $f_{GR}$) in order to solve for the deviation $\eta$. This comes from solving eq.(5.12) with $\xi = 1$, which in general must be done numerically.

Let us interpret eq.(5.23) physically. It says that the fractional deviation from $f_{GR}$ is an integral from early times ($x \rightarrow -\infty$) up to the time of observation. One expects that an observer will detect a stronger signal from non-GR behaviour occurring at times recent to him/her than from non-GR behaviour occurring at high redshift. This sensitivity is encoded in the kernel $K(x, \tilde{x})$ – it gives an exponential suppression factor depending on the interval between the time of the
Figure 5.1: Examples of how the growth rate is affected by different source terms in eq.(5.12), where $\delta \xi = \xi - 1$. Effects on the growth rate are expressed as a percentage deviation from the GR prediction, i.e. $\eta = f/f_{GR} - 1$. Comparable effects on $f$ are obtained from broad, small-amplitude deviations from GR (right-hand panel) and violent, brief ones (left-hand panel). Note also the time lag between the peaks of the source and the response (particularly pronounced in the right-hand panel).

deviation from GR ($\tilde{x}$) and the time of observation ($x$). For deviations occurring at the time of observation there is no suppression, $K(x, x) = 1$. Causality imposes that $K(x, \tilde{x}) = 0$ for $\tilde{x} > x$; clearly observables cannot be affected by deviations from GR that occur after the time of observation.

We can also apply some physical interpretation to function $\delta S$ (eq.(5.22)) that sources corrections to the GR growth rate. It contains three contributing factors:

- The first term, $\delta \xi$, can be interpreted as the modified clustering properties stemming from the modified Poisson equation;

- The second term, $u(x) = \int_0^x \beta(x') dx'$, arises from the modified expansion history;

- The third term in $\beta(x)$ describes the instantaneous modified expansion rate, i.e. at the time that $\delta S$ is being evaluated.

We put eq.(5.23) to use by considering some toy examples. These will illustrate the response of the growth rate to generic deviations from GR; they are not intended to represent any particular theory of modified gravity. Let us first consider a simple case where the background expansion precisely matches that of a cosmological constant (i.e. $\beta = 0$), and only the behaviour of perturbations
is modified. Figs. 5.1 and 5.2 show the fractional deviation of the growth rate (defined in eq. (5.20)) triggered when matter clustering is modified by $\delta \xi \neq 1$:

1. A Gaussian (two examples of different widths and amplitudes are shown for comparison). Clearly we do not anticipate violent, transient, non-GR behaviour as depicted in the left-hand panel of Fig. 5.1 to occur in most viable modified gravity theories. We consider this example only to display the response function of the system$^1$;

2. A smoothed step function;

3. A cubic-order Taylor series in $[1 - \Omega_M^{(0)}(x)]$, where $\Omega_M^{(0)}(x)$ evolves as predicted by the $\Lambda$CDM+GR model.

One could imagine associating these examples to different kinds of gravity theories. For example, 2) might be a suitable model for modifications that switch on when the universe moves from radiation to matter domination, or from deaccelerated to accelerated expansion. A scalar field rolling in a suitable potential might produce slowly-growing effects such as 3). In each case we have assumed a scale-independent $\delta \xi$, but one could construct a more complicated function of $k$ and consider Figs. 5.1 and 5.2 as snapshots at a given scale. It is worth noting, however, that taking the quasistatic limit will tend to result in a scale-independent $\delta \xi$, see the examples in §5.6.2 $^2$.

We see from Fig. 5.1 that after a transient $\delta \xi$ source the growth rate gradually returns to its GR value, decaying approximately as $a^{-\frac{5}{2}}$ for a reasonably narrow Gaussian (the index can be inferred by considering eq.(5.24) during a matter-dominated epoch). The rate of return to GR is slightly suppressed at late times when the background expansion starts to accelerate, which acts to ‘freeze in’ perturbations. As expected from the form of eq.(5.23), both the amplitude and the duration of $\delta S$ play a role in determining the deviation from $f_{GR}$ observed at $z = 0$ – this can be seen by comparing the two panels of Fig. 5.1.

Sustained modifications such as Fig. 5.2 lead to growing deviations from GR, and hence will generally be more tightly-constrained. For example, the sources shown in Fig. 5.2 result in a $\sim 6\%$ effect on the growth rate, substantially larger than the sub-percent deviations shown in Fig. 5.1.

$^1$Where we are considering a narrow Gaussian as the closest physically-achievable approximation to an (unphysical) Dirac delta function. Recall that the response function of a system (e.g. in circuit theory) is its behaviour following a delta-function impulse forcing term.

$^2$This is not intended to be a rigorous statement in any sense. But $\delta \xi$ must be a dimensionless quantity, and dimensionless combinations involving the Hubble factor, such $\mathcal{H}/k$, arise more naturally in the field equations of a theory than combinations like $k \phi/G_N \rho$, for example.
Note that there is a time lag between changes in $\delta \xi$ and the response of the growth rate: this is most pronounced in the right-hand panel of Fig. 5.1. The input $\delta \xi$ is a broad Gaussian peaking at $z = 2.5$, but the response peaks later at $z \sim 1.75$. Similarly, in the right-hand panel of Fig. 5.2 the non-GR source begins to die away after $z \sim 0.5$, but $\eta(z)$ has insufficient time to follow suit. One could imagine generalizations of this situation, in which GR is the correct description of our universe today, but the effects of past non-GR behaviour still persist for a limited time (late-time changes in the dynamics of the dark sector were explored in [16]).

One may justifiably ask what kind of error is introduced by approximating eq.(5.12) as a linear equation. In fact the error is extremely small for the situations we are considering here. The full (non-linearized) evolution equation for $\delta f(x) = f(x) - f_{GR}(x)$ given by eq.(5.12) has a solution with the same form as eq.(5.23), but with a modified kernel:

$$K^{Full}(x, \tilde{x}) = \exp \left\{ - \int_{\tilde{x}}^{x} d\bar{x} \left[ f(\bar{x}) + \frac{3}{2} \Omega_M(\bar{x}) f_{GR}(\bar{x}) \right] \right\}$$ (5.25)

The first term of the integrand is now the modified growth rate instead of $f_{GR}$. That is to say:

$$K(x, \tilde{x}) = K^{Full}(\tilde{x}, x) \exp \left[ - \int_{\tilde{x}}^{x} f_{GR}(\bar{x}) - f(\bar{x}) d\bar{x} \right]$$

$$= K^{Full}(\tilde{x}, x) \exp \left[ \int_{\tilde{x}}^{x} \delta f(\bar{x}) d\bar{x} \right]$$ (5.26)

For the small deviations from $\Lambda$CDM+GR the exponential factor above is of order unity, so the linearized kernel is a very good approximation to the real kernel for $\eta$. For all the examples above the maximum error on $\eta(z)$ introduced by linearizing the growth rate equation is 0.2%.

