Disformal Gravity

A thesis presented for the degree of

Doctor of Philosophy

by

Johannes Joachimov Noller

Department of Physics

Imperial College London

July 2012

Submitted in part fulfillment of the requirements for the degree of Doctor of Philosophy in Physics of Imperial College London and the Diploma of Imperial College London.

To my parents,
to whom I owe everything

Abstract

An intriguing feature of scalar-tensor theories is the emergence of different metrics, e.g. when matter is minimally coupled to a metric non-trivially related to the Einstein metric $g_{\mu\nu}$ used to construct the Ricci scalar. Strong equivalence principle constraints then typically force permissible "many-metric" scenarios to reduce to a bimetric picture. In this thesis we first aim to construct the most general bimetric relation, where the two metrics are related by a single scalar degree of freedom ϕ and its derivatives. This results in the disformal metric relation and a natural extension which we present.

In the context of primordial structure formation, disformal bimetric theories give rise to "general single field inflation" models of the $P(X, \phi)$ type. We investigate the perturbative properties of such disformally motivated models. The focus is on non-Gaussian phenomenology and we establish non-Gaussian fingerprints for inflationary single field models and non-inflationary bimetric setups, also going beyond the slow-roll approximation. Furthermore we show that various dualities exist between disformally motivated $P(X, \phi)$ theories and higher-form models. As an explicit example we use the dual picture to compute non-Gaussian signals for three-form theories.

In the context of dark energy/modified gravity, we show that the conformal subgroup of the general disformal relation can be used to construct a generalized "derivative" Chameleon setup. We present and investigate this setup and study its phenomenology. Finally we show that a natural extension of the disformal relation can generate Galileon solutions from a single geometrical invariant - the first Lovelock term - in four dimensions.

As such the over-arching theme of this thesis is to show that the disformal bimetric picture and its extensions present us with a geometrical understanding of scalar-tensor/single field models. That they provide a unified description of large classes of scenarios linked to accelerated space-time expansion and also point us towards new physically motivated setups.

Declaration

I hereby certify that to the best of my knowledge all the material presented throughout this thesis is the result of my own work or has otherwise been properly acknowledged.

Johannes Joachimov Noller

The research presented in this thesis is based on the following work

1. J. Noller and J. Magueijo, "Non-Gaussianity in single field models without slow-roll," Phys. Rev. D **83**, 103511 (2011) [arXiv:1102.0275]. [1]
2. J. Noller, "Constraining fast-roll inflation," Proceedings for Rencontres de Moriond (2012) [arXiv:1205.5796]. [2]
3. J. Magueijo, J. Noller and F. Piazza, "Bimetric structure formation: non-Gaussian predictions," Phys. Rev. D **82**, 043521 (2010) [arXiv:1006.3216]. [3]
4. J. Noller, "Derivative Chameleons," JCAP **1207**, 013 (2012) [arXiv:1203.6639]. [4]
5. D. J. Mulryne, J. Noller and N. J. Nunes, "Three-form inflation and non-Gaussianity," [arXiv:1209.2156]. [5]
6. J. Noller, "Emergent Galileons," in preparation. [6]

Related papers, to which the author contributed, are

7. J. Noller and J. Magueijo, "Non-adiabatic primordial fluctuations," Class. Quant. Grav. **28**, 105008 (2011) [arXiv:0911.1907]. [7]
8. J. Magueijo and J. Noller, "Primordial fluctuations without scalar fields," Phys. Rev. D **81**, 043509 (2010) [arXiv:0907.1772]. [8]

Acknowledgements

None of the work in this thesis would have been possible without the guidance, support and encouragement of my supervisor, João Magueijo. I am extremely grateful for all his help, collaboration, teaching and mentoring throughout my PhD.

I would like to thank everyone I have shared office 512b with: Ben Hoare, Noppadol Mekareeya, Silvia Nagy, Tom Pugh, Andrej Stepanchuk, Giuseppe Torri and James Yearsley. You have been invaluable both as friends and as colleagues during our many, many discussions, where I have learned so much and which have helped my research enormously.

I am also very thankful to Carlo Contaldi, Pedro Ferreira, Andrew Jaffe, Anupam Mazumdar, Ali Mozaffari, David Mulryne, Nelson Nunes, Federico Piazza, Dan Thomas and many other people from the cosmology community for very useful discussions, collaboration towards, comments on and ideas about the work contained in this thesis.

My teachers throughout the years, especially Ian Craig, Jonathon Hodby, Andre Lukas, Marion Muetzelfeldt, Christopher Palmer, Armin Reichold, Subir Sarkar and David Wallace, deserve special mention. Without your inspiration I would never have even started this PhD.

I am most deeply indebted to my family and especially my parents for their limitless support. Words can't convey how thankful I am for everything you have done for me.

I would like to express my deep thanks to the people that have ensured my PhD time has been incredibly fun, who have kept me grounded, helped me to grow up and supported me in countless ways. Adam, Andre, Andreas, Amy, Anne, Ariel, Assaf, Caroline, Charlie, Daniel, Dermot, Eduardo, Eleni, Irmak, Jan, Janine, Jonathan, Kieron, Kostas, Laura, Leyla, Marc, Marta, Monifa, Naiara, Richard, Rumyana, Sarah, Séverine, Siân, Sönke, Stefano, Zisis and many others, the last four years would not have been the same without you.

Finally this thesis, just as all the papers written before, is also in memory of Andrew Mason - an amazing friend, who is no longer with us, but who will never be forgotten.

Contents

List of Figures

Part I

Introduction

Chapter 1

Introduction

Present-day cosmologists are blessed with vast amounts of rapidly growing and improving observational data. In interpreting this data two fundamental questions have become prominent, both relating to the accelerated expansion of space-time. Firstly, to fully understand the origin of small inhomogeneities in the early universe which act as seeds for the formation of cosmic structure, from clusters all the way down to planets. Here the leading paradigm is an "inflationary" period of accelerated expansion [10, 11, 12], stretching quantum fluctuations to become seeds for classical gravitational structure formation [13]. Secondly, to identify the agent responsible for causing accelerated expansion of space-time today - "dark energy" or indeed a modified theory of gravity (for a review see [14]).

The simplest dynamical theories for both inflation and dark energy/modified gravity introduce a single scalar *dof* (degree of freedom) ϕ in addition to the metric *dof* encoded by $g_{\mu\nu}$. The effective field theories generated in this way may, in a rather wide sense, be described as scalar-tensor theories where the scalar drives accelerated expansion of space-time. Such theories can also arise naturally in a variety of ways even if no new fundamental scalar field is introduced, e.g. as a result of the appearance of induced scalar(s) in compactifying higher-dimensional theories or via the introduction of a Stückelberg scalar from introducing a massive graviton.

An interesting feature of scalar-tensor theories is the emergence of different effective metrics. On the dark energy/modified gravity side, models such as $F(R)$/chameleon theories [14, 15, 16] can be constructed by minimally coupling matter to a metric $\tilde{g}_{\mu\nu}$ conformally related to $g_{\mu\nu}$ via $\tilde{g}_{\mu\nu} = A^2(\phi)g_{\mu\nu}$. Alternative setups such as DGP [17] and DBI Galileons [18] arise as limits of probe brane configurations, where matter on our local 4D brane is minimally coupled to an induced metric $\tilde{g}_{\mu\nu} = g_{\mu\nu} + \partial_\mu\phi\partial_\nu\phi$. On the inflationary side, fluctuations in generic $P(X,\phi)$ models [19, 20] propagate along the geodesics of an effective metric different from $g_{\mu\nu}$, so that bimetric structure emerges perturbatively [21]. On the other hand, fully non-perturbative DBI-type models [22] can be straightforwardly generated by postulating the existence of two

fundamental metrics [23]. Throughout this thesis we will consider these different models and others, making their bimetric structure explicit. We will argue that the bimetric perspective is a particularly intuitive and useful way of understanding and constructing such theories.

Disformal gravity then is an umbrella term for an approach that aims to systematically construct and investigate the most general relations between different emergent metrics in scalar-tensor theories of the type described. We will review the disformal relation, first proposed in [24], and argue that it can be used to provide a unified description of large classes of effective field theories both in the inflationary and dark energy/modified gravity context. That it points us towards new types of physically motivated models and also leads to more general insights into generic single field inflation and dark energy/modified gravity models. As such the overarching aim throughout this thesis will be to use the bimetric perspective to gain a geometric understanding of single field/scalar-tensor theories. We hope that, after perusing the thesis, we have convinced the reader that the disformal approach is a powerful one, both in classifying and understanding known theories as well as in constructing new scenarios and opening up new vistas.

This thesis is organized as follows. The remainder of the introduction consists of three prologues, mainly reviewing the necessary background for what will follow. Prologue I introduces the disformal relation from first principles and discusses possible generalizations. Prologue II reviews basic single field inflationary models, their generalizations and some of the associated phenomenology. Here we also discuss how bimetric structures manifest themselves in such setups and how they are related to disformal transformations. Prologue III then recaps scalar-tensor theories in the context of dark energy/modified gravity. A particular focus are screening mechanisms [25], i.e. the idea that a cosmological scalar, which drives accelerated space-time expansion on cosmic scales, needs to be screened away on small scales in order to satisfy e.g. solar system constraints on the presence of fifth forces [26]. Once again we point out how this relates to disformal setups.

In chapter 2 we investigate the perturbative properties of inflationary effective field theories with a single *dof* , as arising in a disformal context. We derive their non-Gaussian signals and show how this computation can be extended to cases when slow-roll conditions do not hold. As we will see, "fast-roll" effects significantly alter non-Gaussian phenomenology, leading to a reduced $f_{\rm NL}^{equi}$, modified shape-spaces and less stringent parameter-space constraints when confronting the model with observational data. This allows us to place bounds on slow-roll parameters both from three-point statistics as well as from tensor modes.

In chapter 3 we extend the results of the previous chapter, making use of the fact that the disformal bimetric relation also allows us to construct non-inflationary structure formation models with "superluminally" propagating perturbations without violating causality constraints. We show how such models can be constructed consistently and investigate their non-Gaussian

properties.

In chapter 4 we explore the existence of a large class of dualities relating various inflationary n-form models with single field setups. We focus on a particular set of three-form models [27, 28] and explicitly construct their power spectra and non-Gaussian features. Interestingly we find that the simplest set of three-form models with a power-law potential, DBI-type setups and minimal disformal models all give rise to closely related observational signatures.

In chapter 5 we investigate whether the disformal relation allows new ways of implementing chameleon-screening in a dark energy/modified gravity context. Chameleon models [15, 16] and their features are reviewed and discussed from the bimetric disformal perspective. We establish a no-go theorem for purely disformal chameleon theories but also find that the conformal subgroup of the disformal relation generates a new type of generalized chameleon model. We investigate this model in detail and point out how it alters standard chameleon phenomenology.

In chapter 6 we discover the relationship between Vainshtein-screened solutions of the Galileon type [29] and disformal bimetric setups. We show that an extension of disformal metric relations, dubbed triformal relations (and introduced in prologue I), naturally generates all Galileon terms from a single geometric invariant - the first Lovelock term [30, 31]. We investigate the role the disformal subgroup of triformal transformations can play in this process and conjecture how other ghost-free scalar-tensor theories are related to the triformal setup.

Finally, in chapter 7, we summarize our findings and discuss possible research avenues for the future. Various appendices follow which we have reserved for further explanations and technical details that are relevant to the computations carried out throughout this thesis.

1.1 Prologue I : The relation between gravitational and matter geometries

1.1.1 The bimetric perspective

In general relativity nature is endowed with a single global Riemannian geometry. Its metric $g_{\mu\nu}$ is used to construct both the Einstein-Hilbert action as well as to (minimally) couple matter fields to gravity. However, in general the situation may not be this simple and nature may allow the existence of more than one geometry. In fact, even if at a fundamental level there is only a single metric $g_{\mu\nu}$, new effective metrics can arise when additional *dof* are introduced into the theory. These new effective metrics will then be functions of the fundamental metric and of the additional *dof* present in the theory. A promising starting point would then be an action of the form

$$\mathcal{S} = \int d^4x\sqrt{g}\frac{M^2}{2}R\left[g_{\mu\nu}\right] + \mathcal{S}_m\left(\tilde{g}^{(i)}_{\mu\nu}, \Psi_{(i)}\right). \tag{1.1}$$

Here we have built the Einstein-Hilbert term R out of an Einstein metric $g_{\mu\nu}$, but minimally coupled matter fields $\Psi_{(i)}$ to in principle distinct metrics $\tilde{g}^{(i)}_{\mu\nu}$. This allows for gravity to discriminate between different matter species $\Psi_{(i)}$ putting us in immediate tension with the equivalence principle. To show why let us state the equivalence principle in its weakest form [14].

The weak equivalence principle (WEP): All uncharged, freely falling test particles follow the same trajectories, once an initial position and velocity have been prescribed. In other words, locally gravity does not distinguish between different types of test particles.

An action such as (1.1), which does allow locally distinct metrics to couple to different test particles, violates this principle unless all $\tilde{g}^{(i)}_{\mu\nu}$ are identical. Given tight observational constraints on equivalence principle violations [26], we will opt to avoid any such violation here and therefore reduce (1.1) to a bimetric theory

$$\mathcal{S} = \int d^4x\sqrt{g}\frac{M^2}{2}R + \mathcal{S}_m\left(\tilde{g}_{\mu\nu}, \Psi_i\right), \tag{1.2}$$

where now we only have two metrics $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$.

How does this relate to the scalar-tensor setup we were motivating above? If gravity is modified by the presence of a single new scalar *dof* ϕ, then ϕ can alter the way matter couples to gravity, giving rise to a new distinct metric $\tilde{g}_{\mu\nu}$ for matter. Before investigating how $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$ are related (and in particular how ϕ enters this relation), let us first lay out the problem in general terms without imposing any particular metric relation. Consequently we start with a schematic action of the following form

$$\mathcal{S} = \int d^4x\sqrt{g}\frac{M^2}{2}R + \mathcal{S}_m\left(\tilde{g}_{\mu\nu}, \Psi_i\right) + \mathcal{S}_\phi \tag{1.3}$$

where Ψ_i are matter fields minimally coupled to $\tilde{g}_{\mu\nu}$ as before. We have generalized (1.2) by adding $\mathcal{S}_\phi$, which denotes an action giving the scalar field ϕ dynamics of its own. We emphasize that there is no a priori requirement constraining S_ϕ to be formed with either $g_{\mu\nu}$ or $\tilde{g}_{\mu\nu}$.[1] We will also find it useful to schematically write (1.3) as

$$\mathcal{S} = \int d^4x\sqrt{g}\mathcal{L}_E\left(g_{\mu\nu}, \phi\right) + \int d^4x\sqrt{\tilde{g}}\mathcal{L}_m\left(\tilde{g}_{\mu\nu}, \Psi_i\right), \tag{1.4}$$

where we have split the action into two parts which minimally couple to $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$ respectively. As a consequence $\mathcal{S}_\phi$ has been absorbed into $\mathcal{L}_E$ and/or $\mathcal{L}_m$ depending on which metric(s) are used to construct $\mathcal{S}_\phi$.

[1] Note that adding a non-trivial $\mathcal{S}_\phi$ can be seen as modifying the gravity sector itself if ϕ is seen as a gravitational *dof* .

1.1.2 Metric relations

We now turn our interest to the relation between $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$. Explicit examples of theories that have such non-trivial relations depending on ϕ are readily found and we will discuss and extend these throughout this thesis. However, here we are wondering what the most general relation between two such metrics might be. A systematic way of constructing a general matter metric in the presence of a gravitational scalar ϕ is to write

$$\tilde{g}_{\mu\nu} = \tilde{g}_{\mu\nu}(g_{\mu\nu}, \phi, \partial\phi, \partial\partial\phi, ...), \tag{1.5}$$

which is therefore a function of the gravity metric $g_{\mu\nu}$, the field ϕ and its derivatives. Eventually we will want to truncate this derivative expansion at some order. However, for the time being let us focus on what possible valence $(0, 2)$ tensors we can form out of $g_{\mu\nu}, \phi$ and its derivatives. Schematically we can write

$$\tilde{g}_{\mu\nu} = \mathcal{F}_1 g_{\mu\nu} + \mathcal{F}_2 \partial_\mu\phi\partial_\nu\phi + \mathcal{F}_3 \partial_\mu\partial_\nu\phi, \tag{1.6}$$

where $\mathcal{F}_i$ are some scalar functions constructed out of coordinate invariants, which we can group by derivative order $\phi, X, \Box\phi, ...$ (insisting on only using coordinate invariants ensures the appearance of X but not $\partial_\mu\phi$ by itself). Here $X = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi$ is the usual canonical kinetic term for ϕ. Note that the signature of $g_{\mu\nu}$ used here is $(-+++)$, and hence $2X = \dot{\phi}^2 - \left(\vec{\partial}\phi\right)^2$ in Minkowski space. We therefore have

$$\tilde{g}_{\mu\nu} = A^2(\phi, X, \Box\phi, \ldots)g_{\mu\nu} + B^2(\phi, X, \Box\phi, \ldots)\partial_\mu\phi\partial_\nu\phi + C^2(\phi, X, \Box\phi, \ldots)\partial_\mu\partial_\nu\phi, \tag{1.7}$$

We dub this a "triformal" metric relation. This generalized metric relation consists of a conformal piece, with $A^2(\phi, X, \Box\phi, \ldots)$ being the conformal factor, a so-called "disformal" piece, where we shall call $B^2(\phi, X, \Box\phi, \ldots)$ the disformal factor, and a triformal piece with a triformal factor $C^2(\phi, X, \Box\phi, \ldots)$.

If (1.7) arises in an effective field theory then subsequently higher derivatives of the field can be suppressed by higher powers of the cutoff scale. This motivates a truncation of (1.7) at some derivative order. Note, however, that such an argument naturally does not discriminate between e.g. $(\partial\phi)^n$ and $\partial^n\phi$ or between $\partial_\mu\phi\partial_\nu\phi$ and $\partial_\mu\partial_\nu\phi$. However, operators such as $\partial^n\phi$ with $n \geq 3$ are dangerous since they lead to the appearance of equations of motion with higher than second order derivatives of the field, where we would expect ghost-like instabilities via Ostrogradski's theorem [32] (also see appendix A). If $n = 2$, a finite number of terms can be constructed [29, 33, 34] for which dangerous higher order terms in the equation of motion cancel out. We will return to this later in chapter 6, but for now we only permit $n = 0, 1$ trivially

protecting us from having Ostrogradski-type ghosts. In particular this also means we will ignore the triformal piece $\partial_\mu\partial_\nu\phi$ and other higher order terms such as $\Box\phi$. In other words, we are reducing (1.7) by requiring $C = 0$ and A, B to be functions of ϕ and X only. Consequently the triformal relation reduces to a disformal one satisfying

$$\tilde{g}_{\mu\nu} = A^2(\phi, X)g_{\mu\nu} + B^2(\phi, X)\partial_\mu\phi\partial_\nu\phi. \tag{1.8}$$

This relation was first proposed by Bekenstein [24] to be the most general relation between gravitational and matter metrics when a single *dof* ϕ and its first derivative are used to construct a matter metric $\tilde{g}_{\mu\nu}$. Note that the argument he used to derive (1.8) differs from the one presented here[2] - for details see appendix B and [24].

In the remaining two prologues we will introduce some of the relevant background for exploring scalar-tensor theories and their effective bimetric descriptions in inflationary and dark energy/modified gravity settings. In particular this will show how a disformal description arises abundantly for wide classes of such theories. The associated bimetric description equips us with a geometric understanding for these setups, which also makes their causal structure explicit (e.g. particles travel on geodesics of whichever metric they are minimally coupled to). We will also find that there are strong hints that a bimetric formulation serves as an excellent guide to construct theories whose interaction terms are protected from renormalizations [35, 36] - i.e. they permit a strong coupling regime while maintaining quantum corrections under control. Perhaps most importantly we will find that the structure of conformal and disformal factors $A^2(\phi, X)$ and $B^2(\phi, X)$ is intimately tied to the phenomenology exhibited by the associated scalar-tensor theory. For example non-trivial such factors straightforwardly give rise to a varying speed of sound in inflation models and hence the generation of large non-Gaussian signals. On the dark energy/modified gravity side, on the other hand, we will find that the conformal and disformal terms are associated with the appearance of chameleon and Vainshtein screening respectively. The remainder of the thesis is dedicated to showing how the disformal relation provides an extremely powerful unified description of known single field inflationary and dark energy/modified gravity models. In addition we will find it to be an excellent means to explore these models further *and* to construct alternative and/or more general single field setups capable of causing cosmic acceleration in agreement with observational constraints.

[2] As a brief summary: Bekenstein assumes what is essentially a 4D effective field theory in an overall Finslerian geometry with gravity and matter metrics related by a single degree of freedom ϕ up to first order in its derivatives. He then proceeds to show that this has to reduce to a Riemannian geometry described by (1.8).

1.2 Prologue II : Inflation and primordial physics

1.2.1 The FLRW universe and Big Bang problems

The cosmological principle requires the universe to be homogeneous and isotropic about every point. In other words, there are no preferred locations or directions in the universe. While non-linear evolution and structure formation dynamically breaks these symmetries at small scales, there is strong evidence for this principle on large scales. Namely the cosmic microwave background (CMB) is highly homogeneous up to one part in 10^5 [9].

We can formalize the cosmological principle by constructing a cosmological metric subject to the requirement of overall homogeneity and isotropy. The result is the FLRW (Friedmann-Lemaitre-Robertson-Walker) metric, which in spherical polar co-ordinates (t, r, θ, ϕ) takes on the following form:

$$ds^2 = -dt^2 + a^2(t)\left(\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi)\right). \tag{1.9}$$

$a(t)$ is the so-called scale factor of the metric, t is proper time and altering the parameter $k = -1, 0$ or 1 describes open, flat or closed universes respectively. Note that $a(t_i)$ essentially measures the size of space-like hypersurfaces at a given time t_i. The Hubble parameter is then defined as $H = \dot{a}/a$ and serves as a convenient measure of the expansion of the universe, since $\dot{a} > 0$ in the expanding case. A dot $\dot{}$ denotes differentiation with respect to proper time t. It will prove useful to also introduce conformal time η which satisfies

$$\eta = \int \frac{dt}{a}, \tag{1.10}$$

defining co-ordinates that are "comoving" with the expansion of the universe. We will denote differentiation with respect to conformal time η by a dash $'$. The conformal Hubble parameter $\mathcal{H}$ then satisfies $\mathcal{H} = aH$.

Now we can expand out the Einstein tensor $G_{\mu\nu}$ and write Einstein equations as

$$G_{\mu\nu} = R_{\mu\nu} - (R/2 - \Lambda)g_{\mu\nu} = 8\pi G T_{\mu\nu}, \tag{1.11}$$

$R_{\mu\nu}$ denotes the Ricci tensor, $R = R^{\mu}_{\mu}$, G is Newton's gravitational constant and Λ is a potential cosmological constant, which is however sub-dominant in the primordial universe. $T_{\mu\nu}$ is the stress energy tensor which can be obtained by varying the matter Lagrangian with respect to the metric.

Modeling the stress-energy tensor by assuming the universe is filled with a perfect fluid one

can write

$$T_{\mu\nu} = (\rho + P)u_\mu u_\nu + P g_{\mu\nu}, \tag{1.12}$$

where the four-velocity u^μ satisfies $u^\mu u_\mu = -1$, P is pressure and ρ denotes the energy density. The Friedmann equations governing the evolution of the fluid then follow straightforwardly. The 00 equation and the trace of the Einstein tensor then respectively yield

$$\begin{aligned} H^2 &= \frac{8\pi G}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3}, \\ \dot{H} + H^2 &= \frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) + \frac{\Lambda}{3}. \end{aligned} \tag{1.13}$$

The observed expansion of the universe, e.g. as implied by the Hubble law, then allows us to retrodict the universe's past evolution arriving at a very hot and dense initial state at early times in the universe. The evolution from such an initial state with a subsequent expanding universe which dilutes and cools down is the so-called Hot Big Bang model. In this context nucleosynthesis and the existence of a cosmic microwave background stemming from the epoch when radiation and matter decoupled are further pieces of observational evidence in favor of this model. However, such a setup faces a number of important challenges.

The flatness problem: Rewriting the first Friedmann equation in terms of the parameter $\Omega = \rho/\rho_{crit}$, where $\rho_{crit} = 3H^2/(8\pi G)$ is the critical density for which the curvature of the universe is zero, one finds

$$\Omega(t) - 1 = \frac{k}{(aH)^2} = \frac{k}{\dot{a}^2}, \tag{1.14}$$

For an accelerating universe with $\ddot{a} > 0$ then Ω_m asymptotically approaches $\Omega = 1$ over time, whereas if $\ddot{a} < 0$ it diverges from this value. Current constraints find $\Omega = 1$ within roughly 2%, at a 2σ confidence [9]. Thus there's an immediate tension with a decelerating universe, where any primordial $\Omega \neq 1$ will rapidly diverge away from 1. More precisely, assuming the standard cosmological evolution where a radiation-dominated era is followed by matter-domination, $\Omega - 1 \propto a^2 \propto t$ during radiation-domination and $\Omega - 1 \propto a \propto t^{2/3}$ during matter/dust-domination. At the electroweak scale a fine-tuning of $|\Omega - 1| \sim 10^{-27}$ is therefore required. Note that the argument here has ignored dark energy, which, since it only becomes relevant in the late universe, will only mildly alleviate the fine-tuning discussed, however.

The horizon problem: The cosmic microwave background is observed to be highly isotropic with a smooth temperature to within one part in 10^5. Yet regions in the CMB separated by more than ~ 2 degrees could not have been in causal contact with each other at the time of the formation of the CMB, assuming the universe had been continuously decelerating in accordance

with the hot big bang model. More precisely one may define the particle or comoving horizon

$$\eta = \int_{a_i}^{a} \frac{da}{a} \frac{1}{aH(a)}, \tag{1.15}$$

which gives the maximum separation between two points that could have been in causal contact in the past, where a_i is the initial value of the scale factor. Evaluating this at photon decoupling shows that regions in the CMB separated by more than ~ 2 degrees could not have been in causal contact, as mentioned above. The fact that they still appear highly thermalized, in the sense that the temperature across the whole CMB is highly uniform, without an apparent causal mechanism to explain this uniformity is known as the horizon problem.

The monopole problem: Typically, as a consequence of spontaneous symmetry breaking, high energy physics predicts that topological defects will generate magnetic monopoles, cosmic strings, domain walls and textures. The fact that none of these objects have been observed demands an explanation. In fact, historically this problem was one of the main motivations for inflation to be proposed.

1.2.2 Single field inflation

Inflation is a period of accelerated space-time expansion in the early universe

$$\frac{d}{dt}\left(\frac{1}{aH}\right) < 0 \quad \Rightarrow \quad \ddot{a} > 0 \quad \Rightarrow \quad \rho + 3P < 0\,. \tag{1.16}$$

A prolonged period of inflation can then resolve the big bang problems outlined above. We have already seen that accelerated expansion drives Ω close to 1, so that $\Omega \sim 1$ as an "initial value" for the evolution post-inflation is indeed a generic consequence of inflation models. Objects such as magnetic monopoles will be diluted away by a phase of rapid expansion, alleviating concerns about the non-detection of such objects. And finally the horizon problem is resolved, since all scales which enter the universe in our observable window, e.g. at CMB scales, would have been within the comoving horizon (and hence in causal contact) pre-inflation. Different regions in the CMB would therefore have had enough time to thermalize and reach an equilibrium energy density and temperature.

In order to realize inflation, the universe has to be dominated by matter exhibiting negative isotropic pressure. Standard model matter (excluding the Higgs) or radiation does not have this property. At least one extra degree of freedom is therefore typically introduced in order to source accelerated space-time expansion. In the simplest case we introduce a single scalar field ϕ and

consider have an action of the form [19, 20]

$$S = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{R}{2} + P(X,\phi)\right] \ , \tag{1.17}$$

where the pressure P is a general function of a single scalar field ϕ and its kinetic term $X = -\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$. For a canonical field $P(X,\phi) = X - V(\phi)$ and the energy density is given by $\rho = X + V(\phi)$, from which we can see that canonical inflation corresponds to a period where the potential $V(\phi)$ dominates over the canonical kinetic term X. In principle P can contain higher derivatives of the field ϕ as well, however such terms are expected to be suppressed by the UV cut-off scale in an effective field theory. This motivates studying an action of the form (1.17), which is the most general Lorentz invariant action for a single scalar field minimally coupled to gravity that contains at most first derivatives of the field. The pressure p and energy density ρ for these models are given by

$$\begin{aligned} p &= P(X,\phi)\,, \\ \rho &= 2XP_{,X} - P(X,\phi)\ , \end{aligned} \tag{1.18}$$

where $P_{,X} \equiv \partial P/\partial X$.

We will find it useful to characterize the non-canonical nature of (1.17) with the following quantities [37]

$$\begin{aligned} c_s^2 &= \frac{P_{,X}}{\rho_{,X}} = \frac{P_{,X}}{P_{,X} + 2XP_{,XX}}\,, \\ \Sigma &= XP_{,X} + 2X^2P_{,XX} = \frac{H^2\epsilon}{c_s^2}\ , \\ \lambda &= X^2P_{,XX} + \frac{2}{3}X^3P_{,XXX}\ , \end{aligned} \tag{1.19}$$

where c_s is the so-called adiabatic speed of sound (i.e. the speed with which perturbations propagate). An infinite hierarchy of slow-roll parameters ϵ and ϵ_s and their derivatives can then be constructed for a theory (1.17) with an FRW metric with scale factor $a(t)$ and corresponding Hubble rate $H(t) = \dot{a}/a$. These hierarchies are

$$\epsilon \equiv -\frac{\dot{H}}{H^2},\ \eta \equiv \frac{\dot{\epsilon}}{\epsilon H}, ... \qquad ; \qquad \epsilon_s \equiv \frac{\dot{c}_s}{c_s H},\ \eta_s \equiv \frac{\dot{\epsilon_s}}{\epsilon_s H}, ... \tag{1.20}$$

The slow-roll approximation requires that $\epsilon, \eta ... \epsilon_s, \eta_s ... \ll 1$ and accelerated expansion takes place when $\epsilon < 1$. Requiring a prolonged period of inflation therefore constrains the values taken by slow-roll parameters. Note that these definitions (1.20) are more general than alternative definitions in terms of the scalar potential and its derivatives, which assume canonical kinetic

terms. Perhaps more fittingly (1.20) have therefore also been dubbed "slow-variation parameters" [38, 39, 40]. DBI inflation [22], for example, is slow-varying in this sense yet allows steep potentials and hence the inflaton is not slow-rolling. However, in order to agree with convention and avoid confusion, we will refer to (1.20) as slow-roll parameters from now on. Note that truncating slow-roll parameters at some derivative order, setting the remainder to zero, allows us to approximate H and c_s with Taylor expansions.

1.2.3 The classical curvature perturbation

Here we consider scalar perturbations about an FLRW metric. In order to seed structure formation we require the presence of some initial fluctuations, which we will consider classically in this section. The quantization and evaluation of quantum correlators between such fluctuations is left to chapter 2. We now perturb the FLRW metric to first order and find [13]

$$ds^2 = -a^2\left(1+2\Phi\right)d\eta^2 + 2a^2 B_{,i} d\eta dx^i + a^2\left[\left(1-2\Psi\right)\delta_{ij} + 2E_{,ij}\right]dx^i dx^j\,, \tag{1.21}$$

where we have assumed a flat background metric δ_{ij}, Ψ is the curvature perturbation and Φ denotes the lapse function, B_1 and E_1 are the scalar part of the shear.

Under a gauge transformation $x^\mu \to x^{\mu'} = x^\mu + \xi^\mu = x^\mu + (\xi^0, \xi_{,i})$ metric perturbations $\delta g_{\mu\nu}$ transform as

$$\delta\tilde{g}_{\mu\nu}(x^{\mu'}) = \delta g_{\mu\nu}(x^\mu) - g_{\mu\rho}\xi^\rho_{,\nu} - g_{\rho\nu}\xi^\rho_{,\mu}. \tag{1.22}$$

Different gauges correspond to different ways of foliating our space-time. Since a background-perturbation split is not a covariant operation, gauge degrees of freedom are introduced in this way. The resulting co-ordinate ambiguity can be resolved by fixing a gauge. For this section we will employ the uniform curvature gauge in which spatial hypersurfaces are flat (otherwise known as the flat gauge) and we have $\Psi = 0$ and $E = 0$. However, in general we may transform between different gauges by using the following relations (to linear order).

$$\begin{aligned}
\tilde{\Phi} &= \Phi + \mathcal{H}\xi_0 + \xi' \\
\tilde{B} &= B - \xi_0 + \xi', \\
\tilde{\Psi} &= \Psi - \mathcal{H}\xi_0, \\
\tilde{E} &= E + \xi.
\end{aligned} \tag{1.23}$$

Fixing a gauge, we may use the gauge transformations to eliminate two of the four functions Ψ, Φ, B, E. In the flat gauge, as mentioned above, these are chosen to be $\Psi = 0$ and $E = 0$.