### 5.4.2 The Response Function of the Observable Growth Rate $f\sigma_8(z)$

Extending the linear response analysis of the previous subsection to the observable growth rate, $f\sigma_8$, is fairly straightforward. We will continue to focus on scale-independent modifications. Using eq.(5.14), the fractional deviation of $f\sigma_8$ from its value in GR is given by:

$$\delta_{f\sigma} = \frac{\delta[f\sigma_8](z)}{f\sigma_8|_{GR}(z)}$$
Figure 5.2: As Fig. 5.1, but for sustained deviations from GR. For the right-hand panel $\delta \xi = 1 + 0.75 (1 - \Omega_M) - 1.5 (1 - \Omega_M)^2 + 0.75 (1 - \Omega_M)^3$.

$$\eta(z, k) = \frac{\delta \Delta M(z, k)}{\Delta M(z)|_{GR}}$$

(5.27)

where the first equality defines $\delta f \sigma$. We have already calculated the first term above, so we now tackle the second. For convenience we define a new symbol for this:

$$\delta \Delta(z) = \frac{\delta \Delta M(z)}{\Delta M(z)|_{GR}}$$

(5.28)

By perturbing eq.(5.10) we obtain:

$$\delta'' \Delta + \delta' \Delta \left(1 + \frac{\mathcal{H}'}{\mathcal{H}} + 2 f_{GR}\right) - \frac{3}{2} \Omega_M^{(0)} \delta S(x) = 0$$

(5.29)

where $\delta S(x)$ is given by eq.(5.22) again.

Eq.(5.29) can be solved for $\delta' \Delta$ using a simple integrating factor, then integrated once more to obtain $\delta \Delta$. Reversing the order of the integrations allows us to write the solution in the form ‘source $\times$ kernel’, as we did for $\eta$ in the previous subsection:

$$\delta \Delta(x) = \frac{3}{2} \int_{-\infty}^{x} \delta S(\tilde{x}) I(x, \tilde{x}) d\tilde{x}$$

(5.30)

where, as before, the kernel $I(x, \tilde{x})$ is a function of the background cosmology.
only:

\[
I(x, \tilde{x}) = \int_{\tilde{x}}^{x} dy \exp \left[ - \int_{\tilde{x}}^{y} d\tilde{x} \left( 2 - \frac{3}{2} \Omega_M(\tilde{x}) + 2 f_{GR}(\tilde{x}) \right) \right]
\]  

(5.31)

and we have used the Friedmann equation for the GR background en route.

Finally, using eq.(5.27), the fractional deviation of \( f \sigma_8 \) from its \( \Lambda \)CDM value can be expressed in a Green’s function-like form:

\[
\frac{\delta[f(x)\sigma_8(x)]}{f(x)\sigma_8(x)|_{GR}} = \int_{-\infty}^{x} \delta S(\tilde{x}) G(x, \tilde{x}) d\tilde{x}
\]

(5.32)

where the kernel \( G(x, \tilde{x}) \) is:

\[
G(x, \tilde{x}) = \frac{3}{2} \Omega_M(\tilde{x}) \left[ \frac{K(x, \tilde{x})}{f_{GR}(\tilde{x})} + I(x, \tilde{x}) \right]
\]

(5.33)

and the factors \( K(x, \tilde{x}) \) and \( I(x, \tilde{x}) \) are given by eqs.(5.24) and (5.31). Fig. 5.3 shows uses of this formula. The left-hand panel shows the same case considered in the left panel of Fig. 5.1, where \( \delta \xi \) has a Gaussian form. Whilst in Fig. 5.1 \( \eta(z) \) declined to zero, \( \delta f_{\sigma}(z) \) instead settles to a constant. The difference in behaviour arises from the second term of eq.(5.27), as follows: during the time the source is ‘switched on’ the growth of density perturbations (\( \delta M \)) is either enhanced or suppressed relative to the \( \Lambda \)CDM+GR case. When the source switches off density perturbations return to growing at the GR rate, but their absolute value has now been shifted from that of a pure \( \Lambda \)CDM universe. This shift is the constant term seen in the figure.

The right-hand panel of Fig. 5.3 shows the effect of allowing the background effective equation of state to evolve too, ie. \( \beta(x) \neq 0 \). Even after the Gaussian transient has passed, \( \delta S < 0 \) will cause \( \delta f_{\sigma} \) to decrease, see eq.(5.29). \( \delta f_{\sigma} \) begins to upturn at late times again when \( \delta S > 0 \) for \( z < 0.5 \), but again we see the time lag between the low-redshift peak of \( \delta S \) and the response of \( \delta f_{\sigma} \). It is worth noting that a very small modification to the background (see caption for the values used) soon rapidly erases the deviations caused by the transient \( \delta \xi \) modification – such is the power of an extended integration time.

Eqs.(5.32) and (5.33) are a key result of this chapter, so let us summarize what has been achieved. Accepting that the \( \Lambda \)CDM+GR model is an excellent description of the universe at leading order, we have found a general way to calculate the impact that modifications to the General Relativistic field equations have on the observable growth rate of structure. All the modifications are encapsulated in a single function \( \delta S(x) \), which can be matched to a specific gravity theory or
Figure 5.3: Lower panels show the fractional deviation of the density-weighted growth rate from its ΛCDM+GR model caused by the source functions in the upper panels, see eq.(5.32). The left panel shows the same Gaussian considered in the left panel of Fig. 5.1. The right panel shows the effect of adding an evolving background of the kind considered in §5.5, specifically, $\beta(x) = 0.05(1-e^{x})$.

Constrained in a model-independent, phenomenological manner. On a practical note, we re-iterate that the kernel $G(x, \tilde{x})$ is a function of the zeroth-order ΛCDM cosmology only, and hence is relatively simple to compute. It only needs to be calculated once and stored (as a function of $x$ and $\tilde{x}$) to allow rapid calculation with different source functions. Furthermore, for the examples considered in this chapter we have found our method is remarkably accurate; the maximum error incurred using our linearized treatment (instead of the exact calculation) is of order $0.2\%$, which is well within the error bars on $f\sigma_8(z)$ forecast for next-generation galaxy surveys.

### 5.5 The Degeneracy Between $\delta\xi$ and $\beta$

In the previous sections we have seen how we can, for small deviations from ΛCDM, divide the evolution of the observable growth rate into a response kernel and a source, $\delta S$. We now focus on the source term, which incorporates effects from both the clustering properties of the theory (via $\delta\xi$) and the expansion rate (via $\beta$). We want to understand how the degeneracy between $\delta\xi$ and $\beta$ evolves in time.