Note that a scalar field quantity such as inflaton perturbations will transform as

$$\delta\tilde{\phi}_1 = \delta\phi_1 + \phi_0'\xi^0. \tag{1.24}$$

We can now define two gauge-invariant curvature perturbations, which coincide on super-horizon scales. Firstly the curvature perturbation on uniform density hypersurfaces ζ [41] is given by

$$-\zeta \equiv \Psi + \frac{\mathcal{H}}{\rho_0'}\delta\rho\,, \tag{1.25}$$

which, when evaluated in flat gauge, simplifies to $-\zeta = \frac{\mathcal{H}}{\rho_0'}\delta\rho$. Secondly, the comoving curvature perturbation R [42], i.e. the curvature perturbation on comoving or uniform field slices, is

$$\mathcal{R} \equiv \Psi + \frac{\mathcal{H}}{\phi_0'}\delta\phi\,, \tag{1.26}$$

where we can once again simplify this in the flat gauge to obtain $\mathcal{R} = \frac{\mathcal{H}}{\phi_0'}\delta\phi$. These variables will prove useful when evaluating inflationary observables. In particular we will be interested in computing n-point correlation functions such as $\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\rangle$ and $\langle\zeta(t,\mathbf{k}_1)\zeta(t,\mathbf{k}_2)\zeta(t,\mathbf{k}_3)\rangle$, here shown in Fourier space. Observationally these correspond to correlating temperature fluctuations on different scales k, e.g. ones measurable via the CMB. For further details on related perturbation theory see [43, 44].

1.2.4 Perturbations and emergent geometry

We begin by considering the general $P(X,\phi)$ theory introduced above (1.17) with action

$$S = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{R}{2} + P(X,\phi)\right]\,. \tag{1.27}$$

Following [21] we work out the equation of motion for the scalar field

$$-\frac{\delta S}{\delta\phi} = \tilde{G}^{\mu\nu}\partial_\mu\partial_\nu\phi + 2XP_{,X\phi} - P_{,\phi} = 0, \tag{1.28}$$

where the effective metric $\tilde{G}^{\mu\nu}$ is given by

$$\tilde{G}^{\mu\nu} \equiv P_{,X}g^{\mu\nu} + P_{,XX}\partial^\mu\phi\partial^\nu\phi. \tag{1.29}$$

We now consider perturbations $\delta\phi$ around a given background ϕ_0. By Leray's theorem these propagate causally as determined by the metric $\tilde{G}^{\mu\nu}$. As shown in [21] we may write the equations of motion for the perturbations in terms of a metric conformally related to $\tilde{G}^{\mu\nu}$ (and which

consequently preserves the causal structure of $\tilde{G}^{\mu\nu}$). This metric $G^{\mu\nu}$ is defined as

$$G^{\mu\nu} \equiv \frac{c_s}{P_{,X}^2}\tilde{G}^{\mu\nu}, \quad \sqrt{-G} \equiv \sqrt{-\text{det}G_{\mu\nu}^{-1}} \quad \text{where} \;\; G_{\mu\lambda}^{-1}G^{\lambda\nu} = \delta_\mu^\nu, \tag{1.30}$$

Working out the corresponding inverse metric one finds

$$G_{\mu\nu} = \frac{P_{,X}}{c_s}\left[g_{\mu\nu} - c_s^2\left(\frac{P_{,XX}}{P_{,X}}\right)\partial_\mu\phi\partial_\nu\phi\right], \tag{1.31}$$

where c_s is defined as in (1.19). This is evidently of the disformal form with $A^2(\phi, X) = \frac{P_{,X}}{c_s}$ and $B^2(\phi, X) = -c_s P_{,XX}$. The action for perturbations $\delta\phi$ can now be written

$$S_{\delta\phi} = \frac{1}{2}\int d^4x\sqrt{-G}\,\left[G^{\mu\nu}\partial_\mu\delta\phi\partial_\nu\delta\phi - M_{\text{eff}}^2\delta\phi^2\right], \tag{1.32}$$

where the effective mass term is

$$M_{\text{eff}}^2 \equiv \frac{c_s}{P_{,X}^2}\left(2XP_{,X\phi\phi} - P_{,\phi\phi} + \frac{\partial\tilde{G}^{\mu\nu}}{\partial\phi}\nabla_\mu\nabla_\nu\phi_0\right). \tag{1.33}$$

Defining a covariant derivative for the metric $G^{\mu\nu}$, i.e. $D_\alpha G^{\mu\nu} = 0$, the equation of motion for perturbations can be written as a massive Klein-Gordon equation

$$G^{\mu\nu}D_\mu D_\nu\delta\phi + M_{\text{eff}}^2\delta\phi = 0. \tag{1.34}$$

Perturbations therefore propagate on geodesics of $G^{\mu\nu}$. In this way an emergent metric disformally related to $g_{\mu\nu}$ becomes important when considering perturbations in scalar field models of primordial physics with non-canonical kinetic terms. Moreover the speed of sound c_s with which perturbations propagate on the background ϕ_0 is intimately tied to the conformal and disformal factors in the disformal relation (1.8). Specifically we have

$$A^2(\phi, X) = \frac{P_{,X}}{c_s} \qquad \text{and} \qquad B^2(\phi, X) = -c_s P_{,XX} \tag{1.35}$$

$$c_s^2 = 1 + 2X\frac{B^2}{A^2}. \tag{1.36}$$

In this way disformal structure manifests itself for non-canonical scalar field models and a bimetric formalism of the type outlined becomes appropriate when dealing with perturbations $\delta\phi$. Two final comments: A well-posed Cauchy problem now requires initial conditions to be posed on a hyper-surface which is space-like with respect to both metrics ($g^{\mu\nu}$ and $\tilde{G}^{\mu\nu}$). Furthermore,

requiring $G^{\mu\nu}$ to have Lorentzian signature enforces

$$\frac{1}{c_s^2} = 1 + 2X\frac{P_{,XX}}{P_{,X}} > 0. \tag{1.37}$$

1.2.5 Non-perturbative bimetric structures

In the previous section we saw how a bimetric description becomes applicable when considering perturbations in generic $P(X,\phi)$ theories. But what if the two metrics really are fundamental and already present at the level of background dynamics? Let us recall the generic bimetric theory in (1.3)

$$\mathcal{S} = \int d^4x\sqrt{g}\frac{M^2}{2}R + \mathcal{S}_m\left(\tilde{g}_{\mu\nu}, \Psi_i\right) + \mathcal{S}_\phi. \tag{1.38}$$

When we first introduced this action we stressed that $\mathcal{S}_\phi$ can be minimally coupled to $g_{\mu\nu}$ or $\tilde{g}_{\mu\nu}$ or indeed it may be a hybrid so that we can split $\mathcal{S}_\phi = \mathcal{S}_\phi\left[g_{\mu\nu}\right] + \mathcal{S}_\phi\left[\tilde{g}_{\mu\nu}\right]$.[3] Now in the case of dark energy/modified gravity theories, which we will discuss in the following section, non-perturbative bimetric structures of the type discussed are readily at hand, since $\mathcal{S}_m$ plays a crucial role in the dynamics. However, in considering early universe theories prior to reheating, one tends to ignore the effect of other matter particles. Instead it is typically assumed that universal dynamics are dominated by the scalar field, which only later on decays into standard model particles during reheating. We are thus effectively dealing with an action

$$\mathcal{S} = \int d^4x\sqrt{g}\frac{M^2}{2}R + \mathcal{S}_\phi. \tag{1.39}$$

An explicitly bimetric theory at the non-perturbative level therefore means we are dealing with either $\mathcal{S}_\phi = \mathcal{S}_\phi\left[\tilde{g}_{\mu\nu}\right]$ or $\mathcal{S}_\phi = \mathcal{S}_\phi\left[g_{\mu\nu}\right] + \mathcal{S}_\phi\left[\tilde{g}_{\mu\nu}\right]$. Let us immediately stress at this point that, while perturbations around a background specified by such a theory will be captured by the emergent bimetric structure outlined in the previous section, the presence of any other type of matter coupled to gravity during this epoch at all should break any degeneracies.

As a particularly simple case study for such a bimetric model we now focus on the scenario introduced in [23]. Introducing a cosmological constant into the $\tilde{g}_{\mu\nu}$-dependent $\mathcal{S}_\phi$ we have:

$$\mathcal{S}_\phi = \int d^4x\sqrt{\tilde{g}}(-2\tilde{\Lambda})\ . \tag{1.40}$$

Since $\sqrt{\tilde{g}}$ effectively generates a kinetic term plus higher derivative terms (we may confirm this

[3] Note that, while formulating such actions explicitly in terms of $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$ simplifies calculations and does bring out the causal structure of the theory out more clearly, we can of course always map the whole action into either frame, since $\tilde{g}_{\mu\nu} = \tilde{g}_{\mu\nu}\left(g_{\mu\nu}, \phi, X\right)$. No frame is more physical than the other, in the same sense that Jordan and Einstein frames are equally physical, but different features of the theory may of course be more readily visible in either frame. The salient point being that it is really a representational issue which frame(s) are chosen for the description of the dynamics.

by expanding out the root), this is sufficient to endow ϕ with dynamics, even in the absence of an explicit potential or kinetic term for ϕ. In fact, since $\tilde{T}^{\mu\nu} = \Lambda\tilde{g}^{\mu\nu}$ we find a Klein-Gordon equation for ϕ in the $\tilde{g}_{\mu\nu}$ frame [23]

$$\tilde{T}^{\mu\nu}\tilde{\nabla}_\mu\tilde{\nabla}_\nu\,\phi = \tilde{\Lambda}\tilde{g}^{\mu\nu}\tilde{\nabla}_\mu\tilde{\nabla}_\nu\,\phi = 0\,, \tag{1.41}$$

Notice that we refrain from calling $\tilde{g}_{\mu\nu}$ the matter frame now, since we have not specified anything about $\mathcal{S}_m$ yet.

It is interesting to notice, that (1.40) is the unique action yielding a Klein-Gordon equation of motion for ϕ in the $\tilde{g}_{\mu\nu}$ frame [23]. Mapping the ϕ-action into the Einstein frame we find

$$\mathcal{S}_\phi = \int d^4x\sqrt{-g}\sqrt{1+2B^2X}(-2\tilde{\Lambda})\,. \tag{1.42}$$

As shown in [23] such an action is already sufficient in order to generate scale-invariant fluctuations (we will return to this when investigating n-point correlation functions in chapter 3). Moreover something rather interesting happens when we pair the $\tilde{g}_{\mu\nu}$-frame action with a balancing $\mathcal{S}_\phi$ action minimally coupled to $g_{\mu\nu}$ and also impose $\tilde{\Lambda} = -1/2B^2$. As such we are now considering an action $\mathcal{S}_\phi$ which satisfies

$$\mathcal{S}_\phi = \int d^4x\sqrt{-\tilde{g}}\frac{1}{B^2} - \int d^4x\sqrt{-g}\frac{1}{B^2}\,, \tag{1.43}$$

When mapping this into the Einstein frame this generates a DBI-type action [22] for which

$$P(X,\phi) = -f^{-1}\sqrt{1-2fX} + f^{-1} - V(\phi) \tag{1.44}$$

where we have identified the DBI warp factor f via $f = -B^2$ and allowed for a potential $V(\phi)$. As such we are considering a minimal DBI-type model with constant warp factor. As we will discuss in chapter 3 this minimal model straightforwardly generates scale-invariant fluctuations. Now in the DBI case the sign of the warp factor is constrained to be positive by the signature of the metric. However, since we are purely considering the effective field theory determined by a DBI-type action and in fact explicitly constructed it without recourse to any higher-dimensional setup, f here is simply a free parameter. In summary, ϕ here has Klein-Gordon dynamics in the $\tilde{g}_{\mu\nu}$ frame, but behaves as a DBI field with constant warp factor in the Einstein frame.

There are various ways in which we can generalize this simple model. Promoting matching "cosmological constants" Λ and $\tilde{\Lambda}$ to potentials with ϕ-dependence, we obtain a DBI-model with ϕ-dependent warp factor

$$P(X,\phi) = -f^{-1}(\phi)\sqrt{1-2f(\phi)X} + f^{-1}(\phi) - V(\phi)\,. \tag{1.45}$$

Finally we may consider a fully bimetric $P(X,\phi)$ model with

$$\mathcal{S}_\phi = \int d^4x\sqrt{-\tilde{g}}\tilde{P}(\tilde{X},\phi) - \int d^4x\sqrt{-g}P(X,\phi). \tag{1.46}$$

This most general setup does not have to retain the DBI-form at all of course, but notice that the DBI-characteristic square-root form for the kinetic term, which generates higher-derivative interactions and consequently interesting phenomenology such as enlarged non-Gaussianities, is retained in all such bimetric models. Moreover, as we shall see later on in chapters 6, they retain second order equations of motion for ϕ, ensuring the absence of Ostrogradski-type ghosts [32] and, at least in the limiting simple cases outlined above, come equipped with an effective non-renormalization theorem to all terms in X^n as generated by the square root in the DBI action [35, 36].

1.2.6 Beyond scalar field inflation

1.2.6.1 Dual non-scalar theories

So far we have only considered early universe scenarios where dynamics is driven by a single scalar field. This simplest scenario may of course be modified ad infinitum by introducing extra degrees of freedom. But even without opening this particular Pandoric box, we have already discussed some such theories in disguise. Examples include multifield models where other more massive scalar degrees of freedom can be integrated out below a given energy scale, yet still influence the dynamics by modifying the structure of the effective kinetic term for the remaining single scalar degree of freedom. This happens for example in the Gelaton scenario [45] where two canonical scalar fields lead to an effective non-canonical single scalar theory below a given energy scale.

We may also consider theories where not a scalar (a 0-form), but some other n-form drives universal dynamics. Then an immediate worry is the breaking of isotropy and homogeneity, e.g. due to the appearance of preferred directions caused by additional degrees of freedom in a vector (1-form). But taking isotropy and homogeneity as constraints, higher n-form models frequently permit an effective scalar field description, since such constraints reduce the effective number of *dof* . This may result in e.g. a vector or vector-scalar theory where the vector can be integrated out consistently (regardless of energy-scale) to yield an effective scalar field description. A promising strategy is therefore to establish dualities between other n-form models, explicitly using a Hodge dual mapping between different n-forms, and then to explore whether dual classes of theories permit an effective scalar description under the assumption of global isotropy and homogeneity. In chapter 4 we will discuss this idea in more detail, focusing on the case of inflation driven by a massive three-form field [27, 28].

1.2.6.2 A bimetric alternative to inflation?

Above we introduced non-perturbative bimetric structures in the context of inflation. But interest in alternatives to inflation has never floundered - after all, even if we were solely interested in testing inflation, the ultimate test of any theory is to see how well it fares against any of its (potential) competitors. Above we have seen that a great variety of inflationary scenarios can be couched in a bimetric formalism. Here we quickly review how "varying speed of light" (VSL) scenarios [46, 47, 48], proposed as an alternative to inflation, also permit a bimetric formulation and how/whether they address the Big Bang problems. For some related recent work on VSL theories see [49, 50, 23, 8, 7].

VSL theories posit that the invariant speed of propagation associated with "matter" in the early universe differs from that we see at present. This can be naturally implemented with an action of the form (1.3), i.e.

$$\mathcal{S} = \int d^4x\sqrt{g}\frac{M^2}{2}R + \mathcal{S}_m\left(\tilde{g}_{\mu\nu}, \Psi_i\right), \tag{1.47}$$

where we ignore any potential additional scalar degree of freedom for now. Then R is constructed out of a gravity metric $g_{\mu\nu}$, with invariant speed $c_{gravity}$ while matter couples to the metric $\tilde{g}_{\mu\nu}$ with invariant speed c_{matter}, where we will be especially interested in c_{photon} which is identical to c_{matter} here. We will now consider how this setup can be used to address the Big Bang problems, closely following [51].

The horizon problem: We have already discussed the horizon problem in a simple inflationary context in section 1.2.1, but here we will need to be more careful since the introduction of different invariant speeds also introduces different horizons. The size of the (gravity) particle horizon when the CMB is formed at t_* is

$$R_{particle-horizon}(t_*) = \int_0^{t_*} \frac{c_{gravity}\, dt}{a(t)}, \tag{1.48}$$

while for photons one finds

$$R_{photon-horizon}(t_*) = \int_0^{t_*} \frac{c_{photon}\, dt}{a(t)} \geq R_{particle-horizon}(t_*). \tag{1.49}$$

where the inequality holds if $c_{photon} \geq c_{gravity}$ throughout the period integrated over. $R_{photon-horizon}$ is of particular interest since it determines the distance scale over which photons can establish causal contact between different regions in the early universe and thermalize them. In order to address the horizon problem this has to be compared to the coordinate distance to the CMB

(surface of last scattering), which at the present time t_0 is

$$R_{last-scattering}(t_*, t_0) = \int_{t_*}^{t_0} \frac{c_{photon}\, dt}{a(t)}. \tag{1.50}$$

In order for all CMB scales which we can observe presently to have been in causal contact when the CMB was formed, we require

$$R_{photon-horizon}(t_*) \geq 2\, R_{last-scattering}(t_*, t_0) \tag{1.51}$$

or equivalently

$$3\, R_{photon-horizon}(t_*) \geq 2\, R_{photon-horizon}(t_0). \tag{1.52}$$

The VSL solution to satisfying this constraint is to require $c_{photon} \gg c_{gravity}$ in the primordial universe. Intuitively this is straightforward to understand, since a "larger" speed of light c in the early universe allows regions separated by greater distances to become causally connected.[4]

The flatness problem: From the Friedmann equations we have

$$\Omega - 1 = \frac{K\, c_{gravity}^2}{\dot{a}^2}, \tag{1.53}$$

as before, but we explicitly keep track of factors of c this time. Now the fact that we have a factor of $c_{gravity}$, and *not* c_{photon}, is important. This is the case since (1.53) comes from varying the Einstein–Hilbert action. In the bimetric picture the c here therefore cannot be c_{photon}. As such we have

$$\dot{\Omega} = -2(\Omega - 1)\left(\frac{\ddot{a}}{\dot{a}}\right), \tag{1.54}$$

which also does not depend on c_{photon} and is therefore unaffected by $c_{photon} \neq c_{gravity}$. This also means that $\ddot{a} > 0$, i.e. inflation, is still required in order for $\Omega = 1$ to become an attractor. As such a bimetric VSL theory of the type considered does not provide an independent solution to the flatness problem. Of course this does not mean that other implementations of the VSL mechanism cannot solve the flatness problem in different ways (for a review of different implementations, e.g. hard breaking of Lorentz invariance via the introduction of preferred frames, see [48]).

Just as for inflation, we defer a discussion of quantizing perturbations and structure formation in the early universe to later chapters (for bimetric theories with $c_s \gg 1$ we will do so in chapter 3). But we conclude this section with two observations. Firstly the fact that we have a metric formulation means that, even though the invariant speed of one metric might appear "superluminal" in the frame defined by the other metric, no closed causal curves (CCC) or other

[4] Typically one require $c_{photon} \approx c_{gravity}$ between last scattering and today in order to minimize modifications to late-time cosmology.

causality problems arise since e.g. light and gravity signals transform under two different copies of the Lorentz group [52, 49, 23]. In fact, as long as there is a hyper-surface which is space-like w.r.t. both metrics, the Cauchy problem is well-defined [21]. Secondly note that the use of the word "superluminal", as frequently employed when describing $c_{photon} > c_{gravity}$, is rather misleading. What is important here is the dimensionless ratio of invariant speeds $c_{photon}/c_{gravity}$ and in the VSL scenario outlined this is always ≥ 1. So signals can propagate faster than the invariant speed of the gravity metric. However, light always propagates with c_{photon} and nothing propagates faster than light here.

1.3 Prologue III : Dark Energy and Modified Gravity

One of the most important discoveries in recent cosmology is the accelerated late-time expansion of the universe. Via the use of different standard rulers (Baryon acoustic oscillations, CMB and Supernovae data) it was found that, assuming general relativity is valid all the way up to cosmological length-scales, around $\sim 70\%$ of the effective energy density of the universe is provided for by a component exerting negative pressure and hence causing accelerated expansion. Moreover its effective equation of state appears to be very close to $w = -1$, signaling that whatever this dark component may be, it very closely mimics a cosmological constant. This leaves us with a few options: Either cosmic acceleration is due to existing *dof* , e.g. sourced by a cosmological constant or back-reaction, or dark energy/modified gravity is dynamical, i.e. due to the presence of new *dof* . This may be due to a modification of the RHS of the Einstein equation - a new type of matter acting as "dark energy" - or due to new gravitational *dof* - a modification of the LHS of the Einstein equation. Clearly this is not a sharp distinction and whether a new *dof* is described as gravitational or "matter" will depend on how it enters into the model under consideration.

In the context of modified gravity the appearance of new *dof* is a generic consequence of departing from GR. This is to be expected given Weinberg's theorem [53] which states that a Lorentz invariant theory of a massless spin two field must be equivalent to GR in the low energy limit. Infrared modifications of gravity therefore necessarily introduce new *dof* . In the simplest case which we are interested in here, there is only one new *dof* and it is a scalar. The way this scalar is introduced is of course model-dependent. In massive gravity models, where the graviton is equipped with a small mass or resonance width, the new scalar *dof* appears when a Stückelberg scalar is introduced in order to make gauge invariance manifest. In higher-dimensional setups new scalar *dof* appear in the effective 4D theory due to dimensional reduction or, in the case of a probe brane, as an induced scalar specifying the position modulus of the brane in the bulk.

The result is a scalar-tensor theory of modified gravity. The notion of screening mecha-

nisms [25] - how a light scalar degree of freedom ϕ can act as dark energy on cosmological scales while being shielded in dense environments such as on earth - has turned out to be especially useful in this context. Implementations of such a mechanism include the following: 1) The chameleon model [15, 16], where a density dependent mass is generated and the field ϕ becomes too massive for detection in dense environments. 2) Vainshtein screened setups [54, 55, 56] such as DGP [17] and Galileon [29]/Horndeski [33] models, where non-linear interactions of ϕ lead to strongly coupled dynamics. A density-dependent (classical) renormalization of the kinetic energy there results in an effectively decoupled scalar in dense environments. 3) Symmetron models [57, 58, 59], where a scalar ϕ is coupled to matter with a coupling strength proportional to the vacuum expectation value of ϕ. This in turn depends on the ambient density, so that the scalar effectively decouples in high-density regions. All these mechanisms reconcile the existence of a light cosmological scalar with tight fifth force constraints on solar-system scales [26]. In this section we review some of the relevant modified gravity models and why one might want to consider them rather than e.g. models where accelerated expansion is sourced by a cosmological constant. We show how they are related to disformal gravity and why bimetric setups naturally generate such modified gravity models.

1.3.1 A cosmological constant?

Having spent some time motivating why extra *dof* may be considered if we want to realize a phase of accelerated late-time expansion, let us briefly step back and recap why the simplest solution - plain GR with a cosmological constant (CC) - may be considered unsatisfactory. Introducing an explicit cosmological constant Λ the Einstein equation becomes

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} + \Lambda g_{\mu\nu}. \tag{1.55}$$

Even if we do not introduce a cosmological constant by hand, the vacuum energy of all matter species contained in $T_{\mu\nu}$ generates an effective cosmological constant, so that overall one finds an effective cosmological constant

$$\Lambda_{eff} = \Lambda + 8\pi G \langle \rho_{vac} \rangle. \tag{1.56}$$

Now, from a quantum field theory point of view, GR is an effective field theory valid up to some cutoff scale, since it is non-renormalizable. The naive expectation for the size of the contribution of the vacuum energy of matter to Λ_{eff} depends on this cutoff scale, so that if we assume GR to be valid all the way up to the Planck scale we expect

$$\Lambda_{eff} \sim \left(10^{27} eV\right)^4. \tag{1.57}$$

This is rather catastrophically different from the observed value

$$\Lambda_{eff} \sim \left(10^{-3} eV\right)^4 . \tag{1.58}$$

Even if we lower the cutoff scale to, say, the supersymmetry breaking scale, a difference of around 60 orders of magnitude remains.

We can see that this is a problem even at very low energies. Following [14] we may consider the low-energy effective theory obtained by integrating out all the contributions down to just above the electron mass m_e. As such Λ_{eff} will include a bare term Λ_1 and terms coming from low-energy loops of particles with masses $\leq m_e$. Schematically this means

$$\Lambda_{eff} = \Lambda_1 + c_e m_e^4 + c_\nu m_\nu^4 \ldots , \tag{1.59}$$

where c_e and c_ν are coefficients determined by computing the loop corrections for each matter species. We now lower the energy and integrate out the electron, arriving at

$$\Lambda_{eff} = \Lambda_0 + c_\nu m_\nu^4 \ldots , \tag{1.60}$$

where Λ_0 is a new bare cosmological constant term. Preserving the same observed value of Λ_{eff}, Λ_1 and Λ_0 need to cancel to over 30 decimal places. In other words, loop corrections to any bare cosmological constant at these energies are generically over 30 orders of magnitude larger than the resulting overall Λ_{eff}, asking for a miraculous cancellation of different contributions to the effective CC. In this sense a CC with the observed value is technically unnatural in the absence of some mechanism/symmetry protecting it. This motivates looking for a different agent causing late-time acceleration. The cosmological constant problem (CCP) of course still needs to be addressed, but interestingly new *dof* might very well be introduced as a consequence of whichever symmetry we impose in order to solve the CCP. Putting the cosmological constant problem to one side, we therefore now move on to investigate how new *dof* might be used in order to drive accelerated space-time expansion.

1.3.2 Screening mechanisms I : Chameleons

Chameleon models [15, 16] are typically constructed with the use of two conformally related metrics, where the conformal factor is a function of the chameleon field ϕ only. In particular we may consider an action of the following form

$$\mathcal{S} = \int d^4x\sqrt{g}\left(\frac{M^2}{2}R + X - V(\phi)\right) + \mathcal{S}_m\left(A^2(\phi)g_{\mu\nu}, \Psi_i\right), \tag{1.61}$$

All matter species Ψ_i are therefore universally coupled to the matter metric $\tilde{g}_{\mu\nu} = A^2(\phi)g_{\mu\nu}$. This universal coupling ensures the validity of the weak equivalence principle by design. In other words we consider the generic action (1.3) subject to the simplifying assumptions of a purely conformal relation between metrics $\tilde{g}_{\mu\nu} = A^2(\phi)g_{\mu\nu}$ and a simple canonical $\mathcal{S}_\phi = \int d^4x\sqrt{g}X - V(\phi)$ minimally coupled to the Einstein metric $g_{\mu\nu}$. As usual a matter stress-energy tensor can be defined with respect to the matter ("Jordan frame") metric $\tilde{g}_{\mu\nu}$, so that

$$\tilde{T}^{\mu\nu} = \frac{2}{\sqrt{\tilde{g}}}\frac{\delta\left(\sqrt{\tilde{g}}\mathcal{L}_m\right)}{\delta\tilde{g}_{\mu\nu}}. \tag{1.62}$$

Since it minimally couples to $\tilde{g}_{\mu\nu}$, it is covariantly conserved with respect to that metric $\tilde{\nabla}_\mu\tilde{T}^{\mu\nu} = 0$. When mapping $\tilde{T}^{\mu\nu}$ into the Einstein frame one finds

$$\tilde{T}^{\mu\nu} = A^{-6}(\phi)T^{\mu\nu}. \tag{1.63}$$

This means $T^{\mu\nu}$ is not covariantly conserved in the Einstein frame $\nabla_\mu T^{\mu\nu} \neq 0$. Turning our attention back to the original action (1.61), one can now work out the equation of motion for the scalar ϕ obtaining

$$\Box\phi = V_{,\phi} - A^3(\phi)A_{,\phi}\tilde{T}, \tag{1.64}$$

where the matter stress-energy tensor is defined as in (1.62) and has been contracted with the matter metric $\tilde{T} = \tilde{T}^{\mu\nu}\tilde{g}_{\mu\nu}$. Specializing to the case of a pressureless, non-relativistic source, the only non-vanishing component of the stress-energy tensor $\tilde{T}^\mu_\nu$ is $\tilde{T}^0_0 = -\tilde{\rho}$. Following from (1.63) the energy density of matter in the Einstein frame, ρ, is given by $\rho = A^4\tilde{\rho}$. As a direct consequence of $\tilde{\nabla}_\mu\tilde{T}^{\mu\nu} = 0$ we then also find a conserved quantity in the Einstein frame $\hat{\rho} = A^3\tilde{\rho} = A^{-1}\rho$. In terms of this conserved quantity (1.64) can therefore be written as

$$\Box\phi = V_{,\phi} + A_{,\phi}\hat{\rho}, \tag{1.65}$$

and one can integrate up to obtain an effective potential for ϕ

$$V_{\text{eff}}(\phi) = V(\phi) + \hat{\rho}A(\phi), \tag{1.66}$$

since $\hat{\rho}$ is conserved and independent of ϕ in the Einstein frame.

This setup has interesting phenomenological consequences. Suppose we start with a runaway potential $V(\phi)$, e.g. $V(\phi) = M_{Pl}^{n+4}/\phi^n$ as desirable from the point of view of quintessence models [60]. This ensures that, in the limit when we can ignore the matter action $\mathcal{S}_m$ (a low-density environment with $\hat{\rho} \to 0$ in the language set out above), we recover a quintessence-like solution to the cosmological constant problem. A slow-rolling light scalar field then drives

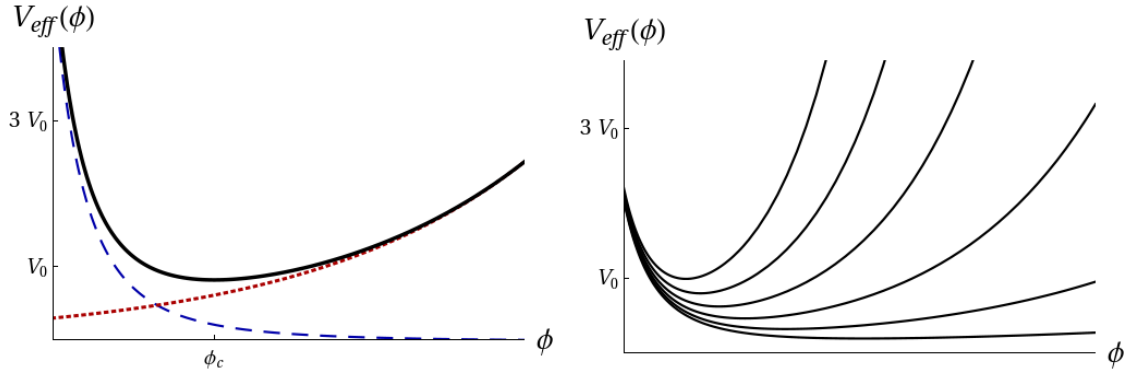


Figure 1.1: **Left**: Plot showing the effective potential $V_{eff}(\phi)$ (solid line), the conformal factor $A(\phi) = e^{k_1\phi}$ (dotted line) and a runaway potential $V(\phi) \sim \phi^{-3}$ (dashed line) in arbitrary units, cf. equation (1.66). **Right**: Plot showing the effective potential for $k_1 = 1 - 6$ from bottom to top. Note how the position of the minimum ϕ_{min} is shifted to smaller ϕ and the curvature and hence mass of oscillations around ϕ_{min} is enhanced as k_1 increases.

accelerated expansion of space-time. But in regions of high density this behavior changes in the following way. Naively ϕ does not possess a mass term at all, since $V(\phi)$ has no minimum (except for the trivial one at $\phi \to \infty$). However, from (1.65) it becomes clear that, for non-zero $\hat{\rho}$, a suitably chosen $A(\phi)$ can result in a V_{eff} which does have a minimum $\phi_{\min}$, such that

$$V_{eff,\phi}(\phi_{\min}) = V_{,\phi}(\phi_{min}) + A_{,\phi}(\phi_{min})\hat{\rho} = 0, \tag{1.67}$$

where the mass of the field m for small oscillations around the minimum ϕ_{min} is given by

$$m^2 \equiv V_{eff,\phi\phi}(\phi_{\min}) = V_{,\phi\phi}(\phi_{\min}) + \hat{\rho}A_{,\phi\phi}(\phi_{\min}). \tag{1.68}$$

This is the essence of the chameleon mechanism: An environmentally-dependent way of generating a large mass for an otherwise very light scalar ϕ. This reconciles a model such as (1.61) with fifth force constraints, since ϕ becomes too heavy for detection in laboratory experiments on earth, yet can act as dark energy on large scales. Figure 1.1 illustrates the chameleon mechanism for a conformal factor of the form $A(\phi) \sim e^{k_1\phi}$.