A first step towards understanding the degeneracy is to take the simplest model for $\delta\xi$ and $\beta$: two step functions that switch on at some transition redshift.
Then for \( z < z_T \) we can rewrite the source term in a simplified form:

\[
\delta S = \delta \xi_0 + \alpha(x) \beta_0 \quad \text{for} \quad z < z_T
\]

\[
\alpha(x) = \left( 1 - \Omega_M^{(0)} \right) \Omega_M^{(0)} \left[ 3 \Omega_M^{(0)} x (1 + f_{GR}) + f_{GR} \right]
\]

(5.34)

where \( \delta \xi_0 \) and \( \beta_0 \) are constants.

In this model a given constraint on \( \delta S \) at each redshift corresponds to a degenerate line in the \((\delta \xi_0, \beta_0)\) plane, whose slope is given by \( \alpha(z) \). This then poses a problem if, say, we wish to infer constraints on the modified clustering properties, \( \delta \xi_0 \). If we make the naive assumption that \( \delta \xi_0 \) and \( \beta_0 \) are uncorrelated, the the uncertainty on \( \delta \xi_0 \) is given approximately by:

\[
\sigma_{\delta \xi_0} \approx \left( \sigma_{\delta S}^2 + \alpha^2(x) \sigma_{\beta_0}^2 \right)^{\frac{1}{2}}
\]

(5.35)

The error on \( \delta S \) will be determined by the experimental errors on measurements of \( f \sigma_8 \), using a discretized version of eq.(5.32) (see §5.7). However, we see that the error we obtain on \( \delta \xi_0 \) will be further inflated by our uncertainty regarding the background expansion rate, \( \sigma_{\beta_0} \).

Of course we can constrain \( \beta_0 \) through other geometric probes, but eq.(5.35) suggests that there is an additional step we can take to help break the degeneracy. In Fig. 5.4 we plot \( \alpha(z) \) as a function of redshift, noting that it has a root at \( z \sim 0.3 \). Targeting this ‘sweet spot’ will minimize the errors on \( \delta \xi_0 \). The location of the sweet spot is dependent on the form we assumed for \( \beta(x) \), so further investigation is needed to determine the extent to which it can vary. Likewise further work is needed to determine the optimum redshift ranges for combining geometric and growth probes, such that the constraints coming from each data set are maximally orthogonal.

A more motivated choice for \( \beta(x) \) might be a form that mimics the well-known CPL parameterization of the equation of state: \( \omega(x) = \omega_0 + \omega_a (1 - a) \). This is reproduced by the choice:

\[
\beta(x) = \beta_0 + \beta_1 (1 - e^x)
\]

(5.36)

where \( \beta_0 \equiv \omega_0 + 1 \) \quad \( \beta_1 \equiv \omega_a \) (5.37)

---

1 Any expansion history can be reproduced by either a dark energy or modified gravity model, so it is frequently advertized that the behaviour of perturbations/clustering will act as a discriminator between the two scenarios.
These functions describe the extent to which modifications to the equation of state of the non-matter sector (i.e. \(\omega X \neq -1\)) source corrections to the growth rate of GR, see eqs. (5.34) and (5.38). Both functions possess a root in the range \(0.3 \lesssim z \lesssim 0.5\), meaning that this redshift range should offer the tightest constraints on modified clustering properties (through constraints on \(\delta\xi\)).

The source \(\delta S\) then has the form:

\[
\delta S = \delta\xi + \alpha_0(x) \beta_0 + \alpha_1(x) \beta_1
\]

\[
\alpha_0(x) = \left(1 - \Omega_M^{(0)}\right) \left[3\Omega_M^{(0)} x (1 + f_{GR}) + f_{GR}\right]
\]

\[
\alpha_1(x) = \left(1 - \Omega_M^{(0)}\right) \left[3\Omega_M^{(0)} (1 + f_{GR})(x + 1 - e^x) + f_{GR}(1 - e^x)\right] \tag{5.38}
\]

The point made above remains true – we now have a degenerate plane in the parameter space of \(\{\delta\xi, \beta_0, \beta_1\}\). \(\alpha_0(x)\) is the same as eq. (5.34) and the left-hand panel of Fig. 5.4; \(\alpha_1(x)\) is shown in the right-hand panel. Note that it has a root at \(z \sim 0.5\). Again, we can minimize the degeneracies through judicious choices of redshift bins and combination with geometric probes [32].

Throughout this discussion we have been assuming that \(\beta\) and \(\delta\xi\) are completely unconnected. In reality, most of the modified gravity theories in the current research literature contain a handful of parameters which are likely to appear in both \(\beta\) and \(\delta\xi\). We will see in §5.7 that throwing away this information has a profound impact on the constraints obtainable from the growth rate.
5.6 Deriving $\delta \xi$ from the Parameterized Post-Friedmann Formalism

5.6.1 The Quasistatic Limit of PPF

In this chapter we aim to keep our treatment of growth observables in modified gravity as theory-independent as possible. So far we have required only that the quasistatic approximation be valid for some range of scales, and made use of the widely-applicable quasistatic equations (5.6) and (5.7).

Nevertheless, one may wish to know how to map our general results onto a specific theory. We can achieve this using the PPF formalism of chapter 2. The quantity $\xi = \mu/\gamma$ that appears in eq.(5.12) can be written in terms of the PPF coefficient functions, thereby linking it to specific theories of gravity. For the present we will consider theories which:

a) contain second-order time derivatives in their equations of motion (a generic stability criterion [312]),

b) contain one new non-GR degree of freedom, which we denote by $\chi$ (this could be a spin-0 perturbation of a new field, for example).

The PPF framework also encompasses theories that lie outside these restrictions, but we will not present such extensions here.

The quasistatic form of the extended PPF field equations in the conformal Newtonian gauge is (compare with eqs.(2.18-2.21)):

\begin{align}
-a^2 \delta G^0_0 &= \kappa a^2 G \rho M \delta M + A_0 k^2 \Phi + \alpha_0 k^2 \hat{\chi} \\
-a^2 \delta G^i_0 &= \nabla_i \left[ \kappa a^2 G \rho M (1 + \omega_M) \theta_M + B_0 k^2 \Phi + \beta_0 k^2 \hat{\chi} \right] \\
a^2 \delta G^i_i &= 3 \kappa a^2 G \rho M \Pi_M + C_0 k^2 \Phi + \gamma_0 k^2 \hat{\chi} \\
a^2 \delta \tilde{G}^i_j &= \kappa a^2 G \rho M (1 + \omega_M) \Sigma_M + D_0 \Phi + \epsilon_0 \hat{\chi}
\end{align}

where $\delta \tilde{G}^i_j = \delta G^i_j - \frac{1}{3} \delta^i_0 \delta G^k_k$ and $D_{ij} = \nabla_i \nabla_j - \frac{1}{3} \delta_{ij} \nabla^2$. The hat over $\hat{\chi}$ indicates that it is a gauge-invariant combination of perturbations constructed using the algorithm of §2.2.3.2, i.e. it contains both $\chi$ and metric perturbations.

For many theories the equation of motion (hereafter e.o.m) of $\chi$ corresponds to a conservation equation\(^1\). This includes scalar-tensor gravity, quintessence, Einstein-Aether theory, Ho\'rava-Lifschitz gravity (see Table 2.1 for references)

\(^1\)More specifically, it is the time-like component of the equation enforcing the divergenceless nature of tensor additions to the field equations of GR. In §2.2.5 we referred to these as Type 1 theories. Theories with explicit matter-scalar coupling introduce some subtlety here.
and most theories that fall into the Horndeski class [97, 102, 134, 159]. When expressed in terms of the PPF coefficients, the quasistatic limit of the e.o.m. for $\chi$ is:

$$\dot{\chi} \left( \dot{\alpha}_0 + k\beta_0 \right) + \Phi \left[ \dot{A}_0 + kB_0 \right] = 0 \quad (5.43)$$

Combining eq. (5.43) with eqs. (5.39) and (5.42), the connection between the quasistatic $\{\mu, \gamma\}$ parameterization and the PPF functions is:

$$\mu(z, k) \approx \left\{ 1 + \frac{A_0}{2} - \frac{\alpha_0}{2} \left( \frac{\dot{A}_0 + kB_0}{\dot{\alpha}_0 + k\beta_0} \right) \right\}^{-1} \quad (5.44)$$

$$\gamma(z, k) \approx \left\{ 1 - D_0 + \epsilon_0 \left( \frac{\dot{A}_0 + kB_0}{\dot{\alpha}_0 + k\beta_0} \right) \right\}^{-1} \quad (5.45)$$

where we have assumed the anisotropic stress of matter to be negligible for the late-time universe.

Below we give examples of the ‘modified clustering’ source $\xi(k, a) = 1 + \delta\xi(k, a)$, for a number of commonly-discussed gravity theories. Although eqs. (5.44) and (5.45) were derived for theories with only one non-GR degree of freedom, more complicated theories can still be mapped onto specifications of $\mu, \gamma$ and $\xi$ under the quasistatic approximation. However, they need to be treated on a case-by-case basis rather than via eqs. (5.44) and (5.45) – see the example of DGP below.

We should also highlight an interesting subtlety here with regards to the popular family of $f(R)$ models. It is often assumed that any results pertaining to scalar-tensor theories automatically incorporate $f(R)$ gravity, since a conformal mapping exists between the two classes of theories (see §2.6.1). Whilst this is true at the action level, the perturbed e.o.m.s for the new degree of freedom derived from their actions are not equivalent. The e.o.m. of scalar-tensor theory is a conservation equation of the kind described above; the e.o.m. for the ‘scalaron’, $\delta f_R$ originates from the trace of the $f(R)$ gravitational field equations. Hence eqs. (5.44) and (5.45) do not apply to $f(R)$ gravity. Nevertheless we can still take the quasistatic limit of the theory, see below.

---

1 The authors of [281] recently derived a result relating the two quasistatic functions $\{\mu, \gamma\}$ for single scalar field-type models (‘Horndeski theories’). When converted into our notation, their result becomes the statement that the numerator of $\mu$ must be equal to 1, which is manifest in eq.(5.44).

2 That is, the equation of motion of the scalar field is equivalent to a conservation law for the effective stress-energy tensor of the scalar.
5.6.2 Examples

Here we present expressions for the ‘clustering function’ $\xi(x)$ for some specific cases. For brevity we present only identities for $\xi(x)$, but all these theories will have an associated $\beta(x) \neq 0$ too. The relevant actions and references were given in §2.3. We note that $\xi$ is scale-independent in all the cases presented here, except possibly scalar-tensor theory, where it depends on the choice of potential $V(\phi)$.

1. Scalar-Tensor Theory

$$\xi(a, k)_{ST} = \left[ \phi + \frac{\dot{\phi}}{\mathcal{H}} - \frac{Y}{Z} \right] \times \left[ \phi + \frac{1}{2} \left( 1 - \frac{a^2}{k^2} V''(\phi) \right) \left( \frac{\dot{\phi}}{\mathcal{H}} - \frac{Y}{Z} \right) \right]^{-1}$$  
(5.46)

where $Y = \frac{a^2}{k^2} V''(\phi) \left( \frac{\ddot{\phi}}{\mathcal{H}} - \frac{\dot{\mathcal{H}}}{\mathcal{H}} \ddot{\phi} \right) + \frac{\dot{\phi}^2}{\mathcal{H}k^2} V''(\phi) + \dot{\phi} \left( \frac{\omega(\phi) \dot{\phi}}{\mathcal{H} \dot{\phi}} - 3 \right)$  
(5.47)

$$Z = \frac{a^2}{k^2} \left( V''(\phi) \dot{\phi} + 2\mathcal{H}V'(\phi) \right) + \omega \frac{\dot{\phi}}{\phi} - \mathcal{H}$$  
(5.48)

The scale-dependence of $\xi$ in this case has arisen because we have been careful not to make any assumptions about the form of the potential $V(\phi)$. It is likely that once a form is chosen for $V(\phi)$ further terms can be dropped due to the quasistatic approximation.

2. Brans-Dicke Theory

In the Brans-Dicke case of scalar-tensor theory ($\omega=$constant, $V(\phi) = 0$) eq.(5.46) simplifies considerably:

$$\xi_{BD}(a) = \left[ \phi + \frac{\dot{\phi}}{\mathcal{H}} - X \right] \times \left[ \phi + \frac{1}{2} \left( \frac{\dot{\phi}}{\mathcal{H}} - X \right) \right]^{-1}$$  
(5.49)

where $X = \frac{\dot{\phi}}{\mathcal{H}} \left( \frac{\omega_{BD} \dot{\phi}}{\mathcal{H} \dot{\phi}} - 3 \right)$  
(5.50)

In particular, for a scalar field following a power-law evolution $\phi \propto a^n$:

$$\xi_{\phi^n} \approx \frac{n(\omega_{BD} + 2) - 1}{n(\omega_{BD} + 1) - 1}$$  
(5.51)
5. CORRECTIONS TO THE GROWTH RATE IN MODIFIED GRAVITY

For the case $\omega_{BD} \gg 1$ (note that GR is recovered in the limit $\omega_{BD} \to \infty$):

$$\delta \xi_{BD}(a) \sim \frac{1}{\omega_{BD}} \quad (5.52)$$

3. \textbf{f (R) Gravity}

$$\xi_{fR}(a) = \frac{1}{3} \left( \frac{4}{f_R} - 1 \right) \quad (5.53)$$

where $f_R = d f (R) / dR$. Note that we are implicitly assuming some form for $f (R)$ that is able to recover GR without a chameleonic screening mechanism. The authors of [304] have shown that the need to achieve screening of the Sun and/or the Galaxy via a chameleon mechanism would render the modifications to the linear-scale growth rate unobservably small.