Two final comments are in order here. Firstly our argument has been purely phenomenological in that a suitably chosen conformal factor can give rise to the chameleon mechanism described, but no arguments for why such an $A(\phi)$ is (technically) natural have been given. This issue is beyond the scope of this thesis, but we refer to [61] for a discussion. Secondly we have not discussed the thin-shell mechanism so far, which is another essential ingredient for the viability of the chameleon mechanism in that it suppresses fifth force modifications to e.g. planetary orbits. We will do so in section 5.5 where radial solutions are investigated in some detail.

1.3.3 Screening mechanisms II : Vainshtein

The Vainshtein mechanism takes a different approach to screening the scalar ϕ in dense environments. Here the idea is that, while ϕ has a natural $\mathcal{O}(1)$ coupling to matter on cosmological scales, in dense environments matter becomes effectively decoupled from ϕ. In order to do so the Vainshtein mechanism relies on the introduction of non-linearities into the kinetic term for ϕ. Equivalently one may explicitly introduce higher derivative interactions into the Lagrangian for a canonical scalar field ϕ, so that above some scale Λ[5] the field becomes (classically) strongly coupled and as a result effectively decouples from other fields or matter in general. Schematically we can picture this mechanism as follows [62]. Consider the theory

$$\mathcal{L} = -\frac{1}{2}Z^{\mu\nu}(\phi)\partial_\mu\phi\partial_\nu\phi + \frac{\phi}{M_{Pl}}T^\mu_\mu\,, \tag{1.69}$$

where $Z^{\mu\nu}$ essentially plays the role of an effective metric and we have introduced a very simple conformal coupling to matter ϕT^μ_μ[6]. We have ignored any potential $V(\phi)$ since its form is irrelevant for the Vainshtein mechanism. Now we introduce a non-standard kinetic term/explicitly add higher derivative interactions to the Lagrangian by writing [62]

$$Z_{\mu\nu}(\phi) \sim g_{\mu\nu} + \frac{1}{\Lambda^3}\partial_\mu\partial_\nu\phi + \frac{1}{\Lambda^6}(\partial_\mu\partial_\nu\phi)^2 + \cdots \tag{1.70}$$

In the low-energy regime, which corresponds to cosmological scales, we can ignore derivative terms $(\partial_\mu\partial_\nu\phi)^n$ since they are suppressed by Λ^{3n}. A canonically normalized kinetic term is then recovered with $Z_{\mu\nu} \sim g_{\mu\nu}$ and the coupling between ϕ and other matter is unsuppressed. At high energies, however, $Z(\phi_0) \gg 1$, where ϕ_0 is the background configuration of the field ϕ. Fluctuations around the background therefore acquire a kinetic term whose dominant contribution is so that the dominant contribution to the kinetic term is of the form $Z(\phi_0)(\partial\varphi)^2 \sim (\partial^2\phi_0/\Lambda^3)^n(\partial\varphi)^2$, for some (model-dependent) n. Canonically normalizing the field one obtains $\hat{\varphi} = \varphi/\sqrt{Z(\phi_0)}$, so that fluctuations $\hat{\varphi}$ have an effective coupling to matter $(M_{Pl}\sqrt{Z(\phi_0)})^{-1} \ll M_{Pl}^{-1}$. When $Z(\phi_0) \gg 1$ the field therefore effectively decouples from other matter and its dynamics becomes dominated by its own self-interactions.

DBI probe brane and "irrelevant" interactions: A particularly interesting example of a theory where higher derivative terms as discussed above are important, is that of a 4D DBI probe brane embedded in a 5D bulk. Why do we choose this example? While the procedure outlined above is perfectly valid, it typically needs an extra ingredient in order to be implemented in an

[5] Note that the scale Λ, in an abuse of notation, has nothing to do with the cosmological constant.

[6] Note, however, that the non shift-symmetric matter coupling ϕT^μ_μ has some interesting consequences. While this coupling is technically natural when just considering ϕ-loops [63], matter loops do not preserve the form of such a coupling [64]. For a discussion of whether or not this is any worse than the normal cosmological constant problem see [63, 64].

EFT . The higher derivative interactions outlined above are irrelevant from an *EFT* point of view and hence non-renormalizable. Typically one would expect the *EFT* description to only be valid as long as such terms are subdominant/suppressed by some scale Λ which is then equivalent to the cutoff scale beyond which the *EFT* is no longer valid. In the approach above, however, we are precisely interested in regimes where the higher-derivative interactions are important, e.g. $\partial^2\phi \sim \Lambda^3$. Generically one would therefore expect loop corrections to spoil the validity of the *EFT* under consideration.

An additional symmetry is needed in order to keep such corrections under control, giving rise to a strong coupling regime where derivative interactions of the type considered are important, yet higher derivative terms $(\partial^n\phi)^m$ with $n, m > 2$ generated by quantum corrections are irrelevant. In the case of an unwarped DBI probe brane a non-linearly realized 5D Poincare invariance $ISO(1,4)$ protects the form of terms in the Lagrangian [18]. Interestingly we can decompose this larger symmetry group into the 4-dimensional Poincare group, a shift symmetry $\pi \to \pi + c$ protecting the canonical kinetic term X and coming from translation invariance in the 5th dimension, and an additional symmetry associated with rotations in the 5th dimension. This extra symmetry will turn out to be crucial and we will meet it again below. One can now prove a non-renormalization theorem [35, 36] stating that terms invariant under these symmetries are not renormalized and establishing the existence of a stable strong-coupling regime.

The probe brane Lagrangian and Lovelock invariants: Now a generic form for a probe brane Lagrangian that will be invariant under all these symmetries is [18]

$$\mathcal{L} = \sqrt{g}F\left(\tilde{g}_{\mu\nu}, \tilde{R}_{\alpha\beta\gamma\delta}, \tilde{K}_{\rho\sigma}, \tilde{\nabla}_\epsilon\right) + A\pi, \tag{1.71}$$

where F is some scalar constructed out of $\tilde{g}_{\mu\nu}$, the induced metric on the probe brane (which we may expect matter fields to couple to minimally - hence the choice of notation $\tilde{g}_{\mu\nu}$), and $\tilde{R}_{\alpha,\beta,\gamma,\delta}, \tilde{K}_{\rho\sigma}, \tilde{\nabla}_\epsilon$, which are the associated Riemann tensor, extrinsic curvature and covariant derivative respectively. $A\pi$ is the tadpole term which is invariant since $\delta_\mu\pi$ is a total derivative. The induced metric on the brane is of the form [18]

$$\tilde{g}_{\mu\nu} = g_{\mu\nu} + \partial_\mu\pi\partial_\nu\pi. \tag{1.72}$$

π is the position modulus of the brane along the 5th dimension and $g_{\mu\nu}$ is the induced metric on constant π hypersurfaces. Here we will focus on the simple case when $g_{\mu\nu} = \eta_{\mu\nu}$, i.e. it is Minkowski. The unique geometrical quantities which are free of any ghost-like degrees of freedom are then given by the Lovelock invariants [30, 65, 18]. In addition to the tadpole term π the

relevant terms are

$$F_\lambda = -\lambda \int \mathrm{d}^4x\sqrt{-g} = -\lambda \int \mathrm{d}^4x\sqrt{1+(\partial\pi)^2} \tag{1.73}$$

$$F_K = -M_5^3 \int \mathrm{d}^4x\sqrt{-g}K = M_5^3 \int \mathrm{d}^4x \ \left([\Pi] - \gamma^2[\phi]\right) \tag{1.74}$$

$$F_R = \frac{M_4^2}{2} \int \mathrm{d}^4x\sqrt{-g}R = \frac{M_4^2}{2} \int \mathrm{d}^4x\, \gamma \left(([\Pi]^2 - [\Pi^2]) + 2\gamma^2 \left([\phi^2] - [\Pi][\phi]\right)\right) \tag{1.75}$$

$$\begin{aligned} F_{GB} &= -\beta\frac{M_5^3}{m^2} \int \mathrm{d}^4x\sqrt{-g}\,\mathcal{K}_{\mathrm{GB}} \\ &= \beta\frac{M_5^3}{m^2} \int \mathrm{d}^4x\,\gamma^2\Big(\frac{2}{3}\left([\Pi]^3 + 2[\Pi^3] - 3[\Pi][\Pi^2]\right) + 4\gamma^2([\Pi][\phi^2] - [\phi^3]) \\ &\qquad - 2\gamma^2([\Pi]^2 - [\Pi^2])[\phi]\Big), \end{aligned} \tag{1.76}$$

where $\gamma = 1/\sqrt{1+(\partial\pi)^2}$ and we follow the notation of [29, 18] with $\Pi_{\mu\nu} = \partial_\mu\partial_\nu\pi$ and square brackets [...] denoting the trace operator (w.r.t. the Minkowski metric $\eta_{\mu\nu}$). Also $[\phi^n] \equiv \partial\pi\,.\,\Pi^n\,.\,\partial\pi$. The form of these interactions is then protected by the nonlinearly realized 5D Poincare symmetry. This becomes especially clear when taking the non-relativistic limit $(\partial\pi)^2 \ll 1$ of the Lovelock invariants shown above. We obtain [18]

$$F_2 = S_\lambda^{NR} = -\frac{\lambda}{2} \int \mathrm{d}^4x\,(\partial\pi)^2 \tag{1.77}$$

$$F_3 = S_K^{NR} = \frac{M_5^3}{2} \int \mathrm{d}^4x\,(\partial\pi)^2\,\Box\pi \tag{1.78}$$

$$F_4 = S_R^{NR} = \frac{M_4^2}{4} \int \mathrm{d}^4x\,(\partial\pi)^2 \ \left((\Box\pi)^2 - (\partial_\mu\partial_\nu\pi)^2\right) \tag{1.79}$$

$$F_5 = S_{GB}^{NR} = \beta\frac{M_5^3}{3m^2} \int \mathrm{d}^4x\,(\partial\pi)^2 \left((\Box\pi)^3 + 2(\partial_\mu\partial_\nu\pi)^3 - 3\Box\pi(\partial_\mu\partial_\nu\pi)^2\right). \tag{1.80}$$

which are the Galileon terms first introduced by [29]. The additional symmetry coming from rotations in the 5th dimension here manifests itself in the form of a non-relativistic Galilean symmetry, from which the theory inherits its name

$$\pi \to \pi + c + b_\mu x^\mu. \tag{1.81}$$

In other words, in addition to invariance (up to total derivatives) under $\pi \to \pi + c$ we now have an extra shift symmetry $\partial_\mu\pi \to \partial_\mu\pi + b_\mu$. We therefore have a theory which has the right type of interactions to exhibit Vainshtein-screening and is also a robust *EFT* due to the fact that the interaction terms are protected from renormalizations by the Galilean shift symmetry (or its relativistic generalization). Note that the equations of motion resulting from these Galileon terms are purely second order as a result of the Galilean shift symmetry. They are therefore

a special case of the setups proposed by [33, 34] which derived the most general scalar-tensor theories with second *and* first order equations of motion. Such terms are still protected from the appearance of Ostrogradski-type ghosts, but do not obey the Galilean shift symmetry. Protection from quantum corrections is therefore lost and one relies on high energy physics to fortuitously cancel quantum corrections and conspire to give the effective Lagrangian under consideration.

1.3.4 Conformal/disformal geometry and screening

Reminding ourselves of the generic disformal relation

$$\tilde{g}_{\mu\nu} = A^2(\phi, X) g_{\mu\nu} + B^2(\phi, X) \partial_\mu \phi \partial_\nu \phi \tag{1.82}$$

an interesting picture emerges as regards screening in modified gravity theories.

Chameleon screening: In the standard scenario [15, 16] this arises in the conformal limit of the disformal relation, i.e. $B^2 = 0$, and for a particularly simple non-derivative conformal factor $A^2 = A^2(\phi)$. In chapter 5 we will show that chameleon screening remains unaffected by the introduction of disformal terms, but that a more generalized conformal factor $A^2(\phi, X)$, i.e. a conformal factor which explicitly does depend on derivatives of the field, can lead to interesting alterations to the chameleon mechanism. As such the conformal limit of the disformal relation becomes associated with a generalized chameleon model.

Vainshtein screening: In the previous section we saw how DBI galileons [18] arise when an induced metric $\tilde{g}_{\mu\nu} = g_{\mu\nu} + \partial_\mu \phi \partial_\nu \phi$ is present. This is of course a trivial disformal metric with $A^2 = B^2 = 1$. In fact the Vainshtein screened terms F_λ and F_R arise geometrically from $\tilde{g}_{\mu\nu}$ from the determinant of $\tilde{g}_{\mu\nu}$ and its corresponding Ricci scalar. As we will show in chapter 6 the remaining terms arise naturally when the full triformal relation is taken into account. Analogously we expect that generalizing the triformal metric (1.7) by considering non-trivial B^2 and C^2 generates the setups alluded to above and proposed by [33, 34], i.e. the most general scalar-tensor theories with second *and* first order equations of motion. The salient point here is that the disformal factor B^2 (and the triformal factor C^2) controls the generation of higher derivative interactions and hence Vainshtein screening.

While chameleon phenomenology is a result of considering non-trivial conformal limits of (1.82), Vainshtein screening therefore arises due to non-trivial disformal factors B^2 (and triformal factors C^2). As such the different terms in (1.82) acquire a rather instructive phenomenological significance.

Part II

Inflation and early universe physics

Chapter 2

Constraining single field inflation (beyond slow-roll)

2.1 Introduction

In prologue II of the introduction 1.2 we saw that disformal bimetric models generically and non-perturbatively generate $P(X, \phi)$ setups. Going the other way, we also discovered that, when considering perturbations in any $P(X, \phi)$, an effective bimetric desription arises. Disformal bimetric models are therefore intimately tied to non-canonical actions. As such we would like to investigate such $P(X, \phi)$ theories in some detail in this chapter, focusing on the inflationary regime. In particular we would like to ask what generic constraints we can place on inflationary parameter space for single-field models of inflation as arising in a disformal context. One of the primary challenges in doing so is to understand inflationary physics beyond the slow-roll paradigm. And indeed, there has been a lot of progress recently in understanding inflationary phenomenology without assuming slow-roll conditions [66, 1, 67, 68, 69], i.e. where inflation is not almost de Sitter. An important question then poses itself: How far may we depart from the standard slow-roll regime without coming into conflict with observational and theoretical constraints? More specifically, what bounds can we place on "slow roll" parameters (which measure the "distance" from purely de Sitter expansion)? In this chapter we present a number of such constraints for generic classes of inflationary single field models.

Departure from pure de Sitter expansion generically breaks the scale invariance of n-point correlation functions for the curvature perturbation ζ. However, present-day data constrain the 2-point function (the power spectrum) to be near scale-invariant [9]. Generic single field models can restore this observed behaviour via the introduction of non-canonical kinetic terms and hence a time-varying "speed of sound" c_s [66]. In doing so, large interaction terms are produced at the level of the cubic action, leading to the generation of large levels of non-Gaussianity.

These will be heavily constrained by CMB and large scale structure surveys in the near future and as such non-Gaussianity becomes an excellent tool for constraining slow-roll parameters and selecting models of primordial structure formation [9, 70, 71, 72, 73, 74]. One of the most interesting questions one may hope to answer in this way is: Can fast-roll models of inflation be distinguished observationally?

We investigate the phenomenology associated with such models, extending the work by [66] and showing that these models have several distinct observable signatures. These include a generically large running of non-Gaussianity n_{NG} with scale, the suppression of the size of non-Gaussianity $f_{\rm NL}$ by non-slow-roll ϵ_i and the modification of non-Gaussian shapes for single field models. These considerations widen the number of potential future observations that could be explained by single field models. And even if future measurements turn out to favor multifield models, it is still important to fully understand the potential effects of individual fields that may participate in such models.

The plan for this chapter is as follows. In section 2.1.1 we quickly review $P(X,\phi)$ theories and discuss the requirement of weak coupling. In section 2.2 we then set the stage by considering the two-point correlation functions for scalar and tensor perturbations in general single field models. We compute the resulting observational constraints on slow-roll parameters and also summarize bounds coming from other sources. Then, in section 2.4, we give the explicit non-Gaussian amplitudes for general single field models without assuming slow-roll conditions (section 2.4.1). Observational signatures are considered in detail for $f_{\rm NL}^{equi}$ (section 2.4.2), non-Gaussianity shapes (section 2.4.3) and running with scale n_{NG} (section 2.4.4). In section 2.5 concrete models are provided for the phenomenologically interesting cases by explicitly constructing their actions. Finally, in section 2.6, we summarize the salient results of this chapter.

2.1.1 The setup

As introduced in section 1.2.2, we consider general single field inflation models described by an action

$$S = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{R}{2} + P(X,\phi)\right] \ , \tag{2.1}$$

where $X = -\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$. We also remind ourselves of the hierarchy of slow-roll parameters

$$\epsilon \equiv -\frac{\dot H}{H^2},\ \eta \equiv \frac{\dot\epsilon}{\epsilon H},\ldots \qquad \epsilon_s \equiv \frac{\dot c_s}{c_s H},\ \eta_s \equiv \frac{\dot\epsilon_s}{\epsilon_s H},\ldots \qquad \text{, where } c_s^2 = \frac{P_{,X}}{P_{,X}+2XP_{,XX}} \tag{2.2}$$

where the non-canonical nature of the action is essentially quantified by

$$c_s^2 = \frac{P_{,X}}{\rho_{,X}} = \frac{P_{,X}}{P_{,X} + 2XP_{,XX}}\,, \tag{2.3}$$

$$\Sigma = XP_{,X} + 2X^2P_{,XX} = \frac{H^2\epsilon}{c_s^2}\,, \tag{2.4}$$

$$\lambda = X^2P_{,XX} + \frac{2}{3}X^3P_{,XXX}\,, \tag{2.5}$$

and c_s is the speed of sound with which perturbations propagate. A near de Sitter expansion is associated with the slow-roll regime, where $\epsilon, \eta ... \epsilon_s, \eta_s ... \ll 1$ and accelerated expansion takes place as long as $\epsilon < 1$. We now wish to understand what constraints can be placed on these parameters. For simplicity we here focus on the case where slow-roll is broken at the first level in the hierarchy, for ϵ and ϵ_s, but assume that slow-roll conditions still hold for higher order parameters, setting $\eta, \eta_s \sim 0$. In other words, following [66, 3, 1] we consider models where slow-roll conditions are broken by $\epsilon, \epsilon_s \sim \mathcal{O}(1)$ and constant.

In deriving the constraints presented here we will firstly map present-day observational bounds, especially those coming from the WMAP experiment [9], onto the parameter space of fast-rolling models. As a second guidance principle we will impose a minimal theoretical constraint: The fluctuations described by (2.1) should remain weakly coupled for at least ~ 10 e-folds. This range corresponds to the observable window of scales where primordial non-Gaussianity may be measured (running from CMB, $k^{-1} \sim 10^3$ Mpc, to galactic scales, $k^{-1} \sim 1$ Mpc).

Why require weak coupling at all? Strong coupling scales are frequently associated with the appearance of new physics. In the standard model, for example, the would-be strong coupling scale lies around $\sim 1\ TeV$, before the Higgs is introduced. We may expect an analogue to be true for single field inflation models, especially given the generic presence of other massive degrees of freedom (dof) in UV completions of primordial physics. Such *dofs* may be integrated out at low energies, but can become relevant around the would-be strong coupling scale [45, 75]. If so, predictions beyond this scale will depend on exactly how and which *dofs* enter. Flipping the argument around, even if we were able to calculate the dynamics for generic strongly coupled systems, one should remain cautious whether the effective field theory under consideration is valid anymore in such circumstances. As such we will require weak coupling to ensure that (2.1) is predictive over at least the observable window of scales where primordial fluctuations may be measured.

2.2 Power spectra

2.2.1 The spectral index n_s

In this section we sketch the calculation of the primordial power spectrum of scalar perturbations. In prologue II we defined perturbations in the FLRW metric and the associated curvature perturbation on uniform density hypersurfaces ζ (see equations (1.21) and (1.25)). We now perturb (2.1) to quadratic order in ζ and obtain [20]

$$S_2 = \frac{M_{\rm Pl}^2}{2}\int \mathrm{d}^3x\mathrm{d}\tau\; z^2\left[\left(\frac{d\zeta}{d\tau}\right)^2 - c_s^2(\vec{\nabla}\zeta)^2\right], \tag{2.6}$$

where τ is conformal time, and z is defined as $z = a\sqrt{2\epsilon}/c_s$.

As a consequence of considering non-canonical kinetic terms, we have introduced a dynamically varying speed of sound. This means it is convenient to work in terms of the "sound-horizon" time $\mathrm{d}y = c_s\mathrm{d}\tau$ instead of τ [66]. For the case relevant here, i.e. $\eta = \eta_s = 0$ or (equivalently) requiring $\dot{\epsilon} = \dot{\epsilon_s} = 0$, one finds

$$y = \frac{c_s}{(\epsilon + \epsilon_s - 1)aH}\,. \tag{2.7}$$

In terms of y-time the quadratic action takes on the familiar form

$$S_2 = \frac{M_{\rm Pl}^2}{2}\int \mathrm{d}^3x\mathrm{d}y\; q^2\left[\zeta'^2 - (\vec{\nabla}\zeta)^2\right] \qquad , \qquad q \equiv \sqrt{c_s}z = \frac{a\sqrt{2\epsilon}}{\sqrt{c_s}}\,, \tag{2.8}$$

where $' \equiv \mathrm{d}/\mathrm{d}y$. It is useful to list the behavior in proper time τ and in y-time of some relevant quantities for later reference:

$$\begin{aligned} a &\sim (-\tau)^{\frac{1}{\epsilon-1}} \quad ; \quad c_s \sim (-\tau)^{\frac{\epsilon_s}{\epsilon-1}} \quad ; \quad H \sim (-\tau)^{\frac{-\epsilon}{\epsilon-1}}, \\ a &\sim (-y)^{\frac{1}{\epsilon_s+\epsilon-1}} \quad ; \quad c_s \sim (-y)^{\frac{\epsilon_s}{\epsilon_s+\epsilon-1}} \quad ; \quad H \sim (-y)^{\frac{-\epsilon}{\epsilon_s+\epsilon-1}}. \end{aligned} \tag{2.9}$$

Upon quantization the perturbations are expressed through creation and annihilation operators as follows,

$$\zeta(y,\mathbf{k}) = u_k(y)a(\mathbf{k}) + u_k^*(y)a^\dagger(-\mathbf{k}) \qquad , \qquad [a(\mathbf{k}), a^\dagger(\mathbf{k}')] = (2\pi)^3\delta^3(\mathbf{k}-\mathbf{k}'). \tag{2.10}$$

In terms of the canonically-normalized scalar variable $v = M_{\rm Pl}q\zeta$ the equations of motion for the Fourier modes are

$$v_k'' + \left(k^2 - \frac{q''}{q}\right)v_k = 0\,. \tag{2.11}$$

If $\epsilon, \eta, \epsilon_s, \eta_s$ are constant, as is the case here, one can solve for q''/q exactly [76]:

$$\frac{q''}{q} = \frac{1}{y^2}\left(\nu^2 - \frac{1}{4}\right). \tag{2.12}$$

The precise solution for $v_k(y)$ depends on which initial vacuum is chosen. Typically this is the Bunch-Davis vacuum, which is defined to be the state annihilated by $a(\mathbf{k})$ as $t \to -\infty$. With this vacuum choice the solution for $v_k(y)$ is

$$v_k(y) = \frac{\sqrt{\pi}}{2}\sqrt{-y}\, H_\nu^{(1)}(-ky). \tag{2.13}$$

where $H_\nu^{(1)}$ are Hankel functions of the first kind. However, we note that different adiabatic vacuum choices are possible in principle and also have distinctive observational signatures (particularly non-Gaussianities) [77, 78, 79, 80]. For the modes defined in (2.10) one finds

$$u_k(y) = \frac{c_s^{1/2}}{aM_{\rm Pl}\sqrt{2\epsilon}} v_k(y) = \frac{c_s^{1/2}}{aM_{\rm Pl}2^{3/2}}\sqrt{\frac{\pi}{\epsilon}}\sqrt{-y}H_\nu^{(1)}(-ky). \tag{2.14}$$

Therefore the 2-point correlation function for ζ is:

$$\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\rangle = (2\pi)^5\delta^3(\mathbf{k}_1+\mathbf{k}_2)\frac{P_\zeta}{2k_1^3}\,, \tag{2.15}$$

where the expression for the power spectrum P_ζ is[1]

$$P_\zeta \equiv \frac{1}{2\pi^2}k^3\,|\zeta_k|^2 = \frac{(\epsilon_s+\epsilon-1)^2\,2^{2\nu-3}}{2(2\pi)^2\epsilon}\frac{\bar{H}^2}{\bar{c}_s M_{\rm Pl}^2}\,. \tag{2.16}$$

Note that we have approximated the propagator $u_k(y)$ assuming $n_s - 1 \ll 1$ in accordance with observations here (for details see appendix C - a more general expression for P_ζ can be found in [76]). The bar symbol means that the corresponding quantity has to be evaluated, for each mode k, at sound horizon exit, i.e. when $y = k^{-1}$. The spectral index n_s is finally given by

$$n_s - 1 \equiv \frac{d\ln P_\zeta}{d\ln k} = 3 - 2\nu = \frac{2\epsilon+\epsilon_s}{\epsilon_s+\epsilon-1}\,. \tag{2.17}$$

Therefore exact scale invariance obtains when $n_s = 1$, i.e. $\epsilon_s = -2\epsilon$. Constraints on n_s consequently translate into relational constraints on ϵ and ϵ_s. Importantly this means no direct bound on either ϵ or ϵ_s can be obtained in this way, because for any parameter choice ϵ one can find an appropriate ϵ_s to yield a (near)scale-invariant solution. This is in contrast to the case of canonical fields, where the smallness of $n_s - 1$ requires ϵ to be small. Note that, if we drop our

[1] This can be evaluated straightforwardly according to the prescription outlined in appendix C.

assumption $\eta = 0$ and consider a fully general canonical field, then the constraint from $n_s - 1$ maps to a more complicated constraint on the full hierarchy of slow-roll parameters associated with derivatives of H. Generalizing to a general non-canonical field theory and dropping η_s then opens up the new hierarchy of slow-roll parameters associated with derivatives of c_s. The salient point in all of these cases is that, when considering non-canonical theories additional degrees of freedom enter and hence the need for additional ways of measuring/fixing these new parameters becomes more pressing. In what follows we will use the tensor-to-scalar ratio r and three-point correlation functions for ζ to probe the non-canonical theories in question and mostly focus on the case of an exactly scale-invariant 2-point function $n_s = 1$ following [66], requiring $\epsilon_s = -2\epsilon$ and $\eta, \eta_s = 0$. $(n_s - 1)$-dependent corrections in this setting are considered in appendix D.

2.2.2 The tensor-to-scalar ratio r

In inflation [10, 11, 12, 13] primordial quantum tensor fluctuations are sourced on sub-horizon scales before leaving the horizon and freezing out in analogy to scalar fluctuations. Present observational upper bounds on the tensor-to-scalar ratio r therefore constrain the amount of gravity waves that could have been generated in the early universe. Here we use bounds on r to constrain slow-roll parameter ϵ. Note that non-inflationary bimetric models of the primordial epoch, as discussed in chapter 3 do not resolve the horizon problem for gravitational waves [23]. As no quantum tensor fluctuations are amplified to become classical outside the horizon one does not expect any significant primordial gravitational wave contribution for these models. Such models are therefore typically assumed to give $r \approx 0$ as no explicit mechanism to generate tensor perturbations exists for them/footnoteIn addition this of course relies on imposing appropriate initial conditions such that no significant tensor fluctuations persist to the present era..

The action governing tensor modes is [76]

$$S = \frac{1}{2}\int d^3\mathbf{k}\sum_{n=1}^{2}\int\left(\left|\tilde{v}'_{\mathbf{k},n}\right|^2 - \left(k^2 - \frac{a''}{a}\right)|\tilde{v}_{\mathbf{k},n}|^2\right)d\eta \tag{2.18}$$

where the sum runs over the two distinct polarization modes and we have introduced $\tilde{v}$ to distinguish between variables $\tilde{v}$ and v associated with tensor and scalar perturbations respectively. The action is written in terms of a Mukhanov-Sasaki variable $\tilde{v}_{\mathbf{k},n} = \frac{a(\eta)M_{Pl}}{2}h_{\mathbf{k},n}$, where a is the scale factor of the FRW metric.

Referring back to our discussion of curvature perturbations in prologue II (see in particular (1.21)), tensor perturbations correspond to the the $\delta g_{ij} = h_{ij}$ component [39]. We express the associated Fourier modes as $h_{\mathbf{k},n}$ where as mentioned above the index n runs over the two independent polarization modes. Upon quantization $\tilde{v}$ can be expressed in terms of creation and

annihilation operators as

$$\hat{\tilde{v}}_{\mathbf{k},n} = \tilde{v}_k(\eta)\hat{\tilde{a}}_{\mathbf{k},n} + \tilde{v}_k^*(\eta)\hat{\tilde{a}}^\dagger_{\mathbf{k},n}. \tag{2.19}$$

The equation of motion for tensor modes can then be written as

$$\tilde{v}_k'' + \left(k^2 - \frac{a''}{a}\right)\tilde{v}_k = 0. \tag{2.20}$$

Accordingly we can now write down the power spectrum of tensor perturbations, where the sum again runs over polarization modes

$$P_h = \frac{k^3}{2\pi^2}\sum_n^2 \langle |h_{\mathbf{k},n}|^2\rangle, \tag{2.21}$$

which can be found to be [39, 76]

$$P_h = 2^{2\mu-3}\left|\frac{\Gamma(\mu)}{\Gamma(3/2)}\right|^2 (1-\epsilon)^{2\mu-1}\left|\frac{\sqrt{2}H}{\pi M_{pl}}\right|^2_{k=aH}, \tag{2.22}$$

where

$$\mu = \frac{1}{1-\epsilon} + \frac{1}{2}. \tag{2.23}$$

Importantly the tensor power spectrum is evaluated at horizon crossing for the tensor modes. For general theories with a non-canonical kinetic term (2.1) this horizon is different from the corresponding horizon for scalar perturbations. The spectral index of tensor perturbations n_t then is

$$n_t = \frac{-2\epsilon}{1-\epsilon}. \tag{2.24}$$

The tensor-to-scalar ratio r is given by

$$r \approx 2^{2\mu-3}\frac{(1-\epsilon)^{2\mu-1}}{(\epsilon_s+\epsilon-1)^2}\left|\frac{\Gamma(\mu)}{\Gamma\left(\frac{3}{2}\right)}\right|^2 16c_s(k_\zeta)\epsilon\left(\frac{H_h}{H_\zeta}\right)^2 \tag{2.25}$$

where c_s has to be evaluated at scalar horizon crossing and H_h and H_ζ correspond to the Hubble factor at freeze-out for tensor and scalar modes respectively [39].[2] Note that the dependence on the ratio of H at different horizon crossings can induce a strong scale dependence for r.

Present-day observational constraints limit $r \lesssim 0.3$. The WMAP 95 % confidence level bound is $r < 0.24$ [9].[3] For a given value of n_s this can therefore be used to constrain $\epsilon c_s(k_\zeta)$. Ref [81]

[2] As a comparison with (2.16) shows, we have assumed $n_s - 1 = 3 - 2\nu \ll 1$ here, i.e. near-scale-invariance.

[3] This bound is obtained assuming $n_t = \frac{-r}{8}$, however. The relation between n_t and r is of course more involved for general single field models as a comparison between (2.24) and (2.25) shows. However, given that we obtain a red tilt of tensor modes (and P_h therefore peaks on the largest scales) the bound on r should only weakly depend

perform an MCMC sampling for general single field models (2.1) to find $\epsilon c_s < 0.023$ at 2σ for a flat prior (note that the result is somewhat prior dependent).