4. \textbf{Einstein-Aether Theory}

$$\xi_{AE}(a) = \frac{1 - \frac{\dot{c}}{ct}(c_1 + c_3)}{1 - \alpha + c_1 \left( 1 - \frac{\dot{c}}{ct} \right)} \quad (5.54)$$

$$\delta \xi = -\frac{\left[ \alpha - (c_1 + c_3) \frac{\dot{c}}{ct} + c_1 \left( \frac{\dot{c}}{ct} - 1 \right) \right]}{\alpha - 1 + c_1 \left( \frac{\dot{c}}{ct} - 1 \right)} \quad (5.55)$$

where $c_i$ are parameters of the theory and $\alpha = c_1 + 3c_2 + c_3$.

5. \textbf{DGP}

$$\xi_{DGP}(a) = 1 + \frac{1}{3 \beta} \quad (5.56)$$

where $\beta = 1 - \frac{2}{3} \mathcal{H} \tilde{r}_c \left( 2 + \frac{\dot{c}}{\mathcal{H} \tilde{c}^2} \right) \quad (5.57)$

and $\tilde{r}_c = r_c / a$ is the comoving crossover scale.

6. \textbf{Hořava-Lifschitz Gravity}

$$\xi_{HL}(a) = \frac{1}{1 + \frac{\dot{c}}{2 \mathcal{H} \tilde{c}^2}} \quad (5.58)$$

$$\delta \xi \approx -\frac{\dot{c}}{2 \mathcal{H} \tilde{c}^2} \quad (5.59)$$
where $\tilde{\alpha}$ is a parameter appearing in the potential of the ‘healthy’ theory proposed by Blas, Pujolàs and Sibiryakov [50], and the last equality assumes $\tilde{\alpha} \ll 1$. In the quasistatic regime the Planck mass-suppressed higher derivative operators of Hořava-Lifschitz gravity are safely negligible.

### 7. Horndeski’s Theory

Even in the simplified quasistatic limit, the relevant expressions for Horndeski’s most general second-order scalar-tensor theory are non-trivial. Therefore we have chosen to display them at the end of this chapter (§5.9.2), borrowing heavily from the results of [52].

#### 5.7 Constraints from Future Data

Experimental results to date have already told us that only small deviations from GR are still acceptable, at least on the lengthscales and timescales that our observations have probed so far. What do these tight constraints imply for theories of modified gravity? We can use the model-independent formalism of this chapter to answer this question in a general way. We will not attempt a detailed or experiment-specific analysis here, but instead will consider some general order-of-magnitude calculations.

Consider a galaxy survey divided into $N$ redshift bins with centers $x_i$ (recall $x = \ln a$) and widths $w_i$, where $i = 1, \ldots, N$. Eq.(5.32) must be discretized for use:

$$
\delta f_\sigma = \sum_{j=1}^{i} \left[ \int_{x_j + \frac{w_j}{2}}^{x_j + \frac{w_j}{2}} G(\tilde{x}, x) \, d\tilde{x} \times \delta S_j \right] = \sum_{j=1}^{i} G_{ij} \delta S_j \tag{5.60}
$$

$\delta f_\sigma$ and $\delta S_i$ are vectors containing the mean values of $\delta f_\sigma(x)$ and $\delta S(x)$ in each redshift bin. If we choose $i = 1$ to be the highest-redshift bin, $G_{ij}$ must be a lower-left triangular matrix, i.e. it has vanishing entries for $j > i$; this is the causality requirement discussed in §5.4.

Inverting eq.(5.60), the errors on $\delta S_i$ are given by:

$$
C_{\delta S} = \langle \delta S \, \delta S^T \rangle = \langle G^{-1} \, \delta f_\sigma \, (\delta f_\sigma)^T \, (G^{-1})^T \rangle \tag{5.61}
$$

$$
= G^{-1} \, C_{f_\sigma} \, (G^{-1})^T \tag{5.62}
$$

where $C_{\delta S}$ and $C_{f_\sigma}$ are covariance matrices, the latter being determined by the survey specifications. Fig. 5.5 shows projected constraints on $\delta S$ obtainable with a representative next-generation galaxy redshift survey [15, 217].

Fig. 5.5 represents what can be achieved in a model-independent analysis; we
5. CORRECTIONS TO THE GROWTH RATE IN MODIFIED GRAVITY

Figure 5.5: Forecast constraints on $\delta S$ for a typical next-generation galaxy survey, where $\delta S$ sources deviations from the GR growth rate (see eq.(5.22)). This figure represents the most agnostic case, where we assume only that $\delta S$ is monotonic. The contours shown represent 1 $\sigma$ and 2 $\sigma$ uncertainties.

have assumed only that $\delta S$ must vary smoothly between redshift bins. Tighter bounds could be obtained from the same data if we chose to specialize our analysis to a particular theory. Most modified gravitational actions contain a small handful of parameters, and the appearance of these parameters in the gravitational field equations will restrict how $\delta S$ can vary between redshift bins.

In this sense Fig. 5.5 represents the ‘worst-case’ scenario, where we remain as agnostic as possible about the theories being tested. The situation is very different, then, to that of PPN. In PPN the same constraints are obtained whether one directly tests theory-specific parameters, or first constrains the PPN parameters ($\gamma, \beta, \ldots$) and then uses the appropriate correspondences to map these onto specific theories. The loss of constraining power experienced here is a consequence of using functions of time and scale, rather than simple constants as in PPN. Alas this is an unavoidable feature of testing gravity in a time-dependent cosmology. We note that [290] found a similar loss of constraining power when testing an comparable parameterized framework for quintessence models.
5.8 Conclusions

It has been argued [216] that current measurements of the growth rate are somewhat low compared to those expected from the current favoured theory of large scale structure – the ΛCDM model. The discrepancy is not particularly statistically significant, but it points to the need for a clear understanding of factors influencing the growth rate, both fundamental (eg. modified gravity/dark energy) and astrophysical (bias/non-linearities).

In this chapter we have developed an efficient and widely-applicable method of calculating corrections to the growth rate in non-GR gravitational scenarios. In particular, our method elucidates the mechanism through which modifications to the field equations feed into the growth rate, which is sensitive to a degenerate combination of clustering and expansion-related affects. A future extension of this work could be to trace the origin of the functions $\beta(x)$ and $\delta\xi(x)$ back to the level of the action, to determine if there are any general relations between them.

The tools developed here have clear applications in determining how well next-generation galaxy surveys like Euclid [2, 15, 217] and LSST [3] will test GR on cosmological scales. The generality of our parameterized approach is both its strength and its downfall – indeed, we have seen several times in this thesis that one must often strike a balance between generality and constraining power.

However, the work of this chapter has highlighted a potential way forward. In §5.5 we saw that error-bars on $\delta\xi$ were minimized in particular redshift windows. Although there was some theory-dependence involved in the precise numbers, it may be that families of theories all give similar results. In that case it would be wise to remember these considerations when devising survey strategies and combining growth rate data with complimentary probes.

Whilst the generalized constraints forecast in Fig. 5.5 may seem a little disheartening, we ought to remember that the use of parameterized frameworks in cosmology is a comparatively recent technique, and is likely to be subject to further improvement. Ideas from Principal Component Analysis (PCA) may be of use in developing optimal data compression methods. Likewise, it would be interesting to see if our approach of perturbing in solution space about ΛCDM can be applied in other contexts, for example weak gravitational lensing (§1.2.3.1) or detection of the ISW effect (§1.2.1.3); the best-constrained eigenmodes of these observables may turn out to be largely orthogonal to the growth rate eigenmodes. Further work is in progress [32] to answer these questions.
5.9 Supporting Material for Chapter 5

5.9.1 Perturbation of $\Omega_M(x)$

Here we derive the relation eq.(5.18).

First observe that the general fluid evolution equation:

$$\rho' = -3\rho(1 + \omega(x))$$

has the following solution, where we write $\omega(x) = -1 + \beta(x)$:

$$\rho(x) = \rho(0) e^{-3 \int_0^x \beta(x') dx'}$$

Now consider two universes. The first is a perfect ΛCDM model. In the second, the non-matter sector is not a true a cosmological constant; there is a modification to gravity that can be recast in the form of a perfect fluid with a (likely evolving) equation of state. Recall that nearly any gravity theory can be written in this form, even if the expression for the effective equation of state is extremely complex. Observational viability restricts that the equation of state can only differ from $-1$ by a small amount, ie. $\beta$ in eq.(5.64) must be small for all $x$.

In the first universe we have:

$$\Omega^{(0)}_M(x) = \frac{\rho_M(x)}{\rho_M(x) + \rho_{\Lambda 0}} = \frac{1}{1 + R e^{3x}}$$

where we have used $\rho_M(x) = \rho_{M0} e^{-3x}$ and defined $R = \rho_{\Lambda 0} / \rho_{M0}$. In the second universe, denoting the energy density of the effective fluid by $\tilde{\rho}_X(x)$ and analogously defining $\tilde{R} = \tilde{\rho}_{X0} / \tilde{\rho}_{M0}$:

$$\tilde{\Omega}_M(x) = \frac{\tilde{\rho}_M(x)}{\tilde{\rho}_M(x) + \tilde{\rho}_X(x)} = \frac{1}{1 + \tilde{R} \exp \left[3x - 3 \int_0^x \beta(x') dx'\right]}$$

where we have used eq.(5.64).

We are interested in the small perturbation to the matter fraction $\Omega_M(x)$ that results from perturbing about a ΛCDM universe. This is given by the difference between eqs.(5.65) and (5.66). Since $\beta$ is small at all times we can expand:

$$e^{-3 \int \beta(x') dx'} \approx 1 - 3 \int \beta(x') dx' = 1 - 3 u(x)$$

(5.67)
where the last equality defines \( u(x) \). Furthermore, we argue that \( R = \tilde{R} \), because the ratio of the non-matter energy density to the matter energy density is an experimentally-determined quantity. Whether we are living in the \( \Lambda \)CDM or non-\( \Lambda \)CDM universe, we would simply measure one value for this ratio and call it \( R \).

Collecting expressions then, we have:

\[
\delta \Omega_M(x) = \tilde{\Omega}_M(x) - \Omega_M^{(0)}(x) \\
= \frac{1}{1 + Re^{3x}[1 - 3u(x)]} - \frac{1}{1 + Re^{3x}} \\
\approx \frac{3Re^{3x}u(x)}{[1 + Re^{3x}]^2} \\
= 3u(x)\Omega_M^{(0)}(x) \left( 1 - \Omega_M^{(0)}(x) \right)
\]

where the last step uses eq.(5.65) to eliminate \( Re^{3x} \) in favour of \( \Omega_M^{(0)} \). This is the result stated in §5.4.1.

### 5.9.2 Deviation Source Term for Horndeski’s Theory

Given the complexity of the Horndeski Lagrangian [97, 102, 159], it is beyond the scope of this thesis to carry out the intricate reduction to the quasistatic limit. Fortunately, this calculation has recently been presented in [52], from which we adopt results.

In the expressions below, \( m_0^2 \) and \( \bar{M}_i \) are parameters and \( \Omega(t) \) and \( c(t) \) are functions of time that appear in the Lagrangian of the Effective Field Theory of dark energy - we refer to [52, 54] for precise definitions (see also [138, 144] for an equivalent formalism). In this appendix only we will use dots to denote derivatives with respect to physical time instead of conformal time. The notational equivalence between metric potentials used in this thesis and [52] is \( \Phi \equiv \psi \), \( \Psi \equiv \phi \).