For a horizon crossing sound speed of $\bar{c}_s = 0.1$ this corresponds to $\epsilon \lesssim 0.23$ which agrees with the approximate bound derived by [66]. However, as shown by [82] present observational limits for the speed of sound only place a much weaker bound of $\bar{c}_s \gtrsim 0.011$ at 2σ on generic single field inflation models[4]. In a more model-specific context this bound can be strengthened, e.g. for DBI-inflation we find $\bar{c}_s \gtrsim 0.034$, yielding $\epsilon \lesssim 0.68$. If future observations can improve bounds for generic single field models even further, e.g. $\bar{c}_s \gtrsim 0.05$, this bound on ϵ can be improved to $\epsilon \lesssim 0.4$. However, as we will see in principle far stronger constraints can be obtained from the running with scale of the three-point function.

2.3 Observational constraints on (non)-slow-roll

As discussed in the introduction, large classes of both inflationary and bimetric models can be described by action (2.1). Bimetric models generically have $\epsilon > 1$, necessitating expressions for observable n-point correlation functions which do not assume the smallness of slow-roll parameters. Such models are discussed in detail in chapter 3. For the inflationary branch $\epsilon < 1$, however, observations can be used to constrain the allowed values of slow-roll parameters ϵ, ϵ_s. Establishing bounds on inflationary physics in this way as well as disclosing generic phenomenology in fast-roll inflation is our main aim in this chapter. Further associated phenomenology plus constraints are derived in the following sections and we summarize the resulting bounds upfront here. The observational bounds considered are the following:

- **The spectral index n_s:** The spectral index of scalar perturbations n_s is computed in section 2.2.1. In terms of the slow-roll parameters it can be written as

$$n_s - 1 = \frac{2\epsilon + \epsilon_s}{\epsilon_s + \epsilon - 1} \,. \tag{2.26}$$

  The WMAP bounds on n_s in the presence of tensor perturbations are $n_s = 0.973 \pm 0.028$ at 2σ. Note that these constraints are dependent on e.g. the precise value of H_0 [9] and the physics of reionization [83]. Contrary to the canonical field case, for a general single field model (2.1) the smallness of $n_s - 1 \ll 1$ no longer automatically translates into a constraint on ϵ alone. Instead, for any parameter ϵ, one can find a corresponding speed of sound profile parameterized by ϵ_s, which returns any desired spectral index. Bounds on

on the precise value of the tilt.

[4]Note that these bounds come from non-Gaussian amplitudes, which we will discuss later in this chapter. Importantly they also ignore the fast-roll suppression will find for such amplitudes, so that the actual bound which arises when fully taking fast-roll effects into consideration is expected to be significantly weaker still.

n_s therefore only translate into constraints on the relation between ϵ and ϵ_s, not on either variable by itself.

- **The tensor-to-scalar ratio r:** The tensor-to-scalar ratio r is similarly constrained by CMB observations. The WMAP experiment provides a bound $r < 0.24$ at 2σ [9]. As we have seen above in section 2.2.2, the first order relation $r = 16\epsilon c_s$ is modified in general single field models without slow-roll and one can compute corresponding bounds on slow-roll parameter ϵ. For DBI-type models we find $\epsilon \lesssim 0.68$.

- **The running n_{NG}:** Introducing non-canonical kinetic terms can lead to a strong running of non-Gaussianities with scale n_{NG} [66, 84, 85, 86], irrespective of whether the 2-point correlation function as measured by n_s is near-scale-invariant or not. This can impose further constraints on ϵ. If non-Gaussianities show a blue tilt, this means they peak on smaller scales. For a perturbative treatment to be applicable, one therefore needs to ensure that non-Gaussianities on the smallest observable scales are not too large for a perturbative approach to break down. We discuss this in detail in section 2.4.4. There we also identify under which conditions non-Gaussianities are blue tilted and what constraints on ϵ follow. We show that constraints are in principle model-dependent. However, for a very large subclass of models one can obtain rough bounds

$$\epsilon \lesssim 0.2 \quad \text{if} \quad f_{\rm NL}(\text{CMB}) \sim \mathcal{O}(100) \tag{2.27}$$

$$\epsilon \lesssim 0.3 \quad \text{if} \quad f_{\rm NL}(\text{CMB}) \sim \mathcal{O}(10) \tag{2.28}$$

$$\epsilon \lesssim 0.5 \quad \text{if} \quad f_{\rm NL}(\text{CMB}) \sim \mathcal{O}(1). \tag{2.29}$$

  This result agrees with bounds obtained in the exactly scale invariant case [66]. However, we show that these limits are somewhat dependent on parameters c_s, Σ, λ.

- **The Big Bang problems:** Observational bounds only place constraints on primordial perturbations for a small range of scales: CMB ($\sim 10^3$Mpc) to galactic scales ($\sim$ 1Mpc). n_s,r and n_{NG} as discussed above are therefore only observationally constrained over these scales.

  In an inflationary setup this corresponds to approximately 10 e-folds, where the number of e-folds N of inflation is given by $dN = -Hdt$. In order to resolve the flatness, homogeneity and isotropy problems of cosmology, a greater number of e-folds is needed. Conventionally this is chosen $N \sim 60$, although the minimal amount of e-folds needed is not very well defined and depends e.g. on the energy scale of inflation, details of the reheating mechanism, etc. [87] The inflationary condition $\epsilon < 1$ is sufficient to yield an attractor solution towards flatness, homogeneity and isotropy [66]. A large number of e-folds requires $\epsilon < 1$

for several Hubble times, which is controlled by parameter η. As we consider cases for which $\eta \sim 0 \ll 1$, the fractional change in ϵ per Hubble time is small and hence starting with $\epsilon < 1$ guarantees that a prolonged phase of inflation takes place.

However, we should note that several other solutions are possible as well. For example, the scaling phase responsible for near-scale-invariant perturbations on observable scales may be followed by a different inflationary phase (scaling or not) [66]. Particle physics may suggest that inflation is actually a multiple step process, where one relatively short initial phase is responsible for observable structure on very large scales [88]. Inflation then ends and is followed by (an)other phase(s) of accelerated expansion later. It could also be that an altogether different mechanism[5] resolves the problems in question here [46, 47, 48]. In any case no competitive constraints on ϵ can be derived from consideration of these cosmological problems. Also note that bimetric models resolve these problems in a radically different way [46, 47, 48].

In summary, if near-future CMB experiments discover observable non-Gaussianities, then inflationary models of the type envisaged by (2.1), with $\epsilon \lesssim 0.3$, meet all the observational bounds. In the absence of observable non-Gaussianities, bounds on the tensor-to-scalar ratio r are likely to provide the strongest constraints on ϵ. We aim to explore such bounds further in the future.

2.4 Non-Gaussianity

In this section we compute the 3-point correlation function for the curvature perturbation ζ. The 3-point function then measures the strength of interactions of the field[6] which are described by the interaction vertices in the cubic action. Perturbing (2.1) to cubic order in ζ one obtains this effective action as derived in [37, 80]. The result is valid outside of the slow-roll approximation and for any time-dependent sound speed:

$$\begin{aligned} S_3 &= {M_{Pl}}^2 \int \mathrm{d}t\mathrm{d}^3x \left\{ -a^3 \left[\Sigma \left(1 - \frac{1}{c_s^2} \right) + 2\lambda \right] \frac{\dot{\zeta}^3}{H^3} + \frac{a^3\epsilon}{c_s^4}(\epsilon - 3 + 3c_s^2)\zeta\dot{\zeta}^2 \right. \\ &+ \frac{a\epsilon}{c_s^2}(\epsilon - 2\epsilon_s + 1 - c_s^2)\zeta(\partial\zeta)^2 - 2a\frac{\epsilon}{c_s^2}\dot{\zeta}(\partial\zeta)(\partial\chi) \\ &+ \left. \frac{a^3\epsilon}{2c_s^2}\frac{d}{dt}\left(\frac{\eta}{c_s^2}\right)\zeta^2\dot{\zeta} + \frac{\epsilon}{2a}(\partial\zeta)(\partial\chi)\partial^2\chi + \frac{\epsilon}{4a}(\partial^2\zeta)(\partial\chi)^2 + 2f(\zeta)\left.\frac{\delta L}{\delta\zeta}\right|_1 \right\} , \end{aligned} \tag{2.30}$$

[5] And of course, as the flatness, homogeneity and isotropy problems are essentially initial value problems, the correct initial values may be set by pre-inflationary physics. Whilst such a solution may be very non-definitive and as such less appealing, one should keep in mind that there is certainly no guarantee that all cosmological puzzles can be answered at energy scales we have (in-)direct access to.

[6] A free field is Gaussian, hence the 3-point function captures the non-Gaussian statistics for ζ.

where dots denote derivatives with respect to proper time t, ∂ is a spatial derivative, and χ is defined as

$$\partial^2\chi = \frac{a^2\epsilon}{c_s^2}\dot{\zeta}\,. \tag{2.31}$$

Note that the final term in the cubic action, $2f(\zeta)\frac{\delta L}{\delta\zeta}|_1$ is proportional to the linearized equations of motion and can be absorbed by a field redefinition $\zeta \to \zeta_n + f(\zeta_n)$ - for details see appendix C. In order to aid comparison with the literature note that we may also express the cubic action in an even simpler form [37, 67]

$$S_3 = \int \mathrm{d}^3x\,\mathrm{d}\tau\; a^2\left\{\frac{\Lambda_1}{a}\zeta'^3 + \Lambda_2\zeta\zeta'^2 + \Lambda_3\zeta(\partial\zeta)^2 + \Lambda_4\zeta'\partial_j\zeta\partial_j\partial^{-2}\zeta' + \Lambda_5\partial^2\zeta(\partial_j\partial^{-2}\zeta')(\partial_j\partial^{-2}\zeta')\right\}. \tag{2.32}$$

At first order in perturbation theory and in the interaction picture, the 3-point function is given by [89, 37, 80]

$$\langle\zeta(t,\mathbf{k}_1)\zeta(t,\mathbf{k}_2)\zeta(t,\mathbf{k}_3)\rangle = -i\int_{t_0}^{t}\mathrm{d}t'\langle[\zeta(t,\mathbf{k}_1)\zeta(t,\mathbf{k}_2)\zeta(t,\mathbf{k}_3), H_{\rm int}(t')]\rangle\ , \tag{2.33}$$

where $H_{\rm int}$ is the Hamiltonian evaluated at third order in the perturbations and follows from (2.30). Vacuum expectation values are evaluated with respect to the interacting vacuum $|\Omega\rangle$.

The 3-point correlation function is conventionally expressed through the amplitude $\mathcal{A}$

$$\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle = (2\pi)^7\delta^3(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)P_\zeta{}^2\frac{1}{\Pi_j k_j^3}\mathcal{A}\,. \tag{2.34}$$

Also by convention the power spectrum P_ζ in the above formula is calculated for the mode $K = k_1 + k_2 + k_3$. By using (2.10) we can now calculate the amplitudes for each term appearing in the action (2.30). Using a self-evident notation the overall amplitude is therefore given by

$$\mathcal{A} = \mathcal{A}_{\dot\zeta^3} + \mathcal{A}_{\zeta\dot\zeta^2} + \mathcal{A}_{\zeta(\partial\zeta)^2} + \mathcal{A}_{\dot\zeta\partial\zeta\partial\chi} + \mathcal{A}_{\epsilon^2}, \tag{2.35}$$

where $\mathcal{A}_{\epsilon^2}$ accounts for the $\partial\zeta\partial\chi\partial^2\chi$ and $(\partial^2\zeta)(\partial\chi)^2$ terms in the action (2.30).

In the following subsections we first present the main result (Sec. 2.4.1), which is a general expression for $\mathcal{A}$ where ϵ is *not* constrained to satisfy $\epsilon \ll 1$ are given. We then study three physically interesting properties of these three-point functions: The size of the amplitude as parameterized by the variable f_{NL} (Sec. 2.4.2), its shape (Sec. 2.4.3) and the running of the non-Gaussianity n_{NG} (Sec. 2.4.4). Current observational bounds are also discussed for these properties.

2.4.1 The full amplitudes

We first present expressions for the full non-Gaussian amplitude $\mathcal{A}$ following [66]. Details of the calculation are given in appendix C. We stress that these expressions do not assume slow-roll conditions for ϵ, ϵ_s and are therefore valid even when these parameters become $\epsilon, \epsilon_s \sim \mathcal{O}(1)$. Here we focus on results for models with an exactly scale-invariant 2-point function ($n_s - 1 = 0$). $\mathcal{O}(n_s - 1)$ corrections are discussed in appendix D - also see [3, 1, 67, 69]. Interestingly, as we will see in the following subsections, it is precisely the part of the amplitude that vanishes in DBI type models that is instrumental for much of the new phenomenology we find for general slow-roll violating single field models (2.1).

In what follows it will be useful to define the variable α, satisfying

$$\alpha = \frac{2\epsilon - \epsilon_s}{\epsilon_s + \epsilon - 1} = -\frac{4\epsilon}{1+\epsilon}, \tag{2.36}$$

where the second equality holds in the scale-invariant limit $\epsilon_s = -2\epsilon$. Also note that we can relate Σ, λ by introducing a new parameter f_X [37, 66]

$$\lambda = \frac{\Sigma}{6}\left(\frac{2f_X - 1}{c_s^2} - 1\right) \tag{2.37}$$

where

$$f_X = \frac{\epsilon\epsilon_s}{3\epsilon_X} \qquad , \qquad \epsilon_X = -\frac{\dot{X}}{H^2}\frac{\partial H}{\partial X}, \tag{2.38}$$

and we note that there is in principle no dynamical requirement fixing ϵ_X (and hence f_X) to be small or large, even in the presence of constraints for ϵ [37]. As we will see the size of non-Gaussianities can then be expressed as a function of the parameters $\{c_s^{-2}, f_X, \epsilon, n_s\}$, where f_X essentially measures the strength of the first interaction vertex ζ'^3. Relating it back to the cubic action (2.32) it satisfies [67]

$$\Lambda_1 = \frac{2\epsilon}{3Hc_s^4}\left(1 - c_s^2 - f_X\right). \tag{2.39}$$

For DBI models (discussed in 2.5.1) $f_X^{\rm DBI} = 1 - c_s^2$, so consequently the first term in the action (2.30) vanishes and there is no $\mathcal{A}_{\dot{\zeta}^3}$ contribution to the amplitude, i.e. $\Lambda_1 = 0$. For the computation of $\mathcal{A}_{\dot{\zeta}^3}$ we otherwise assume for simplicity that f_X (and hence the kinetic part of ϵ, ϵ_X) is constant.

The non-Gaussian amplitudes $\mathcal{A}$ are then given by [66]

$$\begin{aligned}
\mathcal{A}_{\dot\zeta^3} &= -\frac{1+\epsilon}{2\bar c_s^2}\left[2\bar c_s^2+(f_X-1)\cos\frac{\alpha\pi}{2}\Gamma(3+\alpha)\right]\frac{k_1^2k_2^2k_3^2}{K^3}\ ;\\
\mathcal{A}_{\zeta\dot\zeta^2} &= \frac{1}{4\bar c_s^2}\left\{\left[6\bar c_s^2+(\epsilon-3)\cos\frac{\alpha\pi}{2}(2+\alpha)\Gamma(1+\alpha)\right]\frac{1}{K}\sum_{i<j}k_i^2k_j^2+\right.\\
&\quad\left.-\left[3\bar c_s^2+(\epsilon-3)\cos\frac{\alpha\pi}{2}\Gamma(2+\alpha)\right]\frac{1}{K^2}\sum_{i\neq j}k_i^2k_j^3\right\}\ ;\\
\mathcal{A}_{\zeta(\partial\zeta)^2} &= \frac{(1+5\epsilon)}{8\bar c_s^2}\cos\frac{\alpha\pi}{2}\Gamma(1+\alpha)\left[\frac{1+\alpha}{1-\alpha}\sum_j k_j^3+\frac{2-\alpha^2}{1-\alpha}\frac{2}{K}\sum_{i<j}k_i^2k_j^2\right.\\
&\quad\left.-2(1+\alpha)\frac{1}{K^2}\sum_{i\neq j}k_i^2k_j^3-\frac{\alpha}{K(1-\alpha)}\sum_i k_i^4+\alpha\frac{1+\alpha}{1-\alpha}k_1k_2k_3\right]\\
&\quad-\frac{1}{8}\left[\sum_j k_j^3+\frac{4}{K}\sum_{i<j}k_i^2k_j^2-\frac{2}{K^2}\sum_{i\neq j}k_i^2k_j^3\right]\ ;\\
\mathcal{A}_{\dot\zeta\partial\zeta\partial\chi} &= -\frac{\epsilon}{4\bar c_s^2}\cos\frac{\alpha\pi}{2}\Gamma(1+\alpha)\left[\sum_j k_j^3+\frac{\alpha-1}{2}\sum_{i\neq j}k_ik_j^2-2\frac{1+\alpha}{K^2}\sum_{i\neq j}k_i^2k_j^3-2\alpha k_1k_2k_3\right]\ ;\\
\mathcal{A}_{\epsilon^2} &= \frac{\epsilon^2}{16\bar c_s^2}\cos\frac{\alpha\pi}{2}\Gamma(1+\alpha)(2+\alpha/2)\left[\sum_j k_j^3-\sum_{i\neq j}k_ik_j^2+2k_1k_2k_3\right]\ ,
\end{aligned}\tag{2.40}$$

where $\mathcal{A}_{\epsilon^2}$ accounts for the $\partial\zeta\partial\chi\partial^2\chi$ and $(\partial^2\zeta)(\partial\chi)^2$ terms.

2.4.2 f_{NL}

The size of the three-point correlation function is usually quoted using parameter $f_{\rm NL}$. Here we focus on the equilateral limit $k_1 = k_2 = k_3 = K/3$ and define $f_{\rm NL}^{equi}$ in terms of the amplitudes $\mathcal{A}$ [66]:

$$f_{\rm NL}^{equi} = 30\frac{\mathcal{A}_{k_1=k_2=k_3}}{K^3}\,,\tag{2.41}$$

where we have matched amplitudes at $k_1 = k_2 = k_3 = K/3$, i.e. in the equilateral limit. Defined in this way $f_{\rm NL}^{equi}$ is essentially the value of the amplitude $\mathcal{A}$ at a particular point in k-space and serves as a convenient single number measure of the amplitude of non-Gaussianity in the equilateral limit $k_1 = k_2 = k_3$

Note that we follow the WMAP sign convention here, where positive $f_{\rm NL}$ physically corresponds to negative-skewness for the temperature fluctuations. In the literature expressions for $f_{\rm NL}^{equi}$ for

a general theory with action (2.1) has been investigated in detail for slowly varying c_s and to first order in slow-variation parameters [80]. For (potentially) rapidly varying c_s when the 2-point correlation function is exactly scale invariant only a small section of the parameter space has been explored so far [66] (namely only cases assuming $f_X \in \{0, 1\}$). Here we generalize these results by exploring the full observationally allowed parameter-space to arbitrary order in slow-variation parameter ϵ, focusing on the novel effects on $f_{\rm NL}^{equi}$ due to the departure from slow-roll.

Since requiring a scale-invariant 2-point function yields $\epsilon_s = -2\epsilon$, fast-roll models with $\epsilon \sim \mathcal{O}(1)$ lead to a rapidly decreasing c_s (as long as inflation is not ghost-like, i.e. $\epsilon \not< 0$). They therefore naturally yield regimes where c_s is small and the 3-point function is large. A useful fitting formula illustrating this point in this context is [66]

$$f_{\rm NL}^{equi} = 0.27 - \frac{0.164}{\bar{c}_s^2} - (0.12 + 0.04 f_X)\frac{1}{\bar{c}_s^2}\left(1 - \frac{4\epsilon}{1+\epsilon}\right). \tag{2.42}$$

In order to establish a connection between different conventions we remind ourselves that f_X can be expressed in terms of λ/Σ [37], given that

$$f_X = \frac{1}{2}\left(\left(6\frac{\lambda}{\Sigma} + 1\right)c_s^2 - 1\right). \tag{2.43}$$

It follows that $|f_X| \gg 1$ in the presence of $c_s^2 \leq 1$ entails $\left|\frac{\lambda}{\Sigma}\right| \gg 1$. From (2.42) a good order of magnitude estimate for models with a large amount of non-Gaussianity as measured by $f_{\rm NL}^{equi} \gg 1$ is

$$f_{\rm NL}^{equi} \sim \mathcal{O}(c_s^{-2}) + \mathcal{O}(\frac{\lambda}{\Sigma}) \sim \mathcal{O}(c_s^{-2}) + \mathcal{O}(\frac{f_X}{c_s^2}). \tag{2.44}$$

Interestingly, this result corresponds to the case where fluctuations in the scalar field dominate over those in gravity [22, 90]. Equation (2.44) suggests that deviations from slow-roll and breaking of scale invariance alone are not sufficient to generate significant levels of non-Gaussianity by themselves.

However, departures from slow-roll can have a significant effect for non-Gaussianities of detectable size $f_{\rm NL}^{equi} \gtrsim \mathcal{O}(1)$. Note that, considering the parametric dependence of the amplitude, ϵ_s is fixed in terms of n_s, ϵ via (2.36). For a given n_s, an $f_{\rm NL}$ increasing with larger ϵ therefore corresponds to ϵ_s decreasing whilst $f_{\rm NL}$ increases and vice versa. The left graph in figure 2.1 show how considering non-slow-roll ϵ can lead to a suppression of $f_{\rm NL}^{equi}$. In fact, if ϵ is allowed to range to $\epsilon \gtrsim 0.3$ this can even result in a sign change in $f_{\rm NL}^{equi}$.[7]

Particularly instructive is the case when a large amplitude $\mathcal{A}$ is primarily due to $\lambda/\Sigma \gg 1$. Then, the exact expression for $f_{\rm NL}^{equi}$ simplifies considerably and we obtain (irrespective of whether

[7] We will see why we only consider $\epsilon \in \{0, 0.3\}$ here when we consider the dependence on scale k of the 3-point function itself.

c_s is large or small)

$$f_{\rm NL}^{f_X\gg 1} \sim -(1+\epsilon)\frac{f_X}{\bar{c}_s^2}\cos\left(\alpha\frac{\pi}{2}\right)\Gamma(3+\alpha). \tag{2.45}$$

At first sight a good estimate for $f_{\rm NL}^{equi}$ in this case is $f_{\rm NL}^{equi} \sim \mathcal{O}(f_X/c_s^2) \sim \mathcal{O}(\frac{\lambda}{\Sigma})$ in agreement with (2.44). However, the amplitude is suppressed by a factor of $\cos\frac{\alpha\pi}{2}\Gamma(3+\alpha)$. For $0 < \epsilon \lesssim 0.33$ this factor indeed always causes a suppression, which can easily reduce the size of the amplitude by more than an order of magnitude for $\epsilon \sim 0.3$, for example. As before, if ϵ is allowed to range all the way up to e.g. $\epsilon \sim 0.4$, this effect will lead to a change in sign for $f_{\rm NL}^{equi}$. Thus slow-roll models with an $f_{\rm NL}^{equi}$ that violate present bounds (see below) can be reconciled with observations when non-slow-roll effects are taken into consideration.

Observational constraints on $f_{\rm NL}^{equi}$ depend on which non-Gaussian shape is being probed [91, 92, 93], as discussed in the next subsection. Frequently these constraints are quoted for three shapes. Firstly, the equilateral shape, peaking in the limit $k_1 = k_2 = k_2$ and going to zero in the flat limit $k_1+k_2 = k_3$. This roughly corresponds to the shapes following from $\mathcal{A}_{\dot\zeta\partial\zeta\partial\chi}, \mathcal{A}_{\epsilon^2}$ and also approximately $\mathcal{A}_{\dot\zeta^3}$ (although for $\mathcal{A}_{\dot\zeta^3}$ there are small differences in the flat limit $k_1 + k_2 = k_3$). A second possibility is the local shape, which peaks in the squeezed limit $k_1 = k_2 \gg k_3$ and goes to zero in the equilateral limit. This is loosely associated with the shapes of $\mathcal{A}_{\zeta\dot\zeta^2}$ and $\mathcal{A}_{\zeta(\partial\zeta)^2}$. Finally, one may consider orthogonal/folded shapes. Here a suitable template shape is one which peaks in the folded configuration $2k_1 = 2k_2 = k_3$ and goes to zero in the equilateral and squeezed limits cf. [82, 94]. This corresponds to the difference between $\mathcal{A}_{\dot\zeta^3}$ and a more exact equilateral shape such as $\mathcal{A}_{\dot\zeta\partial\zeta\partial\chi}$ or $\mathcal{A}_{\epsilon^2}$. For a list of possible factorizable shape ansätze we refer to [95, 82, 77].[8] Current WMAP limits on $f_{\rm NL}$ for each of these shapes are [9]

$$\begin{aligned} f_{\rm NL}^{\rm equil} &= 26 \pm 240 \quad (95\%\ \text{CL}), \\ f_{\rm NL}^{\rm local} &= 32 \pm 42 \quad (95\%\ \text{CL}), \\ f_{\rm NL}^{\rm orth} &= -202 \pm 208 \quad (95\%\ \text{CL}). \end{aligned} \tag{2.46}$$

2.4.3 Shapes

In the previous section the information contained in $\mathcal{A}$ was condensed into a single parameter, $f_{\rm NL}^{equi}$. However, there is vastly more information available, encoded in the functional dependence of $\mathcal{A}$ [91, 92, 93]. This is uniquely specified by the four parameters $\epsilon, n_s, \bar{c}_s, f_X$.[9] In order to disentangle effects from a possible running of $\mathcal{A}$ with wavenumber K and the shapes themselves,

[8] A complete basis capable of describing arbitrary shapes in k-space in principle has infinitely many "shape-members". However, for the models considered here these three basis shapes will be sufficient.

[9] Note that there is an issue of choice here as we have six variables $\epsilon, \epsilon_s, n_s, \bar{c}_s, f_X, \lambda/\Sigma$ and two constraint equations (2.17) and (2.43). For instance specifying $\epsilon, \epsilon_s, \bar{c}_s, \lambda/\Sigma$ therefore also fixes the amplitude.

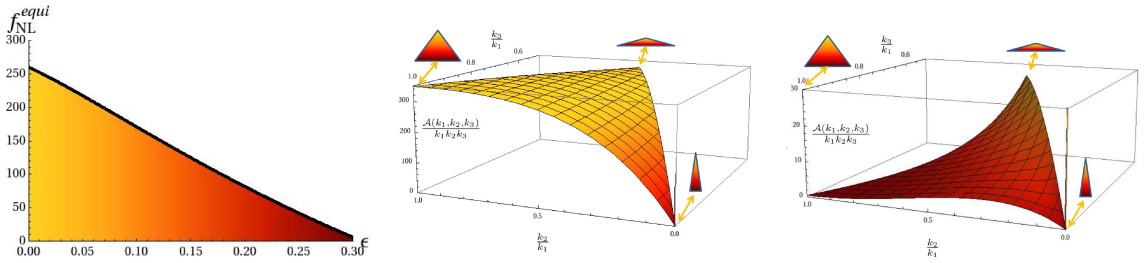


Figure 2.1: **Left**: $f_{\rm NL}^{equi}$ vs. ϵ for a horizon-crossing $\bar{c}_s = 0.1$ and $f_X = -70$. **Middle**: The dimensionless bispectrum $\mathcal{A}(k_1, k_2, k_3)/(k_1k_2k_3)$ plotted in the slow-roll limit $\epsilon \to 0$ for $\bar{c}_s = 0.07$ and $f_X = -53$. Triangular shapes denote the equilateral, enfolded and squeezed/local limit clockwise from top left. **Right**: Analogous plot for $\epsilon = 0.3$. Note how the overall amplitude is suppressed and the shape has changed, now peaking in the enfolded limit.

in this section we present the shape of the amplitude at fixed K. In interpreting the shapes shown in this section it may be useful to emphasize that from (2.41) it follows that a good estimate for the size $f_{\rm NL}$ corresponding to a given plotted shape (we normalize one wavevector k and hence plot $\mathcal{A}(1, k_1, k_2)/(k_1k_2)$) is the value of the plotted amplitude in the equilateral limit.

In the near future the shape and running of $\mathcal{A}$ will hopefully become experimentally accessible and, if a significant amount of non-Gaussianity is detected, will greatly constrain models of structure formation in the early universe. After all, at present these models are only observationally constrained to fit or respect upper limits on a number of single value parameters (The amplitude of scalar perturbations A, its spectral tilt n_s, $f_{\rm NL}$, the tensor-to-scalar ratio r). Fitting a full function would prove far more challenging and provide a strong model selection tool.

The prototypical shape that has been associated with significant non-Gaussianity in single-field models is one of the equilateral type [80, 66, 77, 96, 3], where the amplitude peaks in the limit $k_1 = k_2 = k_3$. If we fine-tune $\mathcal{A}_{\dot{\zeta}^3}$ to cancel out the dominant equilateral contribution of the other shape functions, a small region of parameter-space can also be shown to display dominant orthogonal or folded signals, i.e. non-Gaussian amplitudes peaking in in the limit [82]. A generic single field shape is therefore a superposition of equilateral and orthogonal/folded shapes, with the orthogonal/folded configuration typically being subdominant. This is in contrast to e.g. multifield models which can also give rise to strong local contributions to $\mathcal{A}$ (i.e. peaking in the squeezed limit $k_1 \approx k_2 \gg k_3$) [77] via superhorizon interactions local in position space. This is because superhorizon evolution is local in space, as different regions are not causally connected with each other. Therefore any amount of non-Gaussianity generated after horizon exit is also local in position space and hence not local in k-space. It consequently peaks in the squeezed limit. In single field models, on the other hand, ζ is frozen in upon horizon exit, so no superhorizon evolution takes place. Far inside the horizon modes typically oscillate and their contributions to $\mathcal{A}$ average out. For known exceptions to this see [78, 79, 80, 96, 97, 98]. Thus

$\mathcal{A}$ is almost completely sourced by modes exiting the horizon at similar times and hence with similar wavelengths. Consequently $\mathcal{A}$ peaks in the equilateral limit.

An interesting question to ask is therefore: Do deviations from slow-roll modify non-Gaussian shapes produced by single field models? And, if so, how? But before considering fast-roll scenarios, let us first understand the slow-roll limit a bit better. In this limit the cubic action (2.30) is dominated by its first three terms

$$\begin{aligned} S_{slow-roll} \;\sim\; & {M_{Pl}}^2 \int \mathrm{d}t\mathrm{d}^3x \left\{ -a^3 \left[\Sigma \left(1 - \frac{1}{c_s^2} \right) + 2\lambda \right] \frac{\dot{\zeta}^3}{H^3} \right. \\ & + \left. \frac{a^3\epsilon}{c_s^4}(\epsilon - 3 + 3c_s^2)\zeta\dot{\zeta}^2 + \frac{a\epsilon}{c_s^2}(\epsilon - 2\epsilon_s + 1 - c_s^2)\zeta(\partial\zeta)^2 + ... \right\} , \end{aligned} \tag{2.47}$$

This is visible from (D.15) and (D.16) as we can see there that the amplitudes corresponding to the remaining terms in the action are suppressed by powers of ϵ. The leading order contributions to $\mathcal{A}$ are therefore

$$\mathcal{A}_{slow-roll} \sim \mathcal{A}_{\dot{\zeta}^3} + \mathcal{A}_{\zeta\dot{\zeta}^2} + \mathcal{A}_{\zeta(\partial\zeta)^2}. \tag{2.48}$$

A breaking of scale-invariance will then introduce a local component into the amplitude in accordance with Maldacena's consistency relation for the squeezed/local limit[10]. This states that, in the absence of large subhorizon interactions, the very squeezed limit is dominated by a local shape contribution proportional to $n_s - 1$ [89, 99]

$$\langle \zeta_{k_1}\zeta_{k_2}\zeta_{k_3} \rangle \sim -(n_s - 1)\delta^3(\sum_i \mathbf{k}_i)\frac{P_\zeta(k_1)P_\zeta(k_3)}{4k_1^3k_3^3}. \tag{2.49}$$

The middle and bottom row of Figure 2.2 then show how considering non-canonical fields with $c_s \neq 1$ and/or $|f_X| \gg 1$ can amplify $\mathcal{A}$. As expected, the resulting shapes are predominantly equilateral. It is important to note that, while breaking of scale-invariance does induce a local contribution, this contribution only dominates the signal in the squeezed limit. The extreme squeezed limit is of course not observationally accessible, as it corresponds to considering modes with infinite wavelength. If the local contribution is sufficiently subdominant outside the extreme squeezed limit, as in the case plotted in the middle row of Figure 2.2, it is therefore possible that the observable section of the amplitude $\mathcal{A}$ is consistent with it being purely equilateral, despite there being a (small) local contribution.