\[
\xi_{HD} = \left( B\pi C_\Phi - B\Phi C_\pi - B\Phi C_\pi C_\psi \frac{a^2}{k^2} \right) \\
\times \left[ A_\Phi \left( B\Phi C_\pi + B\Phi C_\pi C_\psi \frac{a^2}{k^2} - B\pi C_\psi \right) + A_\pi \left( B\pi C_\psi - B\Phi C_\phi \right) \right]^{-1} \tag{5.69}
\]

\[
A_\Phi = 2(m_0^2\Omega + \bar{M}_2^2) \tag{5.70}
\]

\[
A_\pi = -(m_0^2\bar{\Omega} + \bar{M}_1^3) \tag{5.71}
\]

\[
B\Phi = -1 \tag{5.72}
\]
5. CORRECTIONS TO THE GROWTH RATE IN MODIFIED GRAVITY

\[ B_\psi = 1 + \frac{\dot{M}_2}{\Omega m_0^2} \]  

\[ B_\pi = \frac{\dot{\Omega}}{\Omega} + \frac{\dot{M}_2}{\Omega m_0^2} \left( H + 2 \frac{\dot{M}_2}{M_2} \right) \]  

\[ C_\Phi = m_0^2 \dot{\Omega} + \dot{M}_2 \left( H + 2 \frac{\dot{M}_2}{M_2} \right) \]  

\[ C_\psi = -\frac{m_0^2}{2} \dot{\Omega} - \frac{\dot{M}_1}{2} \]  

\[ C_\pi = c - \frac{\dot{M}_1}{2} \left( H + 3 \frac{\dot{M}_1}{M_1} \right) + \dot{\bar{M}}_2 \left( H^2 - \dot{H} + 2H \frac{\dot{M}_2}{M_2} \right) \]  

\[ C_{\pi_2} = \frac{m_0^2}{4} \Omega \dot{R}^0 - 3c\dot{H} + \frac{3\dot{M}_1}{2} \left( 3\dot{H} \frac{\dot{M}_1}{M_1} + \dot{H} + 3H\ddot{H} \right) + 3\dot{\bar{M}}_2^2 \dot{H}^2 \]
Chapter 6
Conclusions

We are living in a dangerous time. The past three decades of research in cosmology have been a spectacular success: what was once a highly speculative science has produced a detailed standard paradigm for how our universe began and continues to evolve. The mathematics underlying this paradigm is understood analytically (as far as is tractable), calibrated against computer simulations, and found to be in resounding agreement with a whole toolbox of observations. Furthermore, the $\Lambda$CDM+GR model is pleasingly economical, requiring only six cosmological parameters and a single field to mediate the force of gravity. It is tempting to simply accept that we live in a universe with an unexplainably small cosmological constant, and move onto other astrophysical/cosmological questions.

The danger is that we lull ourselves into thinking that we have have full mastery of the gravity theory that goes hand-in-hand with $\Lambda$CDM, Einstein’s General Relativity. In fact we already know that this cannot be the case, for two reasons. Firstly, the nonrenormalizability of GR – and its ensuing refusal to gel with quantum theory – suggests that it cannot be valid up to Planck energies. Secondly, even in purely classical calculations GR harbours singularities that signal its loss of predictability, namely those found at the Big Bang and in black hole formation. Both these problems hint that there is new physics awaiting us at scales beyond those that we can currently probe, and by implication, corrections to the gravitational action$^1$. I have deliberately chosen to use the ambiguous word ‘scale’ here. If GR is a limit of a larger theory, it is not obvious what might be the correct quantity that divides a GR regime from a non-GR regime. A common intuition is that of an energy scale, but as discussed in §1.3.5, other quantities are

\[1\] Of course it is possible that gravity is modified and we still have a tiny cosmological constant. This thesis has focused principally on modified gravity theories as alternatives to the cosmological constant. However, it is worth noting that not all theories originally constructed with this purpose in mind have achieved their aim (i.e. after further mathematical development it was realized that they still required a cosmological constant to drive acceleration) [304].
also considered for signalling new gravitational phenomena, e.g. density, length, curvature. Clearly these quantities are interconnected – we expect a dense region to be associated with large curvature and a small characteristic lengthscale. The question at hand is which of these quantities is fundamental to the theory, and which is a consequence of the field equations\(^1\).

Faced with this paradoxical situation – the simultaneous observational successes and theoretical failures of the ΛCDM+GR model – let us take stock of the scenarios in which GR has and has not been tested. Fig. 6.1 shows an attempt to visualize this. It places gravitational systems in a two-dimensional parameter space, where the \(x\)-axis indicates the characteristic gravitational potential of the system (divided by \(c^2\) to obtain a dimensionless quantity) and the \(y\)-axis indicates the associated characteristic spacetime curvature. The \(y\)-axis quantity is not the Ricci scalar itself (since that vanishes in vacuum), but a proxy with the correct magnitude and dimensions to be representative of \(R^2\).

For example, an isolated spherically symmetric body of mass \(M\) and radius \(R\) would be placed on the plot according to the potential at its surface \(G_N M / R\), and the curvature proxy used would be \(G_N M / R^3\). For a test particle orbiting a central mass \(M_c\) at radius \(r_o\), the potential and curvature proxy are \(G_N M_c / r_o\) and \(G_N M_c / r_o^3\) respectively. The galaxy lines (apart from the Milky Way) are determined from rotation curves, using the Newtonian approximation \(v^2 = GM / R \sim \Phi / c^2\). The cluster lines were determined from published mass profiles [156, 262]. Ascribing potentials and curvatures to the large-scale matter distribution (cyan lines) is a little more involved - full details are given in [33]. Qualitatively, the curvature assigned to a potential well depends on the matter overdensity \(\delta_M\), which can be obtained from the matter power spectrum.

Acronyms are given in the figure caption. The label ‘CMB first peak’ indicates the potential and curvature on the scale of the sound horizon in the primordial plasma at the time of photon decoupling. The horizontal lines labelled ‘BBN’ and ‘Lambda’ indicates the curvature of the universe at the time of Big Bang Nucleosynthesis and in the far future respectively, calculated using the Ricci scalar for an FRW metric: \(R = 6 / a^2 (3 H^2 + \dot{H})\). Note that there is no value of the potential associated with these (there is no well-defined way of meaningfully assigning a potential to the universe ‘as a whole’). The three cyan lines labelled by redshifts represent the matter power spectrum at different stages; large scales correspond to the vertical sections. Note that the large scales sit at a potential ~

\(^1\)Of course one can then start to debate the importance of this distinction.

\(^2\)One can think of the \(y\)-axis as indicating the magnitude of the largest element of the Riemann curvature tensor near to the system in question.
Figure 6.1: A two-dimensional parameter space for gravitational regimes that exist in our universe, quantified in terms of their characteristic gravitational potential ($x$-axis) and approximate spacetime curvature ($y$-axis). Acronyms used are as follows: SS = planets of the Solar System, PSRs = binary pulsar systems, MS = main sequence stars, WD = white dwarfs, NS = neutron stars, XRB = X-ray binary systems, IMBH = intermediate-mass black holes, AGN = active galactic nuclei, MW = the Milky Way galaxy, BBN = Big Bang Nucleosynthesis. The lilac shading indicates approximately the regimes in which the PPN formalism is applicable; see text for further description. Figure from [33], based upon earlier work by [254].
6. CONCLUSIONS

$10^{-6} - 10^{-5}$. This makes sense, since this is the amplitude of potential fluctuations laid down during inflation (compare with the amplitude of the CMB anisotropies), and these large-scale modes have been frozen outside the horizon for most of the history of the universe.

For spherically symmetric systems as described above, test bodies orbiting the same central mass lie on a straight line with gradient equal to three. Hence the planets of our Solar System lie on a straight line, being considered as test particles orbiting the Sun\(^1\), and the S stars occupy an analogous line where the central mass is the black hole at the centre of the Milky Way.

In the shaded region of this plot – the area populated by the Solar System, binary neutron stars\(^2\) and main sequence stars – deviations from GR are tightly constrained by tests of the Parameterized Post-Newtonian formalism. PPN proceeds as an expansion about the GR metric in terms of the potential, testing whether the first post-Newtonian corrections match those predicted by GR. It is worth noting, then, that constraints on PPN parameters essentially correspond to error ellipses along the $x$-axis of this plot. However, there is no reason to believe that the gravitational potential is the most appropriate expansion parameter for comparing modified gravity theories to GR. Imagine that there are corrections to the gravitational action that scale with curvature (arguably the most natural quantity, since $R$ appears in the Einstein-Hilbert action). These correction terms could be undetectable in the PPN regime and yet become of order unity in the upper or lower regions of this plot. Work is in progress to develop an adapted PPN formalism that is able to rephrase current constraints in terms of other variables such as curvature, density, length, etc. \([33]\).

PPN is not applicable to the systems residing in the lower third and upper-right section of the plot (the strong-field regime and the cosmological regime respectively). In the strong-field regime the potential is no longer small\(^3\), so it cannot be used as an expansion parameter. The most promising avenue for testing gravity here is via direct imaging of nearby black holes (for example, using the Event Horizon Telescope \([176]\)) or from gravitational waves \([86, 315]\), which we

---

\(^1\)In fact the non-negligible mass of Jupiter is taken into account, i.e. it is not treated as a massless test particle. This moves it off the straight line inhabited by the other planets, but the shift is not discernible on the scale of this plot.

\(^2\)Note that binary neutron stars occupy a different location to individual pulsars: measurements of their orbital spin-down rate probe the region of spacetime surrounding the binary, not the strong-field regime near the surface of a neutron star.

\(^3\)If one could find a system occupying the upper left-hand region of the diagram then PPN should in principle be applicable there, since it makes no statements about spacetime curvature. This is why I have extended the shaded region to the top of the plot. However, it is difficult to envision a system that curves spacetime very strongly whilst maintaining a weak characteristic potential.
expect to detect in the next decade from either the Advanced LIGO facility \cite{205} or pulsar timing arrays \cite{154}.

PPN is likewise not applicable to the cosmological regimes. Firstly, PPN requires expansion about a static metric, which is clearly not appropriate for an expanding universe. Secondly, scales comparable to the cosmological horizon are fully relativistic, so the weak-field requirement of PPN is not satisfied\footnote{One can see this by attempting to associate a gravitational potential to these lengthscales. Taking the radius $R \sim 1/H$ and using the Friedmann equation $3H^2 = 8\pi G_N \rho$, one obtains $\Phi \sim \frac{4G_N \pi R^3 \rho}{3R} \sim O(1)$.}. The most important result of this thesis has been the development of a new parameterized formalism, the Parameterized Post-Friedmann (PPF) framework, specifically designed for application to cosmological systems. Although PPF and PPN are different in construction, they are similar in philosophy and motivation. The aim of both is, as far as possible, to exhaust modifications to the perturbed metric (PPN) or gravitational perturbed field equations (PPF) that could arise in an alternative gravity theory, and constrain these generalized modifications with data. This has two major benefits: a) the need for repeated analyses of a single data set in the context of different theories is removed, and b) the (non-)viability of modification terms can be used to direct future theoretical work towards the most fruitful regions of ‘theory space’.

The work presented here has shown that to make optimum use of these benefits, it is crucial to use a parameterization that maintains an exact, analytical mapping between theories and the parameters/functions of the formalism (compare to the phenomenological parameterizations discussed in chapter 3). Another key outcome of this thesis has been to demonstrate that an exact, unified description of modified gravity theories is not simply a theoretical nicety. It is possible to calculate observationally-relevant quantities in a generalized manner, making parameterized frameworks a viable tool for future cosmological surveys.

However, this endeavour – truly agnostic tests of gravity on all scales – remains a work in progress. Indeed, it may be felt that there is something of a disconnect between the observational methods outlined in §1.2 and the extent to which these were discussed in later chapters (or lack thereof) in the context of modified gravity. The fledgling field of parameterized approaches to gravity (in cosmological contexts) has only recently achieved theoretical consistency, and is yet to achieve maximum efficiency. This was made evident in §5.7 of this thesis, where we saw that the constraints obtained from the growth rate of large-scale structure on parameterized formalisms are significantly weaker than the constraints on specific theories. It seems that we have not yet found the optimal data compression
6. CONCLUSIONS

strategies for extracting time-dependent constraints on gravity\(^1\).

Fig. 6.1 demonstrates the extrapolation of scales implied if one adopts the viewpoint that GR is too successful to accommodate new gravitational physics. Such extrapolation flies in the face of our established scientific principles. The conclusion one draws from this is that it is too early to 'give up' on modified gravity and dark energy models and accept the cosmological constant. Of course theorists will continue to generate new gravity models for the foreseeable future but, as I have argued in this thesis, this should not be regarded as the chief hallmark of progress. Instead we should look forward to drawing together new theoretical tools for understanding gravity and the availability of high-precision data from the next generation of experiments.

\(^1\)This should be compared with PPN, where the same constraining power is achieved irrespective of whether one tests theory-specific parameters directly, or first constrains the PPN framework and then maps the information onto the theory-specific parameters. This is not too surprising; the time- and scale-dependent nature of cosmological observables means that PPF is challenged with constraining unknown functions of two variables, whereas PPN tests only constant parameters.
References


REFERENCES


REFERENCES


REFERENCES


[86] N. Cornish, L. Sampson, N. Yunes, and F. Pretorius. Gravitational wave tests of general relativity with the parameterized post-Einsteinian frame-


REFERENCES


REFERENCES


REFERENCES


[115] R. Durrer. The Cosmic Microwave Background. September 2008. 18
REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


[286] Yong-Seon Song, Lukas Hollenstein, Gabriela Caldera-Cabral, and Kazuya


[294] M. Tegmark et al. Cosmological constraints from the SDSS luminous red


[315] N. Yunes and F. Pretorius. Fundamental theoretical bias in gravitational wave astrophysics and the parametrized post-Einsteinian frame-