What happens once we violate the slow-roll conditions and consider cases with $\epsilon \sim \mathcal{O}(1)$ consistent with constraints derived above? Firstly the remaining interaction terms in (2.30),$\mathcal{A}_{\dot{\zeta}\partial\zeta\partial\chi}$ and $\mathcal{A}_{\epsilon^2}$, which were previously suppressed by powers of ϵ, now become relevant. Even more

[10]This is also shown in the left column of Figure 2.2, where the amplitude is plotted for a slow-rolling near-canonical field (i.e. with $c_s \sim 1$ and $f_X \sim 0$) and spectral index $n_s = 0.96$. For the $\mathcal{O}(n_s - 1)$ corrections used to plot these graphs see appendix D

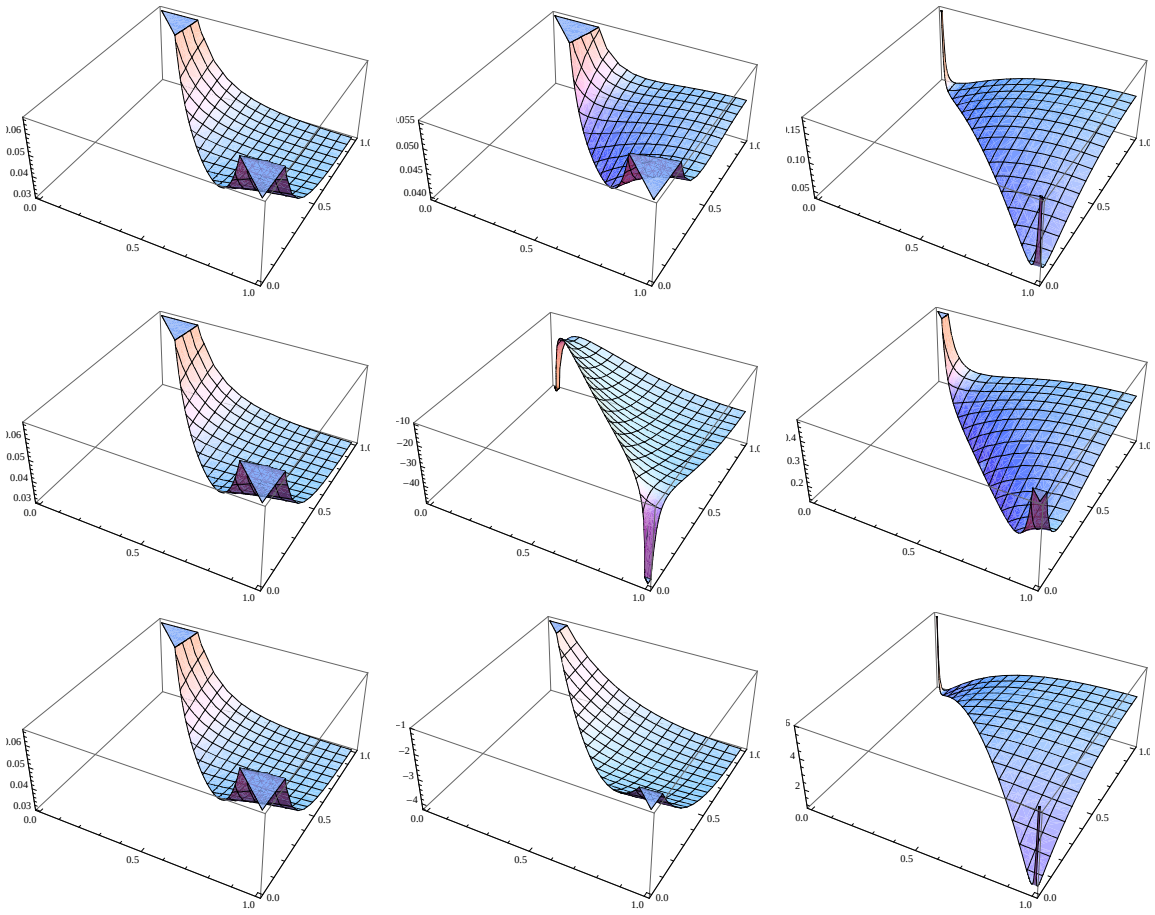


Figure 2.2: Here we plot the Non-Gaussian amplitude $\mathcal{A}(k_1, k_2, k_3)/(k_1 k_2 k_3)$ disentangling effects from $\epsilon, \bar{c_s}$ and f_X. Note that we are plotting a different section of k-space in comparison with figure 2.1. Axes correspond to k_2/k_1 and $k_3/k1$ just as in figure 2.1. Here the top right corner is the equilateral limit, the folded limit is at the center of the box, while the bottom right and top left corners both represent the squeezed limit - hence the symmetry of the plots. **Top row: The effect of varying ϵ.** The amplitude is plotted for $\epsilon = 0.01$, 0.1, 0.3 respectively from left to right. $\bar{c}_s = 1$ and $f_X = 0.01$. **Middle row: The effect of varying $\bar{c}_s$.** The amplitude is plotted for $\bar{c}_s = 1$, 0.1, 10 respectively from left to right. $\epsilon = 0.01$ and $f_X = 0.01$. **Bottom row: The effect of varying f_X.** $f_X = 0.01$, 100, -100 respectively from left to right. $\epsilon = 0.01$ and $\bar{c}_s = 1$.

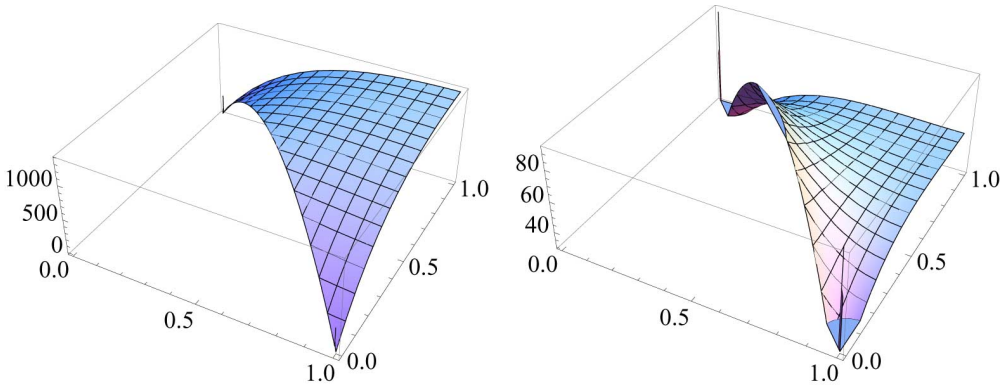


Figure 2.3: Here we plot the Non-Gaussian amplitude $\mathcal{A}(1, k_1, k_2)/(k_1 k_2)$ for $f_X = -100$ and $c_s = 0.05$. Axes label k_2/k_1 and $k_3/k1$ as before. Plots for the left and right columns are for $\epsilon = 0.001, 0.3$ respectively.

importantly the remaining terms receive large corrections from ϵ. This is most easily shown by considering the case when non-Gaussianities are primarily sourced by $f_X \gg 1$. In this case we have

$$\mathcal{A}(f_X \gg 1) \sim \mathcal{A}_{\dot{\zeta}^3} \sim -\frac{f_X - 1}{2\bar{c}_s^2}(\epsilon + 1)\cos\frac{\alpha\pi}{2}\Gamma(3+\alpha)\frac{k_1^2 k_2^2 k_3^2}{K^3}. \tag{2.50}$$

As we already saw in the previous section (2.45) this means the amplitude is suppressed by a factor of $\cos\frac{\alpha\pi}{2}\Gamma(3+\alpha)$. Considering violations of slow-roll can therefore lead to interesting new phenomenology, as previously suppressed contributions become important.

How can one understand the fast-roll suppression of the non-Gaussian amplitude as manifest in (2.50) physically? In essence one can separate the effect of considering large ϵ, ϵ_s on the amplitude into two categories. On the one hand the interaction Hamiltonian that follows directly from (2.30) is modified, since the "coupling constants" for individual interaction terms depend on ϵ, ϵ_s. The interaction terms one can ignore in the slow-roll limit (2.47) are all proportional to ϵ, so that their effect linearly increases upon breaking slow-roll. This explains why in the canonical case with ignorable f_x (as depicted in the left panel of figure 2.1) an increase in ϵ leads to an enhanced non-Gaussian signal. For in the canonical extreme slow-roll limit one is essentially considering a free field and interaction terms "switch on" as larger values of slow-roll parameters are considered.

The second effect of large ϵ, ϵ_s, which lies at the root of the fast-roll suppression of $\mathcal{A}$ discussed above, is a modification to the propagators (2.14). In other words the functional dependence of ζ on slow-roll parameters can then become important. The argument is analogous to the original one presented for the emergence of an equilateral shape for single field models. We recall that this argument invoked modes far inside the horizon oscillating so that their contributions to $\mathcal{A}$ average out. Considering non-slow-roll propagators can now have a similar effect. For the

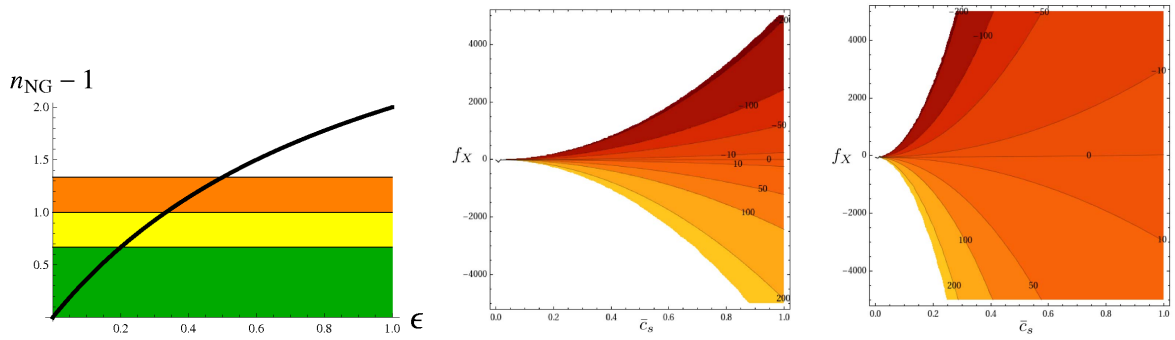


Figure 2.4: **Left**: $n_{\rm NG} - 1$ plotted against ϵ in the small c_s limit. Green, yellow and orange ($< 2/3, < 1, < 4/3$) regions are those allowed by perturbative constraints assuming $f_{\rm NL}^{equi}(\text{CMB}) \sim \mathcal{O}(100), \mathcal{O}(10), \mathcal{O}(1)$ respectively. **Middle**: Contour plot showing the region in parameter space allowed by the WMAP 2σ constraint $f_{\rm NL}^{\rm equil} = 26 \pm 240$ in the slow-roll limit $\epsilon \to 0$. **Right**: Analogous plot for $\epsilon = 0.3$. Note how the allowed region widens.

appearance of suppression factors as shown in equations (2.45) and (2.50) is essentially due to a "destructive interference" of modes. This means cancellations between contributions to the amplitude $\mathcal{A}$ can occur, since the oscillatory behaviour of ζ is modified.

Figures 2.1 and 2.3 show concrete case studies to illustrate our results. In particular they show the ϵ-suppression of otherwise dominant contributions. The left graph in figure 2.3 shows the amplitude for the given choice of parameters in the case of slow-roll. A very large amplitude with $f_{\rm NL} \sim 1000$ is produced, which violates present upper bounds on the level of non-Gaussianity. Deviating away from slow-roll has two important effects here. Firstly the fast-roll suppression moves the amplitude back within observational constraints. Therefore new regions in the parameter space $\{c_s, f_X\}$ open up. This is shown in figure 2.4. Secondly a delicate cancellation between terms can bring out contributions from otherwise suppressed orthogonal/folded shapes. This is shown by the $\epsilon \sim \mathcal{O}(1)$ amplitudes in Figure 2.3 peaking in the folded limit $2k_1 \approx 2k_2 \approx k_3$. A shape which cannot be decomposed into local and equilateral shapes alone, but requires a strong orthogonal/folded component.

Violation of slow-roll therefore naturally leads to the generation of intermediate shapes with equilateral, local (in the case of broken scale-invariance) and orthogonal/folded contributions. Hence fast-rolling single field models with $\epsilon \not\ll 1$ can produce a richer phenomenology than found in the extreme slow-roll limit, as fast-roll effects both lead to a suppression of the overall amplitude (opening up new regions of parameter-space) as well as altering the "shape" of non-Gaussian signals. And whilst general statements about limits of the three-point function for single field models, such as [89, 99] remain true, models beyond the slow-roll paradigm have a more complex fingerprint when considering the full amplitude.

2.4.4 The running n_{NG}

Having studied the shapes of non-Gaussianity at fixed K in the previous section, we now investigate the running of $\mathcal{A}$ with scale K [84, 85, 86]. Introducing a dynamical speed of sound can in principle lead to a strongly scale-dependent non-Gaussian amplitude $\mathcal{A}$. To see why we rewrite $\bar{c}_s$, the speed of sound at sound horizon crossing, in terms of the wavenumber K (which, we remind ourselves, is defined as $K = k_1 + k_2 + k_3$)

$$\bar{c}_s \sim K^{\frac{-\epsilon_s}{\epsilon_s+\epsilon-1}}. \tag{2.51}$$

Different scales K will therefore "see" a different $\bar{c}_s$ upon crossing the horizon. To quantify this difference we follow [84] in defining a spectral index for $f_{\rm NL}^{equi}$ as

$$n_{\rm NG} - 1 \equiv \mathrm{d}\ln|f_{\rm NL}^{equi}|/\mathrm{d}\ln K. \tag{2.52}$$

where we evaluate the running n_{NG} at a fixed point of the amplitude as measured by $f_{\rm NL}^{equi}$ to separate effects from the running and shape.

Expanding around the phenomenologically motivated small c_s limit, we find

$$n_{\rm NG} - 1 = \frac{4\epsilon}{1+\epsilon} + \frac{4\epsilon(8\epsilon-55)\mathrm{Sec}\left[\frac{2\epsilon\pi}{1+\epsilon}\right]c_s^2}{(55+8f_X+2\epsilon(15\epsilon+12\epsilon f_X-47-16f_X))\Gamma\left[\frac{4}{1+\epsilon}-3\right]} + \mathcal{O}(c_s^4), \tag{2.53}$$

which is an exact result in ϵ (the solution to all orders in c_s can be found in [1]). Interestingly this means we have a generically blue running of non-Gaussianities (as long as $\epsilon \not< 0$), resulting in larger interactions and hence enlarged signals on smaller scales. In other words, primordial non-Gaussianities measured on e.g. galaxy cluster scales would be larger than those measured at CMB scales. However, this also means interactions will eventually become strongly coupled for sufficiently small scales. Following [100] we take the ratio of cubic and quadratic Lagrangians as our measure of strong coupling, requiring

$$\frac{\mathcal{L}_3}{\mathcal{L}_2} \sim \mathcal{O}(1,\epsilon,f_X)\frac{\zeta}{c_s^2} \ll 1 \tag{2.54}$$

for fluctuations to be weakly coupled, roughly corresponding to $f_{\rm NL} \ll 10^5$. If this condition breaks (at horizon crossing, where n-point correlation functions are evaluated here), quantum loop corrections are no longer suppressed and a perturbative treatment is no longer applicable.

We now impose a minimal constraint of at least $\sim$ 10 e-folds of weakly coupled inflation governed by action (2.1), corresponding to the window of scales where primordial non-Gaussianity

may be observable.[11] In principle levels of non-Gaussianity are observable for all scales running from CMB scales ($k^{-1} \sim 10^3$Mpc) down to galactic scales ($k^{-1} \sim 1$Mpc) [66, 85, 101, 102]. In terms of the scale K this therefore corresponds to $K_{\rm gal}/K_{\rm CMB} \simeq 10^3$. For the non-Gaussian tilt this means

$$f_{\rm NL}^{equi}(\text{CMB}) \;\approx\; 10^{-3(n_{\rm NG}-1)} f_{\rm NL}^{equi}(\text{Gal})\,. \tag{2.55}$$

Blue/red tilted 3-point functions thus correspond to larger/smaller non-Gaussian amplitudes on smaller scales. As we found above, in order for fluctuations to be weakly coupled so that a perturbative treatment is applicable, the non-Gaussian contribution to ζ must be much smaller than the Gaussian part $f_{\rm NL}^{equi} \ll \zeta^{-1} \simeq 10^5$ for all observable scales. Present observational constraints on CMB scales give $f_{\rm NL}^{equi}(\text{CMB}) \lesssim \mathcal{O}(100)$. In case of a scale-invariant or red-tilted 3-point function, where the amplitude is largest for large scales, a perturbative treatment is valid all the way down to the smallest scales. However, for blue-tilted 3-point functions one can find interesting constraints on n_{NG} and hence on the slow-roll parameters. Specifically we have

$$\begin{aligned} n_{\rm NG} - 1 \lesssim 2/3 \quad &\text{for} \quad |f_{\rm NL}^{equi}(\text{CMB})| \sim \mathcal{O}(100) \\ n_{\rm NG} - 1 \lesssim 4/3 \quad &\text{for} \quad |f_{\rm NL}^{equi}(\text{CMB})| \sim \mathcal{O}(1). \end{aligned} \tag{2.56}$$

Depending on the size of $f_{\rm NL}^{equi}$ at CMB scales, this results in different bounds on $n_{\rm NG}$ as shown in figure 2.4. If the bound on $n_{\rm NG}$ is satisfied, non-Gaussian interactions remain under perturbative control throughout the range of observable scales. In the optimistic scenario with detectable CMB non-Gaussianities, i.e. $f_{\rm NL}^{equi}(\text{CMB}) \gtrsim \mathcal{O}(10)$, we can combine these constraints with equation (2.53) to put an upper bound on ϵ: $\epsilon \lesssim 0.3$. [66, 1] If $f_{\rm NL}^{equi}(\text{CMB}) \gtrsim \mathcal{O}(100)$ the bound is strengthened to $\epsilon \lesssim 0.2$. To cover the full range of models capable of generating interesting CMB non-Gaussianities, the explicit examples and plots we provide throughout the chapter are therefore chosen to satisfy bounds

$$\epsilon \lesssim 0.3 \qquad \text{if} \qquad f_{\rm NL}^{equi}(\text{CMB}) \gtrsim \mathcal{O}(10). \tag{2.57}$$

This shows how one can constrain the amount of slow-roll violation by requiring the action (2.1) to be a valid effective field theory over the observable range of scales for primordial fluctuations.[12]

Constraints from n_{NG} can therefore be used to put bounds on ϵ. These bounds can become important here, since considering non-slow-roll models naturally gives rise to a large running n_{NG} of non-Gaussianity with scale, which are consequently a smoking-gun signature of such fast-roll

[11] Beyond those 10 e-folds several options exist, depending on the UV-completion of the low-energy effective $P(X,\phi)$ theory [75, 45]: new degrees of freedom may become important, resulting in an inflationary weakly coupled multi-field theory, the dispersion relation may change, a strongly coupled phase of inflation may take place,...

[12] Note that eqn. (2.53) shows that these bounds receive $\mathcal{O}(c_s^2)$ corrections - see appendix D.4.

models. Again we reiterate that we use "fast-roll" as a description for an evolution which violates the "slow-variation" condition $\epsilon \ll 1$ and not just as a reference to properties of the potential $V(\phi)$ (see discussion in section 2.2). Note that in the event near-future experiments such as Planck do not observe any CMB non-Gaussianities, the bounds on ϵ from n_{NG} will become weaker and the strongest available bounds on ϵ may very well come from considerations of the tensor-to-scalar ratio r (see previous section).

2.5 Concrete model examples

In the previous sections we derived general expressions for the levels of non-Gaussianity exhibited by single field models 1.17 without assuming slow-roll and discussed their associated phenomenology. Here we consider two concrete model implementations, writing down explicit actions for each model and showing how departure from slow-roll and scale-invariance affects their non-Gaussian signatures. For single field models with large non-Gaussianity there are two interesting limits. From (2.44) we have

$$f_{\rm NL} \sim \mathcal{O}(c_s^{-2}) + \mathcal{O}(\frac{\lambda}{\Sigma}), \tag{2.58}$$

so that models with $c_s \ll 1$ and $|\lambda/\Sigma| \gg 1$ will lead to large non-Gaussianities. The two concrete examples we provide are therefore examples of models with these signatures: (**I**) DBI inflation ($c_s \ll 1$) and (**II**) Models with $\lambda/\Sigma \gg 1$.

2.5.1 Fast-roll DBI inflation

In DBI inflation [22] a 3 + 1 dimensional brane moves in warped extra dimensions, giving rise to an effective 4D theory with action (2.1), where

$$P(X,\phi) = -f^{-1}(\phi)\sqrt{1-2f(\phi)X} + f^{-1}(\phi) - V(\phi)\,. \tag{2.59}$$

$f(\phi)$ is the so-called warp factor and is constrained to be positive by the signature of the space (in this paper we use the convention $-+++$). We only consider the 4-D effective field theory defined by (2.59). If the full dynamics of a particular higher dimensional implementation are considered, far stronger constraints (e.g. from gravitational waves [103]) can be obtained. The speed of sound in DBI models is given by

$$c_s = \sqrt{1-2f(\phi)X} \tag{2.60}$$

which is the inverse of the Lorentz boost factor γ, so $c_s \ll 1$ since the inflaton rolls ultra-relativistically. Importantly one also finds

$$\lambda = \frac{H^2\epsilon}{2c_s^4}(1-c_s^2) \qquad , \qquad f_X^{\rm DBI} = 1 - c_s^2, \tag{2.61}$$

so that the otherwise typically dominant $\mathcal{A}_{\dot\zeta^3}$ contribution vanishes, and the amplitude consequently becomes independent of parameter f_X, resulting in:

$$f_{\rm NL}^{\rm DBI} \sim \mathcal{O}(c_s^{-2}). \tag{2.62}$$

DBI inflation is the leading example of a single field model with large non-Gaussianity to be found in the literature, and it is associated with an equilateral shape $\mathcal{A}$ [80, 77, 96]. This is the case since $\Lambda_1 = 0$ (cf. (2.32)) and hence no dominant orthogonal/folded non-Gaussianities can be produced here. From (2.40) we can therefore write

$$\mathcal{A}^{DBI} = \mathcal{A}_{\zeta\dot\zeta^2} + \mathcal{A}_{\zeta(\partial\zeta)^2} + \mathcal{A}_{\dot\zeta\partial\zeta\partial\chi} + \mathcal{A}_{\epsilon^2}, \tag{2.63}$$

i.e. DBI-type models are a simple subcase of the more general $P(X,\phi)$ setups considered above. As expected the equilateral shape can receive a small squeezed correction upon breaking of scale-invariance (cf. (D.15)), which becomes relevant in the squeezed limit. Also, just as for the more general single field case considered in section 2.4.3, we find that departure from slow-roll suppresses the overall amplitude.

2.5.2 Fast-roll inflation with $\left|\frac{\lambda}{\Sigma}\right| \gg 1$

We now build an explicit action for a model realizing $|\lambda/\Sigma| \gg 1$. First we consider a number of constraints that can be placed on the differential properties of $P(X,\phi)$. We can write slow-roll parameter ϵ as

$$\epsilon = \frac{3}{2 - \frac{P}{XP_{,X}}} \ . \tag{2.64}$$

In order to satisfy the null energy condition $p + \rho \geq 0$ and $1 \leq c_s^2 < 0$ one must require, respectively:

$$2XP_{,X} > 0. \qquad , \qquad P_{,XX} > 0, \tag{2.65}$$

as can be seen from the expression for the speed of sound

$$c_s^2 = \frac{P_{,X}}{\rho_{,X}} = \frac{P_{,X}}{P_{,X} + 2XP_{,XX}}. \tag{2.66}$$

Observational bounds on parameters such as n_s, r, n_t therefore only constrain the first two derivatives $XP_{,X}$ and $X^2P_{,XX}$ [104]. Expressing λ and Σ in terms of derivatives of P with respect to the canonical kinetic term X one finds

$$\frac{\lambda}{\Sigma} = \frac{1 + \frac{2}{3}X\frac{P_{,XXX}}{P_{,XX}}}{2 + \frac{P_{,X}}{XP_{,XX}}}, \tag{2.67}$$

so that in order for $|\lambda/\Sigma| \gg 1$ we require

$$\left|X^2P_{,XXX}\right| \gg |P_{,X}|\,. \tag{2.68}$$

Large non-Gaussianities sourced by λ/Σ (or equivalently f_X) thus open up the doors for constraining $P_{,XXX}$.

Adapting the scheme introduced by [104] we can therefore construct a general form for any action that gives rise to a large non-Gaussian amplitude $\mathcal{A}$ and is consistent with constraints on $XP_{,X}$, $X^2P_{,XX}$ and $X^3P_{,XXX}$. We find:

$$\begin{aligned}\tilde{P}(X,\phi) &= q(X,\phi) + P\left(X_0,\phi\right) - q\left(X_0,\phi\right)\\ &+ \left[P_{,X}\left(X_0,\phi\right) - q_{,X}\left(X_0,\phi\right)\right]\left(X - X_0\right)\\ &+ \frac{1}{2}\left[P_{,XX}\left(X_0,\phi\right) - q_{,XX}\left(X_0,\phi\right)\right]\left(X - X_0\right)^2\\ &+ \frac{1}{6}\left[P_{,XXX}\left(X_0,\phi\right) - q_{,XXX}\left(X_0,\phi\right)\right]\left(X - X_0\right)^3.\end{aligned} \tag{2.69}$$

(where q is an arbitrary function of ϕ and X) so that higher derivatives along X remain unconstrained. X_0 can be fixed by a gauge choice, associated with field redefinitions $\phi \to \tilde{\phi} = g(\phi)$ ([104] use the gauge $X(N_e) = 1/2$ where N_e is the number of e-folds of inflation). Choosing a value for X at a specific time in this way is also equivalent to choosing a normalization of ϕ. Having obtained the general action (2.69) one can thus write down a theory with any given value for λ/Σ as discussed in section 2.4.3, since $P_{,XXX}$ is not constrained in any other way than through the value taken by λ/Σ.

As shown above in (2.50) the dominant shape generated by $|f_X| \gg 1$ is equilateral

$$\mathcal{A}(f_X \gg 1) \sim \mathcal{A}_{\dot{\zeta}^3} \sim -\frac{f_X - 1}{2\bar{c}_s^2}(1+\epsilon)\cos\frac{\alpha\pi}{2}\Gamma(3+\alpha)\frac{k_1^2k_2^2k_3^2}{K^3}. \tag{2.70}$$

and $f_X > 1$ is associated with negative-sign non-Gaussianities (just as produced by DBI-inflation), whereas $f_X < 1$ yields positive-sign non-Gaussianities (note that this statement is only true as long as $\epsilon \lesssim 0.3$ - otherwise the signs are reversed).

2.6 Summary

In this chapter we have considered the amplitude $\mathcal{A}$, size $f_{\rm NL}$, shape and running n_{NG} of the non-Gaussianity to be found in single field models, without assuming slow-roll. We have shown that observational constraints allow significant violations of slow-roll conditions $\epsilon \ll 1$ and $\epsilon_s \ll 1$ in inflationary single field models. As such we derived explicit bounds on slow-roll parameter ϵ in generic single field inflation scenarios. We found models with ϵ as large as ~ 0.3 satisfying all present bounds in the optimistic case of detectable (near-future) CMB non-Gaussianities (i.e. $f_{\rm NL} \gtrsim \mathcal{O}(10)$). We also found a number of significant consequences for fast-roll phenomenology:

- **Fast-roll suppresses $f_{\rm NL}$.** Figure 2.1 and eqn. (2.42) show that for equilateral non-Gaussianity, $f_{\rm NL}^{equi}$ is generically reduced when departing from the slow-roll regime. This implies that models with e.g. such a small speed of sound c_s that they violate observational constraints in the slow-roll limit, can still be allowed when considering fast-roll scenarios.

- **The shape of the bispectrum is modified.** Figure 2.1 also shows that fast-roll suppression is not an artefact of focusing on the equilateral limit, but that in fact the full bispectrum as described by $\mathcal{A}$ is fast-roll suppressed. Furthermore the shape of the amplitude is modified, where figure 2.1 illustrates a particular case where a predominantly equilateral shape is altered into an "enfolded" shape, peaking in the limit $2k_1 = k_2 = k_3$.

- **The allowed parameter-space for f_X, c_s becomes wider.** As a result of fast-roll suppression observational bounds, e.g. the WMAP result [9] $f_{\rm NL}^{\rm equil} = 26 \pm 240$ at 95% confidence, map onto weaker constraints for parameters f_X, c_s at the expense of enlarged ϵ. Figure 2.4 shows how constraints are altered, cf. [82].

- **Blue running non-Gaussianities and strong coupling constraints.** Fast-roll models generically give rise to a large blue running of non-Gaussianity n_{NG} with scale. Pairing this signal with the requirement of having weakly coupled fluctuations (so that a perturbative treatment is still applicable) allows us to put bounds on the running. We have presented exact values for the running in terms of the model parameters in section 2.4.4 and used those bounds to constrain the amount of slow-roll violation as parameterised by ϵ.

Non-Gaussianity is one of the best tools available for testing theories of primordial structure formation. With upcoming experiments promising to measure not just the size, but the full functional space mapped out by $\mathcal{A}$, constraints from the shape of non-Gaussianity may become even more important in differentiating models. In this chapter we have demonstrated that single scalar fields can source large non-Gaussianities with a much richer phenomenology than the negative $f_{\rm NL}$, pure equilateral type non-Gaussianity produced by e.g. DBI inflation. Only time will tell, but while future observations might show that slow-roll single field models are ruled

out, the simplest explanation could also very well still be a single field model. One that violates slow-roll.

Chapter 3

Constraining non-inflationary single field models

3.1 Introduction

In the previous chapter we considered effective $P(X, \phi)$ theories as arising in disformal bimetric models in an inflationary context. Here we wish to explore potential non-inflationary setups as already introduced in section 1.2.6. Consequently the overarching aim of this chapter is to show that non-inflationary models within the bimetric remit are capable of yielding primordial structure formation with a (near) scale-invariant power spectrum and distinct observational signatures. In the previous chapter we saw that an adiabatic scale invariant spectrum is produced even if the expansion - albeit still inflationary - is far from exponential (slow-roll parameters such as ϵ effectively measured the "distance" from de-Sitter expansion), provided the speed of sound varies appropriately. Here we will consider a "superluminal" phase with $c_s > 1$ which allows us to consider expansions that are not even inflationary, as long as the condition $H^2 \propto c_s$ is satisfied (here H is the Hubble parameter - also see [49].

We begin by reminding ourselves of the salient features of disformal bimetric theories (for which the speed of gravity differs from the speed of light [105, 106]) as important for this chapter. As outlined in the introduction we construct the Einstein-Hilbert action from an "Einstein" metric $g_{\mu\nu}$ (the Einstein frame), whilst minimally coupling the matter fields to a "matter" metric $\hat{g}_{\mu\nu}$ (the matter frame), with:

$$S = \frac{M_{Pl}^2}{2} \int d^4x \sqrt{-g}\, R[g_{\mu\nu}] + \int d^4x \sqrt{-\tilde{g}}\, \mathcal{L}_m[\tilde{g}_{\mu\nu}, \Psi_i] + S_\phi \tag{3.1}$$

in which S_ϕ determines the dynamics. The two metrics are related by a disformal relation, where

we will focus on the particularly simple case:

$$\tilde{g}_{\mu\nu} = g_{\mu\nu} - B(\phi)\partial_\mu\phi\partial_\nu\phi \ , \tag{3.2}$$

where ϕ is the "bi-scalar" field and we have reduced the general disformal relation (1.8) by fixing the conformal factor $A(\phi, X) = 1$ and by considering the non-derivative limit for the disformal factor, i.e. $B(\phi, X) \to B(\phi)$. Here B is chosen to have dimensions of M^{-4}, so that ϕ has dimensions of M. We reiterate that we use metrics with signature $-+++$, and B is defined so that $B > 0$ corresponds to a speed of light *larger* than the speed of gravity. As such there are two light cones at any point, one for massless matter particles, another for gravitons. More generally the two metrics may be seen as independent representations of the local Lorentz group (or two non-equivalent tetrads [106]), one valid for gravitons and the other for matter. Thus, different Lorentz transformations must be used to transform among measurements made with matter and gravity (or equivalently, with clocks and rods operated by matter or gravitational phenomena). For this reason no causality paradoxes arise, in contrast to straightforward tachyonic matter [107]. This argument shows why the bimetric construction is important in interpreting "superluminal" structure formation models.

A number of dynamics S_ϕ for bimetric theories have been considered. It was pointed out in [23] that a Klein-Gordon equation for ϕ in the matter frame translates into DBI dynamics in the gravity frame. As we saw in 1.2.5, its corresponding Lagrangian, however, is not the Klein-Gordon Lagrangian in the matter frame, but simply a cosmological constant. (It was first noted in [105] that for bimetric theories a Klein-Gordon action in the matter frame doesn't translate into a Klein-Gordon equation in that frame). Thus a particularly simple bi-scalar dynamics is generated by

$$S_\phi = \int \sqrt{-\tilde{g}}(-2\Lambda_m) \ , \tag{3.3}$$

and $\Lambda_m < 0$ leads to a speed of light larger than the speed of gravity. If we require the field ϕ to have Klein-Gordon dynamics in the Einstein frame at low energies (when matter and gravity frames coincide), we should consider additionally:

$$S_\phi = \int \sqrt{-\tilde{g}}\frac{1}{B} - \int \sqrt{-g}\frac{1}{B} \tag{3.4}$$

i.e. a positive constant term Λ_m in the Einstein frame balanced by a negative one in the matter frame, both with magnitude tuned to $1/(2B)$. This action maps into the DBI action [108, 22] in the gravity frame with a DBI warp factor $f = -B$. For a choice of sign where the speed of light in the gravity frame is larger than one ($f = -B < 0$), this is sometimes labeled "anti-DBI" (although one should note that "flipping" the sign of f means that this setup cannot be interpreted as portraying a relativistic probe brane embedded in a five dimensional bulk, as

usual for DBI; for an earlier study of anti-DBI theories see also [109]). As we will show in this chapter, when combined with a mass potential in the gravity frame it leads to scaling solutions and scale-invariant fluctuations (for related work also see [110, 111, 49]), without the need for accelerated expansion or a contracting pre-Big-Bang phase.

We will also be interested in the non-Gaussian signatures for such bimetric models. In the previous chapter we saw how, in the presence of a speed of sound $c_s \neq 1$ for adiabatic perturbations, the three-point function contains terms which are proportional to the power spectrum squared and further terms which are amplified by an additional factor c_s^{-2} [37, 80, 66]. This c_s^{-2}-dependence is of course instrumental in generating large non-Gaussian signals for non-canonical scalar field models, since in the subluminal case $c_s < 1$, the "c_s^{-2}" terms dominate. In addition to enhancing non-Gaussianities in e.g. DBI inflation, we also saw that this makes the three-point function scale dependent [101, 85] in the case of a varying speed of sound. This happens since the requirement of (near) scale-invariance sets the combination H^2/c_s to be approximately constant. As a result terms that appear with different powers of c_s will therefore run with the scale. In the opposite limit, the one of large speed of sound that we are considering here, the "c_s^{-2}" terms are suppressed and we will consequently find that the remaining terms inherit the scale invariance from the power spectrum of the two point function. The dimensionless quantity $f_{\rm NL}^{equi}$ is of order 1 and has opposite sign to DBI inflation, i.e. $f_{NL}^{equi} \sim \mathcal{O}(1) > 0$ (with the WMAP sign convention).

The structure of this chapter is as follows. In Section 3.2 we review some of the relevant cosmological perturbation theory as discussed in the previous chapter, but with an eye on difference introduced by a "superluminal" speed of sound. Then, in Section 3.3 we explain how scale invariance may be achieved in these models and also show how a non-scale-invariant power spectrum is generated when a non-trivial disformal factor $B(\phi)$ is considered. We derive the associated non-Gaussian features in Section 3.4 and also show how a departure from an exactly scale-invariant two-point function can lead to interesting non-Gaussian properties here. Finally we will examine our results from a wider perspective in a concluding section 3.5.

3.2 The power spectrum

Here we adapt the calculation from section 2.2, but in the context of disformal bimetric models with $c_s > 1$. Projecting (3.4) onto the Einstein frame leads to the (anti)-DBI action, which belongs to the general class of $P(X, \phi)$ models. We recall that this means we are considering an action of the form:

$$S = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{R}{2} + P(X,\phi)\right]. \tag{3.5}$$

The energy density reads $\rho = 2XP_{,X} - P$ while, as before, the speed of sound is given by

$$c_s^2 = \frac{P_{,X}}{\rho_{,X}} = \frac{P_{,X}}{P_{,X} + 2XP_{,XX}} \,. \tag{3.6}$$

Note that, if we consider a non-inflationary solution i.e. a space-time with decelerating expansion rate $\ddot{a} < 0$, then conformal time can be taken to be positive and starting from zero (in contrast to inflation where it starts at $-\infty$ and inflation ends at $\tau = 0$). Accordingly we know from (2.9) that, in the case of constant ϵ and ϵ_s we can now express the (conformal) time-dependence of the speed of sound as

$$c_s \sim (\tau)^{\frac{\epsilon_s}{\epsilon - 1}} \tag{3.7}$$

We are in a non-inflationary solution, so $\epsilon > 1$ and we may already suspect from 2.2 that a negative ϵ_s is needed in order to obtain a scale-invariant solution. We now show that this is indeed the case, postulating a speed of sound which diverges with conformal time according to $c_s \propto \tau^{-\alpha}$, where we have introduced $\alpha = -\epsilon_s/\epsilon - 1$ as shorthand (i.e. $\alpha > 0$ here). Whether we employ a hydrodynamical or a scalar field description, the density fluctuations are described by the modified harmonic oscillator equation (2.11). Writing this in terms of conformal time we have

$$v'' + \left[c_s^2 k^2 - \frac{z''}{z}\right] v = 0 \,. \tag{3.8}$$

where the curvature perturbation ζ obeys $\zeta = -v/z$ and $z \propto \frac{a}{c_s}$. Written in this form the solution is provided by Bessel functions with[1]

$$v = \sqrt{\beta\eta}(AJ_\nu(\beta c_s k\tau) + BJ_{-\nu}(\beta c_s k\tau)) \,. \tag{3.9}$$

Defining $\beta = 1/(\alpha - 1) > 0$, it was shown in [49] that the order ν is given by

$$\nu = \beta\left(\alpha - \frac{3(1-w)}{2(1+3w)}\right), \tag{3.10}$$

where we have assumed a constant equation of state w. Considering the asymptotic solution inside the sound horizon (i.e. when the $c_s^2 k^2$ term dominates) one can derive a WKB solution for v inside the horizon

$$v \sim \frac{e^{ik\int c_s d\tau}}{\sqrt{c_s k}} \sim \frac{e^{-i\beta c_s k\tau}}{\sqrt{c_s k}} \,. \tag{3.11}$$

Considering the opposite regime, the power spectrum outside of the sound horizon can now be found as well. In this regime

$$v \sim \frac{\sqrt{\beta\tau}}{(c_s k\tau)^\nu} \,. \tag{3.12}$$

[1] As shown in 2.2 this reduces to a Hankel function solution if written in terms of sound horizon time y.

where we have only shown the negative order solution, since $c_s\tau$ being a decreasing function of time, this is the growing mode. Scale-invariance of the curvature fluctuation (i.e. requiring $k^3\zeta^2$ to be constant) is still associated with $\nu = 3/2$, i.e. the order parameter of the Bessel function remains invariant under a change from conformal time τ to sound horizon time y. For a scale-invariant power spectrum we therefore require

$$\alpha = \alpha_0 = 6\frac{1+w}{1+3w} \tag{3.13}$$

and if we rewrite c_s in terms of the density ρ we conclude that this implies $c_s \propto \rho$ for all (constant) w.

Different spectral indices can be obtained by introducing $\alpha < \alpha_0$ or $\alpha > \alpha_0$, which becomes clear when rewriting (2.17) in terms of α and β

$$n_s - 1 = \frac{2\epsilon + \epsilon_s}{\epsilon_s + \epsilon - 1} = \beta(\alpha - \alpha_0)\ . \tag{3.14}$$

[49] also shows that the power spectrum may be written

$$k^3\zeta^2 \sim \frac{(5+3w)^2}{1+w}\frac{\rho}{M_{Pl}^4 c_s}\ , \tag{3.15}$$

reiterating that $c_s \propto \rho$ for a scale invariant power spectrum.[2] In other words, a rapidly decreasing speed of sound c_s can provide a scale-invariant power spectrum even in non-inflationary settings, where the fact that $c_s \gg 1$ initially here allows us to solve the horizon problem as outlined in 1.2.6.

Working explicitly with conformal time τ makes the physical significance of a rapidly decreasing c_s more visible, but now we switch back to the computationally more convenient sound horizon time y. As before the quadratic action then takes the form

$$S = \frac{M_{\rm Pl}^2}{2}\int \mathrm{d}^3x\mathrm{d}y\ q^2\left[\zeta'^2 - (\vec{\nabla}\zeta)^2\right]\ , \tag{3.16}$$

where $' \equiv \mathrm{d}/\mathrm{d}y$, and

$$q \equiv \sqrt{c_s}z = \frac{a\sqrt{2\epsilon}}{\sqrt{c_s}}\,. \tag{3.17}$$

and the solution for canonically-normalized $v = M_{\rm Pl}q\zeta$ is given in terms of Hankel functions

$$v_k(y) = \frac{\sqrt{\pi}}{2}\sqrt{y}\,H_\nu^{(1)}(ky). \tag{3.18}$$

[2]For standard cosmology to arise at low energy densities, we effectively require $c_s = c_0(1+\rho/\rho_\star)$, where c_s is approximately constant at low-energy and $\rho_\star$ is the density scale at which the speed of sound begins to diverge.

The expression for the ζ power Spectrum in terms of slow-roll parameters still reads

$$P_\zeta \equiv \frac{1}{2\pi^2} k^3 \left|\zeta_k\right|^2 = \frac{(\epsilon_s + \epsilon - 1)^2\, 2^{2\nu-3}}{2(2\pi)^2 \epsilon} \frac{\bar{H}^2}{\bar{c}_s M_{\rm Pl}^2}\,. \tag{3.19}$$

This reiterates that for scale-invariance to obtain, $c_s \propto H^2$ and since we can relate H^2 to the energy density ρ via the Friedmann equations, this once again establishes $c_s \propto \rho$ as a requirement for a scale-invariant solution.

3.3 Scaling solutions

Generically the parameter B appearing in the simplified disformal transformation (3.2) considered in this chapter is itself a function of ϕ. In this section we show that non-minimal theories with power-law $B(\phi)$ lead to tilted spectra, without running. We also establish the scale-invariant limit which corresponds to the zeroth order $B =$ constant solution. For the bimetric theories discussed in the above, in the Einstein frame the action takes the DBI-form:

$$P(X,\phi) = -f^{-1}(\phi)\sqrt{1 - 2f(\phi)X} + f^{-1}(\phi) - V(\phi) \tag{3.20}$$

where we have made the identification $B(\phi) = -f(\phi)$. Scaling solutions of this action have been studied in [112, 66, 107]. In particular we can solve for the Hubble parameter H and speed of sound c_s as functions of the scalar field ϕ. For constant ϵ and ϵ_s we have

$$H(\phi) = H(\phi_0)\left(\frac{\phi}{\phi_0}\right)^{-2\epsilon/\epsilon_s} \tag{3.21}$$

$$c_s(\phi) = c_s(\phi_0)\left(\frac{\phi}{\phi_0}\right)^{2}. \tag{3.22}$$

The Friedmann equations then give a solution for $V(\phi)$ and $f(\phi)$

$$V = 3M_{Pl}^2 H^2(\phi)\left(1 - \frac{2\epsilon}{3}\frac{1}{1+c_s(\phi)}\right) \tag{3.23}$$

$$f(\phi) = \left(\frac{1}{2M_{Pl}^2\epsilon}\right)\frac{1 - c_s^2(\phi)}{H^2(\phi)c_s(\phi)} \tag{3.24}$$

where substituting in for H and c_s we find

$$V = 3M_{Pl}^2 H^2(\phi)\left(1-\frac{2\epsilon}{3}\frac{1}{1+c_s(\phi_0)\left(\frac{\phi}{\phi_0}\right)^2}\right) \tag{3.25}$$

$$f(\phi) = \frac{1}{2M_{Pl}^2\epsilon H^2(\phi_0)c_s(\phi_0)}\left(\frac{\phi}{\phi_0}\right)^{2\epsilon/\epsilon_s-2}\left(1-c_s(\phi_0)\left(\frac{\phi}{\phi_0}\right)^2\right) \tag{3.26}$$

In the large ϕ limit, which we can see from (3.21) corresponds to the large c_s limit, we can simplify these expressions to obtain

$$V \sim 3M_{Pl}^2 H^2(\phi) = V_0\left(\frac{\phi}{\phi_0}\right)^{-4\epsilon/\epsilon_s} \tag{3.27}$$

$$f(\phi) \sim -\left(\frac{1}{2M_{Pl}^2\epsilon}\right)\frac{c_s(\phi)}{H^2(\phi)} = -\frac{c_s(\phi_0)}{2M_{Pl}^2\epsilon H^2(\phi_0)}\left(\frac{\phi}{\phi_0}\right)^{2+4\epsilon/\epsilon_s}. \tag{3.28}$$

As before a scale-invariant two-point function is associated with $\epsilon_s = -2\epsilon$, so that the potential V and warp factor f then become

$$V(\phi) \sim 3V_0\left(\frac{\phi}{\phi_0}\right)^2 \tag{3.29}$$

$$f(\phi) \sim -\frac{c_s(\phi_0)}{2M_{Pl}^2\epsilon H^2(\phi_0)}, \tag{3.30}$$

i.e. $V \propto \phi^2$ and $f =$ constant in the scale-invariant limit. Importantly we can explicitly show via the Friedmann equation that $\rho \sim H^2$ and consequently $\rho \sim \phi^2 \propto c_s$ here. This means that, in the large ϕ/large c_s limit and for scale-invariant solutions, we recover the solution discussed in [23]. As we discussed in the previous section this shows how scale-invariance is associated with the universal law $c_s \propto \rho$, constant disformal factor B and a quadratic potential. As equation (3.25) shows the same does not happen if we depart from scale-invariance, in which case in particular the disformal factor B picks up a dependence on ϕ. A power-law B, however, does not generate running and is consistent with constant ϵ and ϵ_s. Naturally, more complicated $B(\phi)$ would lead to more complex spectra, so that this absence of running is not a general feature of models with arbitrary disformal factor $B(\phi)$.

3.4 Non-Gaussian signals

Having worked out solutions for power spectra with various constant spectral indeces n_s we now turn our attention to non-Gaussian features. In the previous chapter we reproduced (2.40) which applied for any $P(X,\phi)$ model regardless of how precisely the speed of sound varied. As we have

seen, in a non-inflationary bimetric model we require $c_s \gg 1$ during structure formation. In this large c_s limit (which corresponds to the large ϕ limit) the amplitude of the three point function $\mathcal{A}$ can be read straightforwardly from (2.40) (we refer the interested reader to chapter 2 and the appendices for details). Comparing with the cubic effective action (2.30) we find that only the $\mathcal{A}_{\zeta\dot{\zeta}^2}$ and $\mathcal{A}_{\zeta(\partial\zeta)^2}$ terms are not subdominant as $c_s \to \infty$. The resulting total amplitude is independent of the parameters (w or ϵ) and reads

$$\mathcal{A}_{c_s\to\infty} = -\frac{1}{8}\sum_i k_i^3 + \frac{1}{K}\sum_{i<j} k_i^2 k_j^2 - \frac{1}{2K^2}\sum_{i\neq j} k_i^2 k_j^3 \,. \tag{3.31}$$

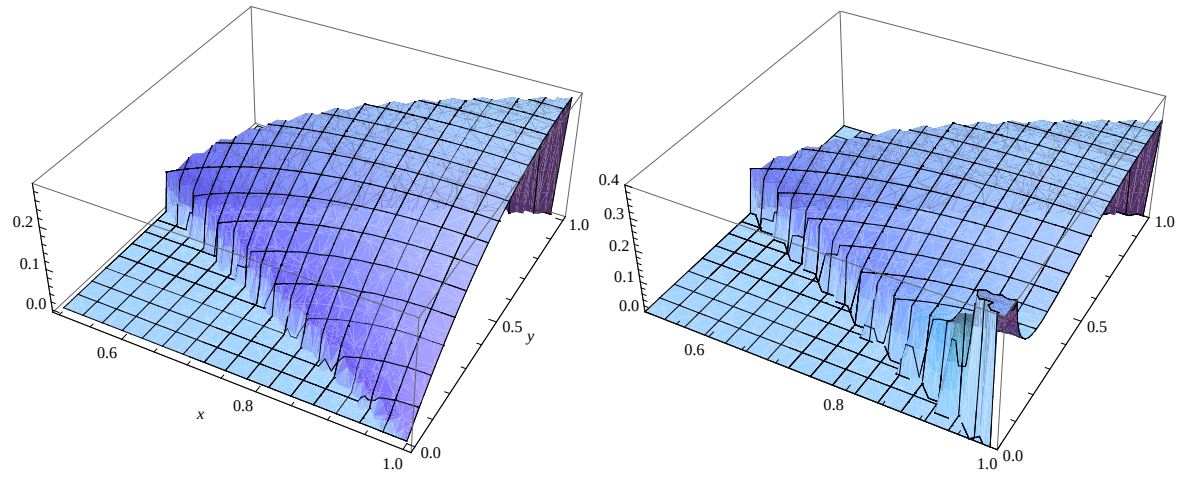
This is precisely the equilateral shape, peaking for $k_1 = k_2 = k_3$, that is also obtained in the scaling solutions considered in [66] in the $\epsilon \to 0$, $\alpha \to 0$ limit. More specifically one obtains

$$\mathcal{A}_{\epsilon\to 0} = \left(1 - \frac{1}{c_s^2}\right)\mathcal{A}_{c_s\to\infty} \;+\; \mathcal{O}(n_s - 1) \,, \tag{3.32}$$

Interestingly this means that the non-Gaussian amplitude is uniquely fixed and does not depend on other background parameters in the scale-invariant limit here. Also note that, while the full three-point function contained terms proportional to P_ζ^2 and further terms which are amplified by an additional factor c_s^{-2}, these additional terms are suppressed in the large c_s limit. Since the relevant non-Gaussian amplitude here is consequently proportional to P_ζ^2, the (near) scale-invariance of the power spectrum straightforwardly maps onto the three-point function. In contrast to the inflationary models considered in the previous chapter, where the additional c_s^{-2} dependence was instrumental both in yielding large non-Gaussian amplitudes as well as in generating a sizeable running of the three-point function $n_{\rm NG}$, $n_{\rm NG} \sim 1$ here.

In terms of $f_{\rm NL}^{equi}$ comparison with (2.42) shows we here have $f_{\rm NL}^{equi} \sim 0.27 + \mathcal{O}(c_s^{-2} + \mathcal{O}(n_s - 1)$. In particular we would like to stress that the f_X-dependence is suppressed by c_s^{-2}. Whilst no $f_{\rm NL} \gg 1$ is obtained in accordance with (2.44), this shows that non-Gaussianities in excess of the levels yielded by slow-roll inflation ($f_{\rm NL}^{sr} \sim \epsilon \ll 1$) are possible here. In other words, in the minimal bimetric model considered here the dimensionless quantity f_{NL}^{equi} is of order 1 and has opposite sign to DBI inflation, i.e. $f_{NL} \sim 0.27 > 0$. Thus the model is quite distinct in this respect to standard slow-roll inflation (for which $f_{NL} \sim \epsilon$) and DBI inflation (for which $f_{NL} \sim -\mathcal{O}(100)$ is a distinct possibility.) Notice there's been some confusion [22], both among theorists and observers, regarding the sign of f_{NL}. Here we adopt the convention used by WMAP, where positive $f_{\rm NL}$ physically corresponds to negative-skewness for the temperature fluctuations: this means we assign a negative f_{NL} to DBI inflation, so that $f_{NL} > 0$ for the anti-DBI models under consideration.

We finish this section by estimating the $\mathcal{O}(n_s - 1)$ corrections to the non-Gaussian amplitude.



(a) $-\mathcal{A}(1, x_2, x_3)/(x_2 x_3)$ for $n_s = 1$ (b) $-\mathcal{A}(1, x_2, x_3)/(x_2 x_3)$ for $n_s = 0.96$

Figure 3.1: We plot the non-Gaussian amplitude from Eq. (3.33) $-\mathcal{A}(1, x_2, x_3)/(x_2 x_3)$ for $n_s = 1$ (left) and $n_s = 0.96$ (right).

Following the procedure outlined in appendix D, we may combine terms from equations (D.15) and (D.16). Taking the small tilt ($n_s - 1 \ll 1$) and $c_s \gg 1$ limits, we find

$$\begin{aligned}\mathcal{A} \;=\; & \left(\frac{k_1 k_2 k_3}{2K^3}\right)^{n_s-1} \left[-\frac{1}{8}\sum_i k_i^3 + \frac{1}{K}\sum_{i<j} k_i^2 k_j^2 - \frac{1}{2K^2}\sum_{i\neq j} k_i^2 k_j^3\right. \\ & + (n_s - 1)\left(-\frac{1}{8}\sum_i k_i^3 - \frac{1}{8}\sum_{i\neq j} k_i k_j^2 + \frac{1}{8}k_1 k_2 k_3 + \frac{1}{2K}\sum_{i<j} k_i^2 k_j^2 - \frac{1}{2K^2}\sum_{i\neq j} k_i^2 k_j^3\right) \\ & \left. + \mathcal{O}\left(\frac{1}{c_s^2}\right)\right] \,, \end{aligned} \tag{3.33}$$

where the only dependence on ϵ and ϵ_s appears either in the "observable" combination $n_s - 1$ or in the subleading $\mathcal{O}(1/c_s^2)$ terms. Upon approaching scale invariance only the first line inside the square brackets stays relevant, ensuring that $\mathcal{A}$ reduces to the equilateral amplitude (3.31) as required. The amplitude (3.33) is plotted in figure 3.1 and peaks in the equilateral limit $k_1 = k_2 = k_3$. In the local limit $k_1 \ll k_2, k_3$, on the other hand, the first line inside the square brackets of Eq. (3.33) goes to zero. In agreement with the consistency relation [89, 99] we then have

$$\mathcal{A}_{k_1 \ll k_2, k_3} \;\approx\; -\frac{1}{2}(n_s - 1)\left(\frac{k_1}{k_2}\right)^{n_s-1} \tag{3.34}$$

The predictive power of our result lies in establishing a consistency relationship between n_s and $\mathcal{A}$. In fact, we find a distinctive Non-Gaussian signal for any given spectral index n_s. Whilst the overall non-Gaussian amplitude $\mathcal{A}$ still peaks in the equilateral limit $k_1 = k_2 = k_3$ in both

red- and blue-tilted cases, its shape is modified when compared with the scale invariant limit. Note that this approach is an estimate in the sense that we have ignored higher order propagator corrections, i.e. we have essentially approximated the Hankel function dependence of the mode function v_k in the limit $|ky| \ll 1$, ignoring additional $\mathcal{O}(n_s - 1)$ corrections (see appendix D for details). However, naturally such extra corrections are expected to modify the (n_s-1) dependence of the amplitude, but they should not re-introduce any dependence on background quantities such as ϵ. One may therefore expect that the non-Gaussian amplitude is still uniquely specified in terms of $n_s - 1$, even as scale-invariance is broken. It will be an interesting task for the future to explicitly check whether higher order propagator corrections leave this feature unmodified. Regardless of whether this is the case or not one should remember that this correction will be very challenging to detect observationally, so that a non-Gaussian amplitude without running and consequently scale-independent and almost constant $f_{\rm NL} \sim 0.27$ is the key observational signal here (still challenging, but perhaps not impossible when additional modes, e.g. via 21cm surveys, are taken into account). Note that the peaks in the squeezed limit in figure 3.1 are solely due to breaking of scale-invariance and thus are a direct measure of n_s. They may still be observationally elusive, if confined too far into the squeezed limit. We once again emphasize that no significant constraints arise for bimetric theories from n_{NG}, as the tilt of the three-point function is strongly suppressed for $c_s \gg 1$.

3.5 Summary

Strict scale-invariance has been associated with superluminal bimetric models, where the speed of light is larger than the speed of gravity in the early Universe [49, 23]. Indeed this is a feature of the minimal bimetric model, but in this chapter we showed how tilted spectra, red or blue, could be generated by a non-minimal bi-scalar coupling $B(\phi)$. At first this might suggest we've fallen into the "theory of anything" trap, but it's not the case. A unique non-Gaussian shape is predicted for the scale-invariant case and distinct distortions away from the scale-invariant equilateral shape appear for each of the tilted cases. These distortions can be seen as "consistency conditions" for this class of models. This is particularly relevant given the absence of gravitational waves for all non-inflationary and "superluminal" bimetric models of this kind. (Note that these models solve the horizon problem for matter but not for gravity, so tensor modes don't start their lives inside the horizon.)

One might wonder where the proposed running coupling $B(\phi)$ comes from. First note that we don't need the full (anti-)DBI action (2.59) resulting from (3.4), unless we impose Klein-Gordon dynamics for ϕ in the Einstein frame at low energies. This may not be necessary, and if we relax this requirement all we need is (3.3), i.e. a negative cosmological constant Λ_m in the matter

frame (which, we stress, does *not* lead to an AdS solution). In fact, if we relax the low-energy requirement, Λ_m doesn't even need to be related to B. If, however, we do insist on Klein-Gordon dynamics for ϕ in the Einstein frame at low energies, then the negative matter frame cosmological constant should be exactly balanced by a positive Einstein frame cosmological constant, and their common magnitude should be $1/(2B)$.

A number of interesting theoretical connections can be made. In the context of emergent geometry, it's been pointed out that different emergent metrics may apply to bosons and fermions [113]. The fact that the vacuum energy is negative for fermions and positive for bosons suggests an action of the proposed form, with a speed of light larger than the speed of gravity (i.e. an anti-DBI action in the Einstein frame). Also these models become asymptotically a cuscaton [110, 111] model, a feature that may be used to support the view that they are a UV-complete alternative to inflation. Finally, it is possible that this construction results from an entirely different set up, such as deformed dispersion relations [114]. It is interesting that the dispersion relations needed for scale-invariance are of the same form as those discussed in the context of Horava-Lifshitz theory [115]. More generally a connection with deformed special relativity remains to be fully explored [116, 117, 118]. Absence of exact scale-invariance could then also be a major clue into the foundations of these theories.

In summary, in this chapter we showed how cosmological structure formation may proceed in non-inflationary bimetric models. (Near) scale-invariant power spectra are readily produced and come associated with a distinctive non-Gaussian signature. Interestingly we also saw that the minimal disformal model with constant disformal factor B leads to scale-invariance, whereas power-law profiles for B are associated with a spectral tilt n_s without running.

Chapter 4

Beyond single field models : Three-form inflation

4.1 Introduction

While invoking the presence of one or multiple scalar fields in order to drive inflation has proven to be a particularly simple and popular approach, several alternatives have been proposed. These include inflation driven by vector fields [119, 120, 121, 122, 123], and higher form fields [124, 125, 126]. And even though, while this thesis is being written, scalar fields (most probably the Higgs field) have been observed in nature for the first time, this hardly proves the case for scalars in cosmology, perhaps serving as motivation to explore non-scalar field scenarios further in this context. However, for non-scalar inflationary models the techniques to make predictions beyond the power spectrum are less well understood than for scalar field models. In this chapter, our interest lies in extending existing techniques for the recently proposed three-form inflationary scenario [27, 28], in which inflation is driven by a massive three-form field.

As we have seen in chapter 1, during inflation quantum fluctuations are promoted to classical perturbations, and the statistics of these perturbations are probed e.g. by observations of the cosmic microwave background and of large scale structure. In contrast to higher order form models, the techniques for making predictions for the statistics of perturbations produced by general $P(X, \phi)$ inflation models as relevant in the disformal bimetric context are extremely well developed (see for example [89, 37, 80] and references therein). The aims in this chapter are therefore two-fold.

Firstly we review and extend previous work [28] on the existence of a series of dualities between different n-form theories. In particular we will show that these dualities relate the actions for several higher-order form field models to that of equivalent scalar field models. Essentially this happens since imposing global symmetries in accordance with an FLRW space-time (homogeneity

and isotropy) reduces the effective degrees of freedom in the theory, so that an effective scalar field description can be recovered. This opens up a tempting avenue for understanding the dynamics and observational predictions of e.g. three-form inflation. Namely, having established a dual scalar field description, we may move to working with the dual non-canonical scalar field model, for which suitable techniques are readily available. Here we point out various subtleties in the mapping between theories (which is generically rather complicated), and use dualities in order to establish a self-contained three-form formalism. In the end this allows computation of, in particular, the bispectrum without explicit reference to the dual scalar field theory (i.e. the final result can be expressed in terms of three-form quantities only). This approach recovers previous results for the two-point function [27, 28, 127] and also allows us to rapidly derive expressions at higher order in perturbation theory. Secondly, we use the resulting formalism to investigate non-Gaussian features of the three-form theory for a particularly simple case - slow-rolling 3-form inflation with a power-law potential. We find that the phenomenology closely resembles that of DBI-inflation [22] and hence of minimal bimetric models as discussed in the previous two chapters.

The map for this chapter is as follows. We begin by reviewing 3-form inflation in 4.2. Given the central importance of dualities to our work, we then discuss these in 4.3. Employing these dualities we then derive non-Gaussian features in 4.4, expressing the cubic action in terms of three-form variables and the curvature perturbation ζ only. We also discuss the explicit example of a power-law three-form potential. Finally we summarize results in 4.5.

4.2 Three-form inflation

We will focus on a canonical theory minimally coupled to Einstein gravity with action [27, 28]

$$S = -\int d^4x\sqrt{-g}\left(R - \frac{1}{48}F^2 - V(A^2)\right), \tag{4.1}$$

where $F = \nabla.A$ is a four-form, A a three-form, and we have adopted a compact notation in which indices are suppressed once the valance of an object is specified, squaring a quantity denotes contraction of all indices (for example e.g. $F^2 \equiv F_{\mu\nu\rho\sigma}F^{\mu\nu\rho\sigma}$), and a dot $(\cdot)$ denotes contraction on the first index (for example $\nabla \cdot A = \nabla_\mu A_{\nu\rho\sigma}$). Here and throughout Greek indices label space-time dimensions and Roman indices spatial dimensions. The potential $V(A^2)$ here generically breaks the gauge $A \to A + \lambda I$, allowing the three-form to become dynamical (otherwise the four-form field strength will simply act as an effective cosmological constant [128]).

In the cosmological context, the unperturbed three-form must be consistent with the universe's symmetries of homogeneity and isotropy. The non-zero components of the background three-form

compatible with this geometry can therefore be written in the form

$$A_{ijk} = a(t)^3 \epsilon_{ijk} \chi(t) \tag{4.2}$$

where $a(t)$ is the scale factor of the universe as a function of cosmic time, $\chi(t)$ is a comoving field and the totally antisymmmetric tensor is defined properly in section 4.3.2.[1] Hence the unperturbed three-form satisfies $A^2 = \chi^2/6^2$. The dynamics of the universe are then governed by the behavior of the scalar quantity $\chi(t)$ which is directly related to the three-form. As emphasized elsewhere [27, 28], when written in terms of this scalar quantity, the role of the three-form potential and the equations of motion which govern the universe are straightforward to interpret. In particular the Friedmann equations are given by (adopting units where $8\pi G = 1$)

$$\begin{aligned} 3H^2 &= \frac{1}{2}\left(\dot{\chi} + 3H\chi\right)^2 + V(\chi^2) \\ \dot{H} &= -\frac{1}{2}\chi V_{,\chi}. \end{aligned} \tag{4.3}$$

The energy density and pressure of the field χ are therefore given by

$$\rho_\chi = \frac{1}{2}(\dot{\chi} + 3H\chi)^2 + V(\chi)\,, \tag{4.4}$$

$$p_\chi = -\frac{1}{2}(\dot{\chi} + 3H\chi)^2 - V(\chi) + \chi V_{,\chi}\,. \tag{4.5}$$

$$\tag{4.6}$$

We can therefore write the equation of state for the three-form, $w_\chi = p_\chi/\rho_\chi$ as

$$w_\chi = -1 + \frac{\chi V_{,\chi}}{\rho_\chi}\,. \tag{4.7}$$

Interestingly this directly shows that cosmological constant-like behavior is directly linked to a "slow-rolling" potential here: $w = -1$ when the potential's slope vanishes. We also note that the equation of state is unbounded from both above and below.

[1] Throughout this chapter we will explore Hodge duals between different n-forms. But, jumping ahead slightly, we may explain the form of A very simply in this way: Consider the dual of A in a four-dimensional theory - a vector B. In order for a single-vector theory to be consistent with isotropy and homogeneity, no preferred directions can exist in the theory, i.e. the vector needs to be aligned with the time-like direction of the cosmological frame. Hence the three-form dual is space-like.

[2] Note that in [27, 28], the three form quantity was denoted by X not χ, but here we reserve X for denoting the canonical kinetic term for the scalar ϕ in order to match notation in previous chapters

4.3 Dual n-form theories

In this section we discuss a number of dualities which inter-relate n-form theories in four space-time dimensions. This will be an essential step in establishing a number of n-form theories which can be described by an effective scalar-tensor theory as arising in disformal bimetric setups. Even outside of the disformal context, these dualities are very useful, since they allow us to analyze n-form theories in terms of equivalent formulations where the degrees of freedom of a particular model are more transparent. Some of these dualities were already pointed out in [28] and we extend results wherever needed. As we will see, the dual picture makes further calculations significantly more tractable, promoting the existence of a class of dual theories from a curiosity to an efficient calculational tool.

4.3.1 Equations of motion for the three-form

We begin by considering a Palatini-type action, which treats the three-form A and the four-form F as independent variables, and which is equivalent to (4.1) up to boundary terms. The Lagrangian for the matter part of the action is given by

$$\mathcal{L}_1 = \frac{1}{48}F^2 - \frac{1}{6}A\nabla\cdot F - V(A^2)\,. \tag{4.8}$$

Variations of the associated action relate A and F. The resulting equations of motion are

$$\begin{aligned} F &= -4\,[\nabla A] \\ \nabla\cdot F &= -12AV'\left(A^2\right), \end{aligned} \tag{4.9}$$

where a prime denotes differentiation with respect to the argument in brackets. Integrating the middle term of $\mathcal{L}_1$ by parts inside the action, one finds

$$\mathcal{L}_2 = \frac{1}{48}F^2 + \frac{1}{6}F\,[\nabla A] - V(A^2), \tag{4.10}$$

which consequently shares the same equations of motion. The first equation of (4.9) now appears as a constraint equation and may be substituted back into $\mathcal{L}_2$ to confirm that we do indeed recover (4.1). We note that we are always free to perform such an integration by parts and this procedure leaves the dynamics of the theory invariant. Moreover, we are free to substitute constraint equations which do not change the order back into the action.

4.3.2 Dual fields

Our aim is to rewrite Lagrangians (4.8) and (4.10) in terms of fields Hodge ($*$) dual to $A_{\mu\nu\rho}$ and $F_{\mu\nu\rho\sigma}$. Here we set up the dual field picture, providing some additional pedagogical detail, since this procedure has led to significant confusion in previous treatments of this topic. We recall that any p-form has a dual d-p form and begin by defining a totally anti-symmetric tensor ϵ as

$$\epsilon^{\alpha_1,\alpha_2...\alpha_p} = \frac{1}{\sqrt{|g|}} \begin{cases} +1 & \text{if } (\alpha_1,\alpha_2...\alpha_p) \text{ is an even permutation of } (1,...d) \\ -1 & \text{if } (\alpha_1,\alpha_2...\alpha_p) \text{ is an odd permutation of } (1,...d) \\ 0 & \text{otherwise,} \end{cases} \tag{4.11}$$

where d is the dimension of the space ϵ lives in and d_t is the number of temporal dimensions, so $d_t = 1$ for the cases we will consider in this chapter which have signature $(-+++)$. Note that for a diagonal metric consequently $(-1)^{d_t} = \text{sign}(g)$, where g is the determinant of $g_{\mu\nu}$. Lowering all indices with the metric $g_{\mu\nu}$ (note this is a valid procedure since we are explicitly dealing with the tensor ϵ, not its associated tensor density) one finds

$$\epsilon_{\alpha_1,\alpha_2...\alpha_p} = (-)^{d_t}\sqrt{|g|} \begin{cases} +1 & \text{if } (\alpha_1,\alpha_2...\alpha_p) \text{ is an even permutation of } (1,...d) \\ -1 & \text{if } (\alpha_1,\alpha_2...\alpha_p) \text{ is an odd permutation of } (1,...d) \\ 0 & \text{otherwise.} \end{cases} \tag{4.12}$$

A particularly useful identity we will use repeatedly is

$$\epsilon_{\alpha_1....\alpha_n}\epsilon^{\beta_1...\beta_n} = (-)^{d_t} d!\delta^{[\beta_1}_{[\alpha_1}\delta^{\beta_2}_{\alpha_2}...\delta^{\beta_n]}_{\alpha_n]}. \tag{4.13}$$

Let us now consider an arbitrary p-form living in an n-dimensional space, where we choose a coordinate basis and write

$$P \equiv \frac{1}{p!}P_{\alpha_1,...\alpha_p}dx^{\alpha_1}\wedge...dx^{\alpha_p}, \tag{4.14}$$

In terms of the totally antisymmetric tensor ϵ the (Hodge $\star$) dual of this p-form is given by

$$(\star P)_{\alpha_1,...\alpha_{d-p}} = \frac{1}{p!}\epsilon_{\alpha_1,...\alpha_{d-p}\beta_1....\beta_p}P^{\beta_1,...\beta_p}. \tag{4.15}$$

As expected this means any p-form has therefore a dual which is a d-p form.

4.3.3 Dual actions

Our primary aim is to rewrite Lagrangians (4.8) and (4.10) in terms of the Hodge ($*$) dual fields to A and F. We recall that any p-form has a dual $(d-p)$-form, where d is the number of space-time dimensions, four in our context. In particular, the three-form A and four-form F that make

up (4.10) can be expressed in terms of their duals as

$$\begin{aligned} (\star F) &= \frac{1}{4!}\epsilon_{\alpha\beta\gamma\delta}F^{\alpha\beta\gamma\delta} \equiv \Phi & F_{\alpha\beta\gamma\delta} &= -\epsilon_{\alpha\beta\gamma\delta}\Phi \\ (\star A)_\alpha &= \frac{1}{3!}\epsilon_{\alpha\beta\gamma\delta}A^{\beta\gamma\delta} \equiv B_\alpha & A_{\beta\gamma\delta} &= -\epsilon_{\alpha\beta\gamma\delta}B^\alpha \end{aligned} \tag{4.16}$$

The Hodge duals to F and A enable us to recast the original theory (4.1) into a scalar-vector description, with Lagrangian

$$\mathcal{L}_3 = -\frac{1}{2}\Phi^2 - \Phi\nabla\cdot B - V\left(-6B^2\right), \tag{4.17}$$

which follows from Lagrangian (4.10). The equations of motion for Φ and B_μ can now be obtained either by varying (4.17) or equivalently by substituting Hodge duals into (4.9). They are

$$\begin{aligned} \Phi &= -\nabla\cdot B, \\ \nabla\Phi &= -12BV'(-6B^2). \end{aligned} \tag{4.18}$$

The first equation of motion now appears as a constraint equation with respect to $\mathcal{L}_3$ and can be substituted back to express our original theory (4.8) as a pure vector theory. Integrating the middle term of $\mathcal{L}_3$ by parts (or equivalently substituting A and F for their duals in $\mathcal{L}_2$), however, one finds that the converse is true, and the second equation of motion in (4.18) appears as a constraint and may be substituted into the action to remove the vector field in favour of the scalar (or equivalently its dual four form). With an eye on calculating inflationary observables, and in particular the 3-point correlation function, the scalar picture is particularly intriguing, since, as we have discussed, it opens up the possibility of using existing machinery for dealing with scalar field models of inflation. The final set of equivalent actions for a four-form, three-form, vector and scalar respectively are

$$\mathcal{L}_{IV}(F,\nabla\cdot F) = -\frac{1}{48}F^2 + 2A^2(\nabla\cdot F)V'\left(A^2(\nabla\cdot F)\right) - V(A^2(\nabla\cdot F)), \tag{4.19}$$

$$\mathcal{L}_{III}(A,\nabla A)) = -\frac{1}{3}\left[\nabla A\right]^2 - V(A^2), \tag{4.20}$$

$$\mathcal{L}_I(B,\nabla\cdot B) = \frac{1}{2}(\nabla\cdot B)^2 - V\left(-6B^2\right), \tag{4.21}$$

$$\mathcal{L}_0(\Phi,\nabla\Phi) = -\frac{1}{2}\Phi^2 - 12B^2(\nabla\Phi)V'\left(-6B^2(\nabla\Phi)\right) - V\left(-6B^2(\nabla\Phi)\right). \tag{4.22}$$

Here $V\left(-6B^2(\nabla\Phi)\right)$ and $V(A^2(\nabla\cdot F))$ indicate that the second equation of motion in (4.18) has to be used in order to express A^2 in terms of $\nabla\cdot F$ and B^2 in terms of $\nabla\Phi$. Note that we cannot simultaneously use the first equation of motion ($\Phi = -\nabla\cdot B$) to substitute for $-\Phi\nabla\cdot B$ in the action in order to replace it with Φ^2, since this would change the order of the action.

It is interesting to examine the form of these dual theories. The potential for A/B essentially gets mapped into a non-canonical kinetic term for the F/Φ (4-form/scalar) theory respectively. The canonical kinetic terms in the A/B picture, on the other hand, give rise to simple quadratic potential terms in the corresponding 4-form/scalar theories. This is important for several reasons. First, in this way an effective non-canonical scalar theory arises from a very simple three-form theory.[3] Second, this immediately tells us that all scalar models dual to the three-form share the same simple quadratic potential. This is particularly important for standard slow-roll inflation[4], where in the dual scalar picture the potential dominates over the kinetic terms. The fact that all models share the same dual scalar potential then implies that the form of the original three-form potential (which turns into a non-canonical kinetic term) is not important when computing, for example, the spectral index n_s. We will return to this point later.

Finally, one may wonder how one can come up with an effective single scalar theory dual to a three-form theory, which in principle possesses more physical degrees of freedom. However, starting with the most general canonical and minimally coupled 3-form action in 4D, as we do here, guarantees that such a dual single scalar field description always exists. This is the case, because 1) the canonical kinetic term for the three-form dualises to a Φ^2 potential 2) the three-form potential is a function of A^2 only, because in 4d this is the only covariant scalar combination that can be built from a 3-form - thus the potential only depends on one effective degree of freedom: A^2 and 3) these two degrees of freedom, Φ and A^2, are related via an equation of motion, leaving only one effective independent degree of freedom, thus explaining the existence of a dual single scalar description.

4.4 Non-Gaussianities

4.4.1 Correlation functions

We will now proceed to utilise our results thus far to compute observables for the three-form inflationary theory under consideration. Because our action is dual to that for a non-canonical scalar field, the calculation follows precisely that of Garriga & Mukhanov [20] to calculate the power spectrum, and Seery & Lidsey [37] and Chen et al. [80] to calculate the three-point function. Here we simply provide an overview of the main steps and the important results.

The 2-point correlation function for the curvature perturbation ζ is defined as

$$\langle \zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\rangle = (2\pi)^5\delta^3(\mathbf{k}_1+\mathbf{k}_2)\frac{P_\zeta}{2k_1^3}\,, \tag{4.23}$$

[3] The reader might have noticed that there are differences in factors and signs when comparing (4.10) and (4.19) with corresponding equations presented in [28]. The version presented here corrects a small number of typographical mistakes in that work.

[4] By this we here mean $\epsilon, \eta... \ll 1$ as well as requiring no rapidly varying speed of sound.

and is calculated using the second order action

$$S_2 = \int \mathrm{d}t\mathrm{d}^3x \left[a^3 \frac{\Sigma}{H^2}\dot{\zeta}^2 - a\epsilon(\partial\zeta)^2 \right] , \tag{4.24}$$

where the quantities c_s^2, Σ, λ (introduced in the previous section) are explicitly given in terms of derivatives of the Lagrangian for theories of the type $\mathcal{L} = P(X, \phi)$ by [37]

$$\begin{aligned} c_s^2 &= \frac{\mathcal{L}_{,X}}{\rho_{,X}} = \frac{\mathcal{L}_{,X}}{\mathcal{L}_{,X} + 2X\mathcal{L}_{,XX}} , \\ \Sigma &= X\mathcal{L}_{,X} + 2X^2\mathcal{L}_{,XX} = \frac{H^2\epsilon}{c_s^2} \\ \lambda &= X^2\mathcal{L}_{,XX} + \frac{2}{3}X^3\mathcal{L}_{,XXX} . \end{aligned} \tag{4.25}$$

Note that for the effective scalar theory (4.19) one can replace all derivatives of $\mathcal{L}$ with respect to X ($\mathcal{L}_{,X}, \mathcal{L}_{,XX}, ..$) with derivatives of the potential V expressed in terms of the scalar, since only terms coming from $V(A^2)$ depend on X.

As found by [20] the power spectrum is

$$P_\zeta \equiv \frac{1}{2\pi^2}k^3 \left|\zeta_k\right|^2 = \frac{1}{2(2\pi)^2\epsilon}\frac{H^2}{c_s M_{\rm Pl}^2}\bigg|_* , \tag{4.26}$$

where $*$ indicates that the expression is evaluated at horizon crossing $c_s k = aH$. The spectral index n_s is then given by

$$1 - n_s = 2\epsilon + \frac{\dot{\epsilon}}{\epsilon H} + \frac{\dot{c}_s}{c_s H} . \tag{4.27}$$

The slow-roll approximation (i.e. neglecting $\ddot{\chi}$ in the equation of motion for χ) implies that

$$\frac{\chi_{,N}}{\chi} \approx -\frac{2}{3}\frac{1}{\chi^2}\left(1 - \frac{3}{2}\chi^2\right)\epsilon. \tag{4.28}$$

Substituting this into the Friedmann equation we obtain [27]

$$\epsilon \equiv -\frac{\dot{H}}{H^2} \approx \frac{3}{2}\chi\frac{V_{,\chi}}{V}\left(1 - \frac{3}{2}\chi^2\right) . \tag{4.29}$$

Eliminating the term in brackets from the last two equations we can write

$$\frac{\chi_{,N}}{\chi} = -\frac{4}{9}\frac{1}{\chi^2}\frac{V}{\chi V_{,\chi}}\epsilon^2. \tag{4.30}$$

This expression allows us to easily compute $\dot{\epsilon}$ from the definition of ϵ to find $\dot{\epsilon}/\epsilon H = 2\epsilon + \mathcal{O}(\epsilon^2)$. Similarly, we find that $\dot{c}_s/c_s H \approx \mathcal{O}(\epsilon^2)$ and also $\dot{\lambda}/\lambda H \approx \mathcal{O}(\epsilon^2)$. This means that to first order

in the slow-roll parameters, the scalar spectral index is simply $n_s = 1 - 4\epsilon$.

The ratio of tensor to scalar perturbations was computed in Ref.[28] and shown to be related to the slow roll parameter ϵ as[5]

$$r = 16 c_s \epsilon, \tag{4.31}$$

mirroring the analogous expression for non-canonical scalar field models. One can also compute the spectral index for tensor perturbations n_t, finding $n_t = -2\epsilon$ to first order in slow-roll parameters as usual.

The third order action, as provided in the previous chapter, is needed to calculate the three-point correlation function. For convenience we recall that it is given by [37, 80]

$$\begin{aligned} S_3 &= \int \mathrm{d}t \mathrm{d}^3x \left\{ -a^3 \left[\Sigma \left(1 - \frac{1}{c_s^2} \right) + 2\lambda \right] \frac{\dot{\zeta}^3}{H^3} + \frac{a^3 \epsilon}{c_s^4} (\epsilon - 3 + 3c_s^2) \zeta \dot{\zeta}^2 \right. \\ &+ \frac{a\epsilon}{c_s^2} (\epsilon - 2\epsilon_s + 1 - c_s^2) \zeta (\partial\zeta)^2 - 2a \frac{\epsilon}{c_s^2} \dot{\zeta} (\partial\zeta)(\partial\sigma) \\ &+ \left. \frac{a^3 \epsilon}{2c_s^2} \frac{d}{dt} \left(\frac{\eta}{c_s^2} \right) \zeta^2 \dot{\zeta} + \frac{\epsilon}{2a} (\partial\zeta)(\partial\sigma)\partial^2\sigma + \frac{\epsilon}{4a} (\partial^2\zeta)(\partial\sigma)^2 + 2f(\zeta) \left. \frac{\delta L}{\delta \zeta} \right|_1 \right\}, \end{aligned} \tag{4.32}$$

where $f(\zeta)\delta L/\delta\zeta|_1$ indicates terms proportional to the functional derivative of the Lagrangian evaluated at first order in ζ. This would be zero if ζ was Gaussian. Such terms can be removed by a field redefinition (though one must recall the redefinition when forming correlation functions so as to form corrections of ζ itself). This form of the action thus identifies the relevant interaction vertices which contribute towards the three-point function.

At tree level in quantum field theory, and in the interaction picture, the In-In (equal time) three-point correlation function is given by the expression

$$\langle \zeta(t, \mathbf{k}_1) \zeta(t, \mathbf{k}_2) \zeta(t, \mathbf{k}_3) \rangle = -i \int_{t_0}^{t} \mathrm{d}t' \langle [\zeta(t, \mathbf{k}_1) \zeta(t, \mathbf{k}_2) \zeta(t, \mathbf{k}_3), H_{\mathrm{int}}(t')] \rangle \ , \tag{4.33}$$

where H_{int} is the Hamiltonian evaluated at third order in the perturbations and follows directly from (4.32). Vacuum expectation values are evaluated with respect to the interacting vacuum $|\Omega\rangle$. By convention the 3-point correlation function is parametrised by the amplitude $\mathcal{A}$.

$$\langle \zeta(\mathbf{k}_1) \zeta(\mathbf{k}_2) \zeta(\mathbf{k}_3) \rangle = (2\pi)^7 \delta^3(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) P_\zeta{}^2 \frac{1}{\Pi_j k_j^3} \mathcal{A} \, , \tag{4.34}$$

where, again by convention the power spectrum P_ζ in the above formula is calculated for the mode $K = k_1 + k_2 + k_3$. Evaluating (4.33), one can determine $\mathcal{A}$

[5] A more exact answer, fully taking into account that tensor and scalar modes freeze out at different times for models with $c_s \neq 1$, can be found in [39]. Here we just note that this extra effect means that the actual tensor-to-scalar ratio r is in fact smaller than naively expected from $r = 16 c_s \epsilon$ for $c_s < 1$.

In principle $\mathcal{A}$ is a general function of the three Fourier modes, which are related by the condition $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$ (within the slow-roll approximation the full form of $\mathcal{A}$ is given in Ref. [80]). For a given shape of non-Gaussianity, however, (see figure 4.3) the size of non-Gaussianity can be adequately characterised by a single-value measure $f_{\rm NL}$. For an equilateral shape (i.e. one peaking in the limit $k_1 \sim k_2 \sim k_3$), this can be defined as [66]

$$f_{\rm NL}^{\rm equil} = 30\frac{\mathcal{A}_{k_1=k_2=k_3}}{K^3}\,, \tag{4.35}$$

where amplitudes are matched at $k_1 = k_2 = k_3 = K/3$. Note that we follow the WMAP sign convention here, where positive $f_{\rm NL}$ physically corresponds to negative-skewness for the temperature fluctuations. As before the parameters controlling the overall size of $f_{\rm NL}^{\rm equil}$ are c_s, Σ, λ. Following [80] we may now compute $f_{\rm NL}^{\rm equil}$ in the slow-roll regime, finding the result at leading order to be[6]

$$f_{\rm NL}^{\rm equil} \approx \frac{5}{81}\left(\frac{1}{c_s^2} - 1 - 2\frac{\lambda}{\Sigma}\right) - \frac{35}{108}\left(\frac{1}{c_s^2} - 1\right) + \mathcal{O}\left(\epsilon, \frac{\epsilon}{c_s^2}, \epsilon\frac{\lambda}{\Sigma}, \frac{\dot{\lambda}}{\lambda H}\right). \tag{4.36}$$

4.4.2 A self-contained three-form formalism

Our primary aim in this section is to use the dual map to devise a self-contained description of perturbative properties at the three-form level. Since the two- and three-point correlation functions are controlled by parameters c_s, Σ, λ it will be useful to write down equivalent expressions in terms of the three-form variable χ. This turns out to be rather straightforward. In particular, we employ the expressions[7]

$$B^2 = -\chi^2, \qquad X \equiv -g^{\mu\nu}\nabla_\mu\Phi\nabla_\nu\Phi = -12^2B^2\left(V'(-6B^2)\right)^2\,, \tag{4.37}$$

which follow from the definition of the Hodge dual to A, Eq. (4.16), and from Eq. (4.18) respectively, to find

$$X = V_{,\chi}^2, \qquad \text{and} \qquad \frac{\partial\chi}{\partial X} = \frac{1}{2V_{,\chi}V_{,\chi\chi}}\,. \tag{4.38}$$

From the discussion of dualities, and in particular $\mathcal{L}_0$, we see that the $P(X,\phi)$ theory dual to our minimally coupled three-form theory was parametrised by

$$\begin{aligned} P(X,\phi) &= -\frac{1}{2}\phi^2 - 12B^2V'(-6B^2) - V(-6B^2) \\ &= -\frac{1}{2}\phi^2 + \chi V_{,\chi} - V(6\chi^2). \end{aligned} \tag{4.39}$$

[6] Note that (4.36) differs from the expression found by Chen et al. [80] by an overall sign. This is because Chen et al. use a sign convention opposite to that used by WMAP and throughout this paper.

[7] Note that the convention for X used in this chapter differs from that used in previous chapters by a factor of 1/2.

Differentiating this expression with respect to X and using (4.38) we find $P_{,X}$, $P_{,XX}$ and $P_{,XXX}$ which upon substitution into (4.25) and by replacing derivatives with respect to X with derivatives w.r.t. χ allows us to write

$$c_s^2 = \frac{V_{,\chi\chi}\chi}{V_{,\chi}}, \qquad \Sigma = \frac{V_{,\chi}^2}{2V_{,\chi\chi}}, \qquad \lambda = -\frac{V_{,\chi}^3 V_{,\chi\chi\chi}}{12V_{,\chi\chi}^3}. \tag{4.40}$$

This constitutes the result we aimed for, since it enables us to write the perturbed action solely in terms of the background dynamics of the three-form, and the perturbed quantity ζ. Comparing results with Kovisto & Nunes [28] we find that the expression for c_s matches exactly that found by them when working directly with the perturbed 3-form theory. (they did not work out Σ, λ since the focus of their paper does not extend beyond the power spectrum). Moreover, formally passing to the dual scalar theory calculating the perturbed action and writing this in terms of ζ, background quantities and c_s, *must* be equivalent to calculating the perturbed action working with the 'raw' 3-form action and writing this in terms of ζ and these quantities. This means that, once we have worked out c_s^2, Σ, λ in terms of three-form quantities, we can immediately extract non-Gaussian features from action (4.32). In other words, having expressed these quantities purely in terms of three-form quantities, we can now work out non-Gaussian signatures generated by any three-form theory

$$S = -\int d^4x\sqrt{-g}\left(R - \frac{1}{48}F^2 - V(A^2)\right). \tag{4.41}$$

directly without the need to explicitly refer to the dual description. The power of this approach is that we can also readily probe higher order statistics, such as the trispectrum, since the quartic action has already been calculated for $P(X,\phi)$ theories [129, 130]. This involves the quantity Π, defined in Ref. [130], which for completeness, we calculate to be

$$\Pi = \frac{1}{40}\frac{V_{,\chi}^3 V_{,\chi\chi\chi}}{V_{,\chi\chi}^3} + \frac{3}{40}\frac{V_{,\chi}^4 V_{,\chi\chi\chi}^2}{V_{,\chi\chi}^5} - \frac{1}{40}\frac{V_{,\chi}^4 V_{,\chi\chi\chi\chi}}{V_{,\chi\chi}^4}, \tag{4.42}$$

in terms of the three-form quantity χ. In [5] the cubic action is worked out in the three-form picture explicitly, establishing the equivalence of cubic actions for the three-form theory and for the equivalent scalar field theory as argued for here.

4.4.3 Example I: A power-law potential

We now consider some concrete examples, and begin by considering a three-form model self-interacting through a simple power-law potential

$$\mathcal{L} = -\frac{1}{48}F^2 - V_0 A^{2p}\,, \tag{4.43}$$

that is, $V(A^2) = V_0 A^{2p} = V_0\left(6\chi^2\right)^p$, where p is a constant.

This is a special example, where the equivalent $P(X,\phi)$ theory is relatively simple, and is given by

$$\mathcal{L}_\phi = (2p-1)\left(\frac{1}{V_0}\right)^{1/(2p-1)}\left(\frac{X}{24p^2}\right)^{\frac{p}{2p-1}} - \frac{1}{2}\phi^2, \tag{4.44}$$

where one now sees explicitly that the 3-form potential has been mapped into a non-canonical kinetic term for the effective scalar ϕ. As long as $2p-1 \neq 0$, one can check that equations (4.25) and (4.40) equivalently yield

$$\begin{aligned} c_s^2 &= 2p-1\,, \\ \frac{\lambda}{\Sigma} &= \frac{1}{3}\frac{1-p}{2p-1}\,, \end{aligned} \tag{4.45}$$

and we can immediately make observational predictions for the theory, finding several interesting results.

Here both the speed of sound and λ/Σ are constant, and we have $n_s - 1 = -4\epsilon$. The constancy of the speed of sound has the interesting consequence that the observational requirement of obtaining a (near)-scale-invariant spectral index n_s forces ϵ to be close to zero and slowly-varying here (compare (4.27) and also [66]). For this power-law model, slow-roll therefore becomes an observational requirement in contrast to models with varying speed of sound.

We now need to calculate this quantity N e-folds before the end of inflation. It was shown in [27] that for a power law potential, the value of the field at this time can be estimated to be

$$\chi_N^2 = \frac{2}{3} - \frac{4}{18p}\frac{1}{1+2N}, \tag{4.46}$$

which upon substitution in (4.29) gives

$$\epsilon_N \approx \frac{1}{1+2N}. \tag{4.47}$$

Assuming that $N \approx 60$ is required to solve the horizon problem, we predict that the spectral

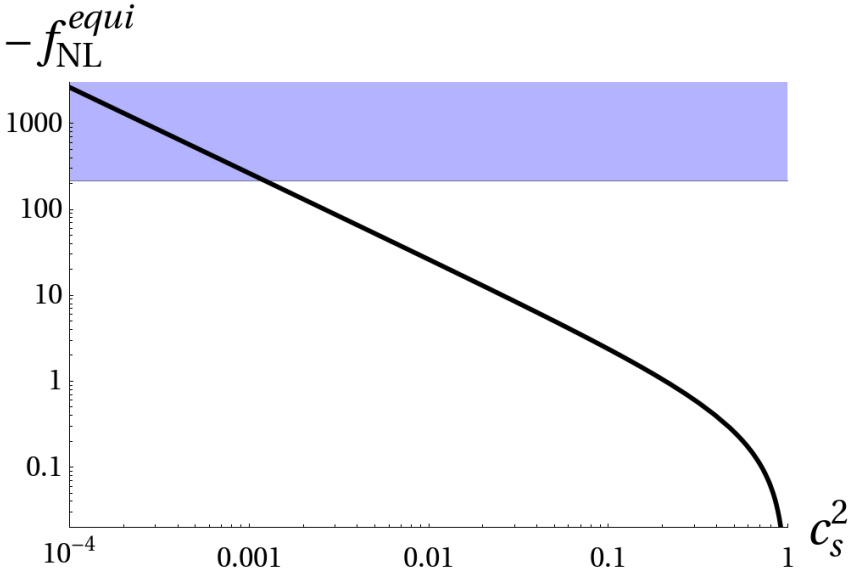


Figure 4.1: Dependence of $f_{\rm NL}^{equil}$ on c_s^2 for the power law potential $V = V_0(6\chi^2)^p$ and $N = 60$. A large and generically negative non-Gaussian amplitude is found. The shaded region is disallowed by the WMAP 2σ bound $-214 < f_{\rm NL}^{equil} < 266$.[9] We recall that $c_s^2 = 2p - 1$ for this model.

index on observationally relevant scales for a three-form with power law potential is

$$n_s \approx 0.97, \tag{4.48}$$

and independent of the value of the power p. Naturally the spectral index will be closer to scale-invariant the longer inflation lasts. Substituting for c_s^2 and λ/Σ into (4.36) we obtain the dependence of $f_{NL}^{\rm equil}$ on c_s^2 illustrated in Fig. 4.1. It clearly shows that a small speed of sound leads to a large $f_{NL}^{\rm equil}$ as expected. Substituting for the speed of sound for a power law potential in Eqs.(4.31) and (4.36) we can relate $f_{NL}^{\rm equil}$ and r. Figure 4.2 illustrates how these two quantities are related, and shows the region of the parameter space (r, f_{NL}) that is allowed given current bounds.

Phenomenologically we find that no large non-Gaussian amplitude of the enfolded or orthogonal types (which have peaks in the folded limit $k_1 \sim k_2 \sim 2k_3$) can be generated here in contrast to generic single field inflation models. To see why, it is useful to notice that one may express λ/Σ as [37, 66]

$$\frac{\lambda}{\Sigma} = \frac{1}{6}\left(\frac{2f_X + 1}{c_s^2} - 1\right) \tag{4.49}$$

where

$$f_X = \frac{\epsilon\epsilon_s}{3\epsilon_X}, \qquad \epsilon_s = \frac{\dot{c_s}}{Hc_s}, \qquad \epsilon_X = -\frac{\dot{X}}{H^2}\frac{\partial H}{\partial X}. \tag{4.50}$$

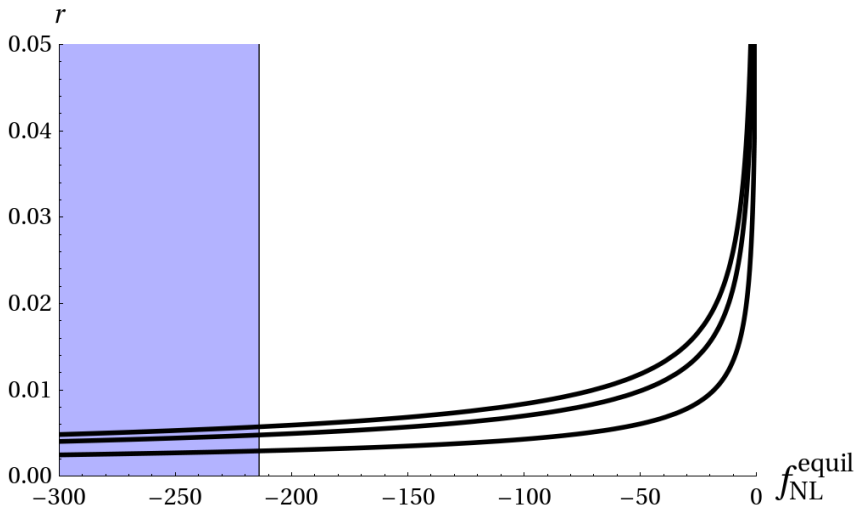


Figure 4.2: The solid lines show how the parameters $f_{NL}^{\rm equil}$ and r are related to each other for the power law potential, $V = V_0(6\chi^2)^p$ and for $N = 50, 60, 80$ from top to bottom. The shaded region is disallowed by the WMAP 2σ bound $-214 < f_{\rm NL}^{equil} < 266$.[9]

Now, since the speed of sound is constant for the power-law model, ϵ_s, f_X are identically zero here. A large and predominantly orthogonal (or enfolded) amplitude, however, requires a negative and non-zero f_X (assuming positive ϵ and $0 \leq c_s \leq 1$)[82][8]. As such the non-Gaussian shape found here is always predominantly equilateral. Furthermore, since parameters λ and Σ are simply related to c_s via (4.49), the non-Gaussian amplitude is completely controlled by c_s^2 for a power-law potential. This means that $f_{\rm NL}^{\rm equil} \sim \mathcal{O}(c_s^{-2})$ for subluminal and positive speed of sound c_s, i.e. $p \in [1/2, 1]$. Figure 4.3 illustrates these points. Finally note that, analogous to e.g. DBI inflation, any sizable level of non-Gaussianity has a negative $f_{\rm NL}^{\rm equil}$ associated with it here.

4.4.4 Example II: Exponential potential

Let us now consider an alternative potential. Even though it is fairly simple in its analytical form, the exponential potential, $V = V_0 \exp(\beta A^2) = V_0 \exp(6\beta\chi^2)$, becomes very complex when written as a $P(X, \phi)$ theory,

$$\mathcal{L}_0 = (W(x) - 1)\, V_0 \exp\left(\frac{1}{2}W(x)\right) - \frac{1}{2}\phi^2, \tag{4.51}$$

[8] In general $P(X, \phi)$ inflation models a predominantly orthogonal or folded shape can be generated by finely balancing the contributions from the $\dot{\zeta}^3$ interaction vertex (which depends on λ/Σ) against the other vertices such that the generically predominant equilateral shape contributions cancel out. In this way the otherwise subdominant orthogonal or folded configurations are brought out. As stated above, this delicate cancellation relies on f_X-dependent contributions though, whereas $f_X = 0$ identically for the power-law example given.

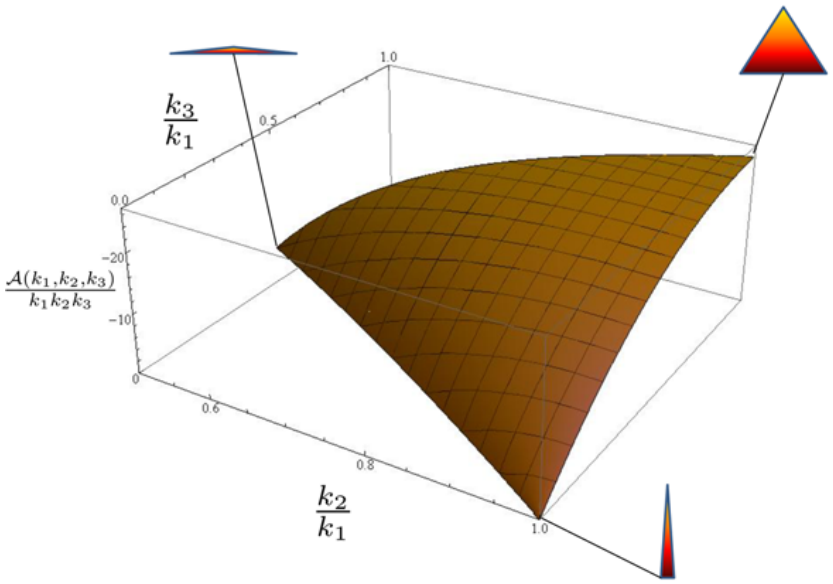


Figure 4.3: The dimensionless bispectrum $\mathcal{A}(k_1, k_2, k_3)/k_1k_2k_3$ in the slow-roll limit for the power law potential, $V = V_0(6\chi^2)^p$ with $p = 0.505$, corresponding to $c_s = 0.1$, and with $N = 60$. Triangular shapes denote the equilateral, squeezed/local and enfolded limit clockwise from top right. A predominantly equilateral shape is found.

where $W(x)$ is the Lambert-W function and $x = X/12\beta V_0^2$. Dealing with this model in the $P(X, \phi)$ description is therefore very difficult. This is a good example showing that performing the calculations in the original three-form theory is far simpler than going to the $P(X, \phi)$ description.

Inflation ends when $\epsilon \approx 1$ which for this potential takes place when

$$\chi_e^2 = \frac{1}{3} + \frac{1}{3}\left(1 - \frac{1}{3\beta}\right)^{1/2}. \tag{4.52}$$

This expression implies that inflation only ends if $\beta > 1/3$. In particular, if $\beta \gg 1/3$ we can approximate

$$\chi_e^2 \approx \frac{2}{3} - \frac{1}{18\beta}. \tag{4.53}$$

In this case, the solution of (4.28) and (4.29) is non-analytical and we must resort to an approximation. Defining $\chi^2 = 2/3 - y$, since we know that inflation occurs very close to 2/3, we obtain

$$\chi_N^2 = \frac{2}{3} - \frac{1}{18\beta}\frac{1}{1 + \sqrt{6}N}. \tag{4.54}$$

The slow-roll parameter N e-folds before the end of inflation is then $\epsilon_N \approx 1/(1 + \sqrt{6}N)$. Here,

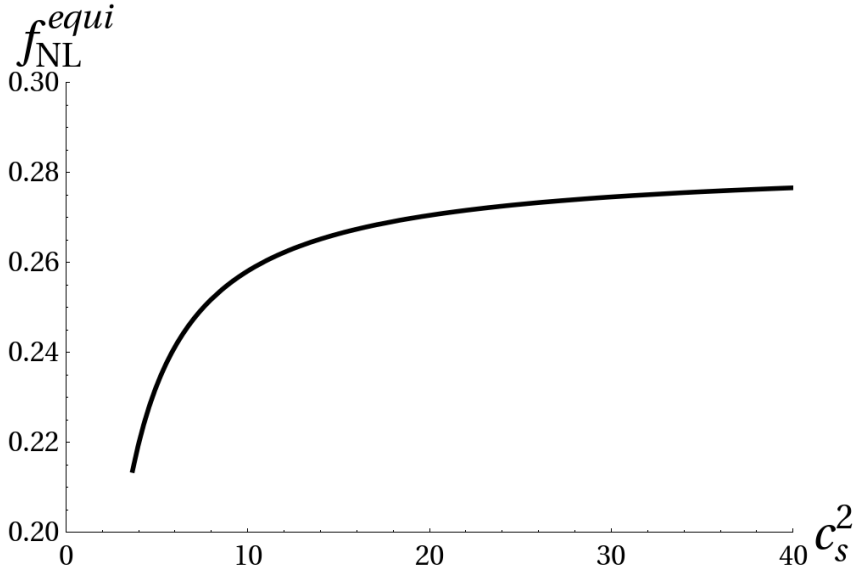


Figure 4.4: Dependence of $f_{\rm NL}^{equil}$ on c_s^2 for the exponential potential, $V = V_0 \exp(6\beta\chi^2)$ and $N = 60$. A small and generically positive non-Gaussian amplitude is found. Notice the lower bound $c_s^2 \gtrsim 11/3$ for this model.

$c_s^2 = 1 + 12\beta\chi^2$ and it can be verified that $\dot{c}_s/c_s H = -\epsilon^2/\chi^2 \sim \mathcal{O}(\epsilon^2)$, hence, for $N = 60$, we obtain for this model

$$n_s \approx 0.97, \tag{4.55}$$

independent of the value of β. The fact that the value of the scalar spectral index is independent of parameter β, or p in the case of the power law potential, and we obtain the same value in both examples, should not come as a surprise. The reason becomes clear when considering the dual $P(X, \phi)$ theory, which for any three-form potential, has the same quadratic scalar potential. Since we are in a slow-roll regime, the functional form of the kinetic term is not important with respect to the potential which leads to identical results for the spectral index.

It was found above that the choice of the exponential three-form potential offers an exit from inflation only if $\beta > 1/3$. Taking $\chi_N^2 \approx 2/3$, this puts the lower bound, $c_s^2 \gtrsim 11/3$, which means that the speed of sound is superluminous.

The amplitude of the three-point function is controlled by c_s^2 and λ/Σ, which are given by

$$c_s^2 = 1 + 12\beta\chi^2, \qquad \frac{\lambda}{\Sigma} = -\frac{6\beta^2\chi^2(1 + 4\beta\chi^2)}{c_s^4}. \tag{4.56}$$

Varying β and substituting for c_s^2 and λ/Σ N e-folds before the end of inflation into (4.36), we obtain the dependence of $f_{\rm NL}^{equil}$ on c_s^2 which we show in Fig. 4.4 for $N = 60$.

In Fig. 4.5 we see how r and $f_{\rm NL}^{equil}$ relate to each other. In particular, we observe that the

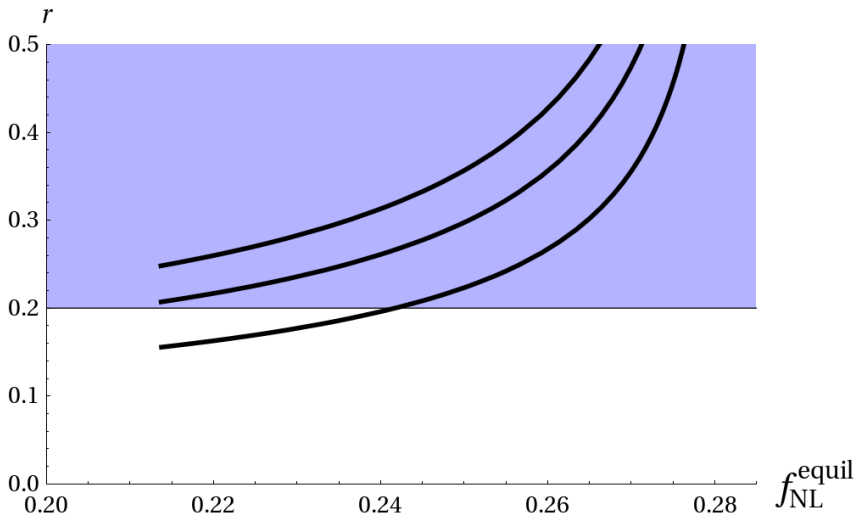


Figure 4.5: The solid lines show how the parameters $f_{NL}^{\rm equil}$ and r are related to each other for the exponential potential, $V = V_0 \exp(6\beta\chi^2)$ and for $N = 50, 60, 80$ from top to bottom. The shaded region is disallowed by the WMAP 2σ bound $r \lesssim 0.2$.[9]. End points for the solid lines at $f_{\rm NL}^{equil} \sim 0.215$ correspond to the lower bound $c_s^2 \gtrsim 11/3$ for this model.

large values of c_s^2 render this model disfavored by current bounds on the ratio of tensor to scalar perturbations, $r \lesssim 0.2$ for $N = 60$.[9] Of course, if we allow for more e-folds of inflation this reduces the relevant ϵ and hence r. A crude estimate for the minimal amount of inflation to bring this model into agreement with current 2σ bounds on r may be obtained by assuming a speed of sound right at the lower bound of $c_s \gtrsim \sqrt{11/3}$ and $r \sim 0.2$, yielding $N \gtrsim 62$ for the exponential model considered here. Finally we see in Fig. 4.6 that the sign of $f_{\rm NL}^{equil} \sim 0.2$ is positive, albeit with a rather small amplitude. The non-Gaussian shape here is predominantly equilateral, but also picks up contributions in the enfolded/orthogonal limit.[10]

4.5 Summary

In this chapter we have considered 3-form theories of inflation and their non-Gaussian features. The salient results may be summarized as follows.

- We explored dualities between various n-form models in four dimensional space-times. In particular these allowed linking the 3-form theory under consideration to an effective scalar-tensor description with a $P(X, \phi)$ action, due to the reduction of the effective number of *dof*

[9] Note that this bound significantly depends on the assumption of no running spectral index n_s. If a running n_s is allowed the bound weakens to $r < 0.49$.

[10] Note that considering departure from slow-roll could modify this statement, e.g. potentially rendering the shape predominantly enfolded or orthogonal[1, 2]. For further details on fast-roll corrections also see[67, 69].

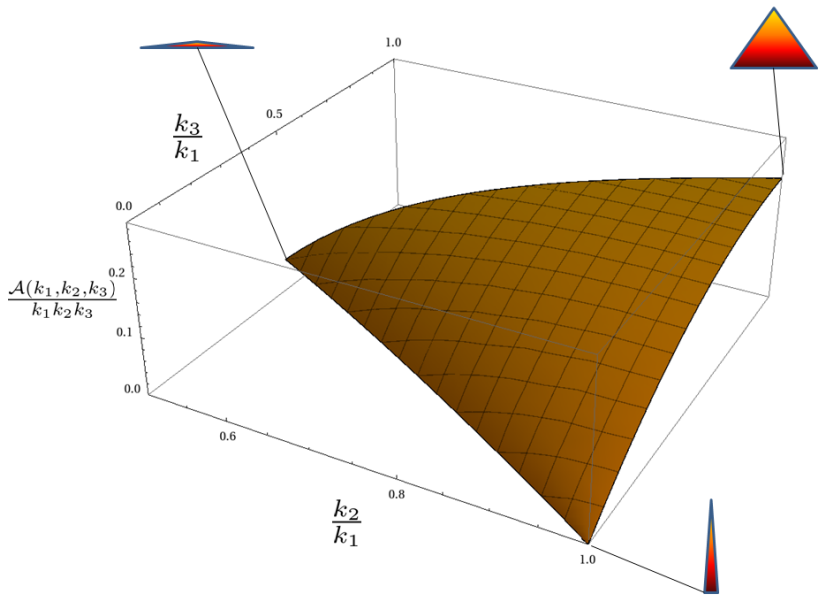


Figure 4.6: The dimensionless bispectrum $\mathcal{A}(k_1, k_2, k_3)/k_1k_2k_3$ in the slow-roll limit for the exponential potential, $V = V_0 \exp(6\beta\chi^2)$ with $\beta = 1$, corresponding to $c_s \approx 3$ and for $N = 60$. Triangular shapes denote the equilateral, squeezed/local and enfolded limit clockwise from top right. $f_{\rm NL}^{equil} \sim 0.2$ and a predominantly equilateral shape is found.

from requiring isotropy and homogeneity. The three-form theory can therefore effectively be described by a disformal bimetric setup. Since the three-form potential is mapped to the scalar fields kinetic term in a non-trivial way, such setups generically result in highly non-canonical $P(X, \phi)$ theories.

- Using these dualities the cubic action necessary to compute non-Gaussian features is simply that of generic $P(X, \phi)$ theories. However, we can use the mapping between dual theories to express the parameters controlling the non-Gaussian amplitude, such as c_s, Σ, λ, in terms of three-form quantities. This allows us to rewrite the cubic action in terms of the background dynamics of the three-form and the perturbed quantity ζ only. Hence we can capture all non-Gaussian features of the theory in a self-contained three-form language.

- Finally, using the tools developed, we explored the observational and particularly non-Gaussian features of two example three-form models:

  – First, we investigated three-form inflation with a power-law potential. This simple setup produces a constant speed of sound and generically equilateral non-Gaussian shapes with negative $f_{\rm NL}^{equil}$, similar to DBI-type inflation. Observational limits on this form of non-Gaussianity constrain this model.

  - Secondly, we explored the phenomenology for a three-form inflationary model with an exponential potential. We found the non-Gaussianity was once again predominantly of equilateral shape, but in this case was unobservably small. We found, however, that the model can be constrained from limits on the tensor-to-scalar ratio r.

- An interesting general result is that the spectral index for all three-form inflationary models is $n_s \approx 0.97$ to leading order in slow-roll when 60 e-folds of inflation occur. This was shown explicitly for the two models at hand, but will hold much more generally. The value of n_s is therefore uniquely predicted once N, the number of e-folds of inflation, is specified, independent of the exact form of the three-form potential and as long as $\epsilon \ll 1$.

Part III

Dark energy and modified gravity

Chapter 5

Derivative Chameleons

5.1 Introduction

We now turn our attention to scalar-tensor theories in the context of modified gravity/dark energy theories. As discussed in chapter 1, modified gravity theories (see e.g. [14, 131] for reviews) have enjoyed continued interest over the past few decades. With our focus being on disformally motivated scalar-tensor theories, the notion of screening mechanisms - how a light scalar degree of freedom ϕ can act as dark energy on cosmological scales while being shielded in dense environments such as on earth - has turned out to be especially useful in this context. We discussed implementations of such a mechanism in 1.3. They include the following: 1) The chameleon model [15, 16], where a density dependent mass is generated and the field ϕ becomes too massive for detection in dense environments. 2) Vainshtein screened setups [54, 55, 56] such as DGP [17] and Galileon [29]/Horndeski [33] models, where non-linear interactions of ϕ lead to strongly coupled dynamics. A density-dependent (classical) renormalization of the kinetic energy there results in an effectively decoupled scalar in dense environments. 3) Symmetron models [57, 58, 59], where a scalar ϕ is coupled to matter with a coupling strength proportional to the vacuum expectation value of ϕ. This in turn depends on the ambient density, so that the scalar effectively decouples in high-density regions. All these mechanisms reconcile the existence of a light cosmological scalar with tight fifth force constraints on solar-system scales [26].

In this chapter we wish to focus on chameleon models and potential extensions thereof. Chameleon phenomenology has already been studied extensively [132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145] for variants of the model introduced by [15, 16]. Typically such theories are built by universally coupling matter to a metric conformally related to the Einstein metric $g_{\mu\nu}$. Making this bimetric structure explicit we here investigate whether there are any interesting and qualitatively new implementations of the chameleon mechanism that arise when going beyond the simple conformal relationships considered so far. Can the chameleon

effect occur in new guises for generic 4D effective field theories? In other words, we will be asking two questions: What forms can the conformal relation generically take? And does it have to be conformal or are there more general bimetric structures that can produce chameleonic phenomenology?

A systematic way of undertaking this investigation is to proceed as we have done above in section 1.1 and construct a generic matter metric as a function of the Einstein metric $g_{\mu\nu}$, ϕ and derivatives of the field $\partial^n\phi$. By showing how what we dub *derivative chameleons* arise in this framework, this chapter aims to illustrate how such an approach unveils qualitatively new constructions. More specifically we show that derivative chameleons naturally give rise to a new mass-altering mechanism, which changes the mass of oscillations around an effective potential minimum. The mass-lifting branch may help in ensuring ϕ can escape detection in fifth force experiments and consequently may be of use in alleviating fine-tuning constraints for chameleon models. However, we also show that care needs to be taken in order to avoid ghost-like instabilities for mass-lifting solutions. The mass-lowering branch of solutions, on the other hand, is generically stable. Importantly the new mechanism works for purely derivative conformal factors too, opening up the exciting possibility of having a chameleon mechanism which comes endowed with a shift-symmetry in the field $\phi \to \phi + c$, offering better protection from quantum-corrections. Furthermore we discuss modifications to the radial solutions around spherical matter sources and modifications to the thin-shell mechanism. We also point out why non-conformal geometries are not expected to display chameleon-screening, establishing a no-go theorem for so-called disformal geometries.

The plan for the chapter is as follows. In section 5.2 we review the standard chameleon picture and how this gives rise to an effective potential with a large mass for oscillations around the effective minimum. Section 5.3 then reviews the bimetric framework, explicitly formulating metric relations which we then use to construct new types of chameleon models. In doing so we focus on conformally related metrics, pointing out the phenomenology of the resulting generalized chameleon model, in particular the new mass-altering mechanism. In 5.4 this is followed up by presenting a few simple concrete examples that implement this new-found mechanism and show how it can be realized for purely derivative conformal factors. Modifications to thin-shell screening are then discussed while investigating radial solutions around a massive source in section 5.5. In 5.6 we go beyond conformal relationships, arguing that their natural extension - disformally related geometries - cannot produce chameleon phenomenology (they do, however, naturally generate Vainshtein-type screening solutions). Finally we conclude in 5.7.

5.2 Conformal chameleons I: The minimal theory

We already reviewed the standard chameleon theory in section 1.3.2. Here we briefly summarize the essential features with an eye on extending the standard setup in the following section. Chameleon models [15, 16] are a particularly simple conformal bimetric theory, where the conformal factor $A(\phi)$ is a function of the chameleon field ϕ only. The original model proposed by [15, 16] has action

$$\mathcal{S} = \int d^4x\sqrt{g}\left(\frac{M^2}{2}R + X - V(\phi)\right) + \mathcal{S}_m\left(A^2(\phi)g_{\mu\nu}, \Psi_i\right), \tag{5.1}$$

Working out the corresponding equations of motion we find

$$\Box\phi = V_{,\phi} - A^3(\phi)A_{,\phi}\tilde{T}, \tag{5.2}$$

where $\tilde{T} = \tilde{T}^{\mu\nu}\tilde{g}_{\mu\nu}$ is the trace of the matter stress-energy tensor which is defined as usual

$$\tilde{T}^{\mu\nu} = \frac{2}{\sqrt{\tilde{g}}}\frac{\delta\left(\sqrt{\tilde{g}}\mathcal{L}_m\right)}{\delta\tilde{g}_{\mu\nu}}. \tag{5.3}$$

Importantly it is therefore defined with respect to the matter ("Jordan frame") metric $\tilde{g}_{\mu\nu}$. Note that $\tilde{T}^{\mu\nu}$ is covariantly conserved with respect to the matter metric $\tilde{\nabla}_\mu\tilde{T}^{\mu\nu} = 0$, since matter minimally couples to that metric.[1]

For a pressureless, non-relativistic source, the only non-vanishing component of the stress-energy tensor is $\tilde{T}^0_0 = -\tilde{\rho}$. The equation of motion (5.2) can therefore be written as

$$\Box\phi = V_{,\phi} + A_{,\phi}\hat{\rho}, \tag{5.4}$$

where $\hat{\rho}$ is a conserved quantity in the Einstein frame (in particular it is independent of ϕ). This follows from the covariant conservation of $\tilde{T}^{\mu\nu}$. Requiring $\tilde{\nabla}_\mu\tilde{T}^{\mu\nu} = 0$ directly leads to a corresponding conserved quantity in the Einstein frame, which is $\hat{\rho} = A^3\tilde{\rho} = A^{-1}\rho$. Note that this is not the energy density of matter in the Einstein frame, which we denote by ρ and which is given by $\rho = A^4\tilde{\rho}$. Finally we can use the fact that $\hat{\rho}$ is independent of ϕ and integrate up the equation of motion to obtain an effective potential for ϕ

$$V_{\text{eff}}(\phi) = V(\phi) + \hat{\rho}A(\phi). \tag{5.5}$$

Note that, even though this equation gives the impression that the sign of $A(\phi)$ is relevant, this

[1] When mapping $\tilde{T}^{\mu\nu}$ into the Einstein frame one finds $\tilde{T}^{\mu\nu} = A^{-6}(\phi)T^{\mu\nu}$, so that $T^{\mu\nu}$ is *not* covariantly conserved in the Einstein frame $\nabla_\mu T^{\mu\nu} \neq 0$.

is not the case as would be expected from a theory that only depends on $A^2(\phi)$. We can confirm this by noting that equation (5.5) depends on $\hat{\rho}A(\phi) = A^4\tilde{\rho}$. As long as the energy density is well-defined in the matter frame $\tilde{g}_{\mu\nu}$, there is consequently no sign ambiguity introduced related to the sign of $A(\phi)$.

A particularly interesting example is when we start with a runaway potential $V(\phi)$, e.g. $V(\phi) = M_{Pl}^{n+4}/\phi^n$ as desirable from the point of view of quintessence models [60]. This ensures that, in the limit when we can ignore the matter action $\mathcal{S}_m$ (a low-density environment with $\hat{\rho} \to 0$ in the language set out above), we recover a quintessence-like solution which leads to accelerated expansion of space-time. In analogy with inflation a slow-rolling light scalar field drives accelerated cosmic expansion. But in regions of high density this behavior changes, because of the $\hat{\rho}A(\phi)$ term in the effective potential. From (5.4) we can see that, for non-zero $\hat{\rho}$, V_{eff} can acquire a minimum $\phi_{\min}$, subject to a suitably chosen conformal factor $A(\phi)$. Specifically we require

$$V_{eff,\phi}(\phi_{\min}) = V_{,\phi}(\phi_{min}) + A_{,\phi}(\phi_{min})\hat{\rho} = 0, \tag{5.6}$$

Oscillations around the minimum ϕ_{min} of the effective potential then acquire a mass

$$m^2 \equiv V_{eff,\phi\phi}(\phi_{\min}) = V_{,\phi\phi}(\phi_{\min}) + \hat{\rho}A_{,\phi\phi}(\phi_{\min}). \tag{5.7}$$

Interestingly therefore, even though ϕ naively does not possess a mass term at all, the conformal factor $A(\phi)$ causes the field to become massive in dense environments. This is the essence of the chameleon mechanism: An environmentally-dependent way of generating a large mass for an otherwise very light scalar ϕ. This reconciles a model such as (5.1) with fifth force constraints, since ϕ becomes too heavy for detection in laboratory experiments on earth, yet can act as dark energy on large scales.[2]

The chameleon mechanism therefore hinges on the existence of two conformally related metrics $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$, respectively used to construct the gravitational part of the action (the Ricci scalar) and the matter part of the action (i.e. matter fields Ψ_i are minimally coupled to $\tilde{g}_{\mu\nu}$). Having introduced the minimal (standard) chameleon model in this section, the question we now want to answer in this chapter is: What happens when this relationship is modified? To be more specific, are there other classes of scalar-tensor theories that offer qualitatively distinct implementations of the chameleon mechanism (which we take to be an environmentally dependent generation of mass for an otherwise light cosmological scalar ϕ)?

[2]Note that we have said nothing about the thin-shell mechanism so far. This mechanism is important in order to ensure the viability of the chameleon mechanism in that it suppresses fifth force modifications to e.g. planetary orbits. We will review this mechanism in detail in section 5.5.

5.3 Conformal chameleons II: Derivative setups

5.3.1 Derivative bimetric setups

In this section we want to lay out the problem in more general terms without imposing as restrictive a metric relation as in section 5.2. Consequently let us start with a schematic action of the following form

$$\mathcal{S} = \int d^4x\sqrt{g}\frac{M^2}{2}R + \mathcal{S}_m\left(\tilde{g}_{\mu\nu}, \Psi_i\right) + \mathcal{S}_\phi \tag{5.8}$$

where Ψ_i are matter fields minimally coupled to $\tilde{g}_{\mu\nu}$ as before and $\mathcal{S}_\phi$ denotes an action giving the scalar field ϕ dynamics of its own. We emphasize that there is no a priori requirement constraining S_ϕ to be formed with either $g_{\mu\nu}$ or $\tilde{g}_{\mu\nu}$. To enable comparison with the existing chameleon literature we fix the form of $\mathcal{S}_\phi$ such that

$$\mathcal{S} = \int d^4x\sqrt{g}\left(\frac{M^2}{2}R + X - V(\phi)\right) + \mathcal{S}_m\left(\tilde{g}_{\mu\nu}, \Psi_i\right), \tag{5.9}$$

i.e. $\mathcal{S}_\phi$ equips ϕ with a canonical kinetic term and a potential minimally coupled to $g_{\mu\nu}$. In order to investigate (5.9) it now becomes necessary to specify how $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$ are related, and in particular how ϕ enters this relation. For this it will be useful to schematically write (5.9) as

$$\mathcal{S} = \int d^4x\sqrt{g}\mathcal{L}_E\left(g_{\mu\nu}, \phi\right) + \int d^4x\sqrt{\tilde{g}}\mathcal{L}_m\left(\tilde{g}_{\mu\nu}, \Psi_i\right). \tag{5.10}$$

In the previous section 5.2 gravity and matter metrics were conformally related by $\tilde{g}_{\mu\nu} = A^2(\phi)g_{\mu\nu}$. Here we are interested in investigating more general metric relations. To this end we recall the disformal relationship where $\tilde{g}_{\mu\nu}$ and $g_{\mu\nu}$ are "disformally" related by

$$\tilde{g}_{\mu\nu} = A^2(\phi, X)g_{\mu\nu} + B^2(\phi, X)\partial_\mu\phi\partial_\nu\phi, \tag{5.11}$$

where we again recognize X as ϕ's kinetic term. In this chapter our focus lies on considering the conformal subset of this full disformal relation. In other words, we will reduce the bimetric relationship (5.11) to the simple case $B(\phi, X) = 0$. This is the most general purely conformal relation permitted by (5.11). As such the conformal factor $A^2(\phi, X)$ is an arbitrary function of the field ϕ and the first higher order coordinate invariant X, ϕ's kinetic term, and we have

$$\tilde{g}_{\mu\nu} = A^2(\phi, X)g_{\mu\nu}. \tag{5.12}$$

This is in contrast to the minimal chameleon model, where the conformal factor is a function of ϕ only. As such this means the full action under consideration is that of a generalized conformal

chameleon

$$\mathcal{S} = \int d^4x\sqrt{g}\left(\frac{M^2}{2}R + X - V(\phi)\right) + \mathcal{S}_m\left(A^2(\phi, X)g_{\mu\nu}, \Psi_i\right). \tag{5.13}$$

The relation between matter stress-energy tensors in different frames straightforwardly generalizes to $\tilde{T}^{\mu\nu} = A^{-6}(\phi, X)T^{\mu\nu}$. We will now first compute the associated equations of motion for ϕ and then comment on the impact a generalized conformal factor has on the chameleon mechanism.

5.3.2 Equation of motion and effective potential

With a matter stress-energy tensor defined as in (5.3) and metrics related by (5.11), we can write down the equations of motion for the general action (5.9). Varying (5.9) we find

$$\frac{\partial\mathcal{L}_E}{\partial\phi} + \frac{1}{2}\sqrt{\frac{\tilde{g}}{g}}\tilde{T}^{\mu\nu}\frac{\partial\tilde{g}_{\mu\nu}}{\partial\phi} = \nabla_\alpha\frac{\partial\mathcal{L}_E}{\partial\phi_{,\alpha}} + \frac{1}{2}\frac{\partial\tilde{g}_{\mu\nu}}{\partial\phi_{,\alpha}}\nabla_\alpha\left(\sqrt{\frac{\tilde{g}}{g}}\tilde{T}^{\mu\nu}\right) + \frac{1}{2}\sqrt{\frac{\tilde{g}}{g}}\tilde{T}^{\mu\nu}\nabla_\alpha\frac{\partial\tilde{g}_{\mu\nu}}{\partial\phi_{,\alpha}}. \tag{5.14}$$

Note that these expressions in fact neither assume anything about the form of $\tilde{T}^{\mu\nu}$ nor rely on a conformal relationship between $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$ as considered above. It is, however, worth emphasizing that $\tilde{T}^{\mu\nu}$ is explicitly a stress energy tensor not including any contributions from $\mathcal{S}_\phi$ in (5.8), so it does not satisfy the Einstein equations by itself. Instead, after having mapped all quantities into the Einstein frame, one finds $G_{\mu\nu} = 8\pi G\left(T^{(\text{matter})}_{\mu\nu} + T^{(\phi)}_{\mu\nu}\right)$.

Specializing to the conformal case with $\tilde{g}_{\mu\nu} = A^2(\phi, X)g_{\mu\nu}$, the equation of motion for ϕ that follows from (5.13) can be written in the following form using (5.14)

$$-V_{,\phi} + A_{,\phi}A^3\tilde{T} = -\Box\phi + \sum_i \mathcal{J}_i, \tag{5.15}$$

where $\Box \equiv \nabla^\mu\nabla_\mu$ and

$$\begin{aligned}
\mathcal{J}_1 &= -A_{,X}A^3\tilde{T}\Box\phi \\
\mathcal{J}_2 &= 2A^2\tilde{T}\left(AA_{,X\phi} + 3A_{,X}A_{,\phi}\right)X \\
\mathcal{J}_3 &= A^2\tilde{T}\left(AA_{,XX} + 3A^2_{,X}\right)\Pi \\
\mathcal{J}_4 &= -A^3A_{,X}\tilde{g}_{\mu\nu}\partial^\alpha\phi\tilde{\nabla}_\alpha\tilde{T}^{\mu\nu},
\end{aligned} \tag{5.16}$$

and we have defined $\Pi \equiv \nabla_\mu\nabla_\nu\phi\nabla^\mu\phi\nabla^\nu\phi = \nabla_\mu\partial_\nu\phi\partial^\mu\phi\partial^\nu\phi$. We have assumed that $\tilde{T}^{\mu\nu}$ is symmetric and expressed derivatives acting on $\tilde{T}^{\mu\nu}$ in terms of matter frame variables to avoid mixing operators and variables defined for different frames. More explicitly this means using

$$\nabla_\alpha\tilde{T}^{\mu\nu} = \tilde{\nabla}_\alpha\tilde{T}^{\mu\nu} - 2\Gamma^{(\mu}_{\beta\alpha}\tilde{T}^{\nu)\beta}, \tag{5.17}$$

where $\Gamma^{\mu}_{\beta\alpha}$ is the connection associated with transformations between $\tilde{g}_{\mu\nu}$ ($\sim$ Jordan) and $g_{\mu\nu}$ (Einstein) frames (for details please see the appendix). Additionally there are two further types of terms one might naively expect to arise from varying (5.13)

$$\begin{aligned} \mathcal{J}_5 &\propto A^4 A_{,X}^2 \tilde{T}^{\beta\nu} \Pi_{\nu\rho} \partial_\beta \phi \partial^\rho \phi \\ \mathcal{J}_6 &\propto A^4 A_{,X} A_{,\phi} \partial_\beta \phi \partial_\nu \phi \tilde{T}^{\beta\nu}, \end{aligned} \tag{5.18}$$

where $\Pi_{\nu\rho} \equiv \nabla_\nu \nabla_\rho \phi = \nabla_\nu \partial_\rho \phi$. However, the symmetry imposed on the conformal factor (i.e. A only being a function of coordinate invariants) means that contributions proportional to $\mathcal{J}_5$ and $\mathcal{J}_6$ cancel and hence do not appear in the equation of motion. As an immediate consequence there is no direct, Vainshtein-like coupling between the stress-energy tensor and derivatives of the field, but matter only enters via $\tilde{T}$.[3] Taking $\tilde{g}_{\mu\nu}$ inside its covariant derivative and using that $\tilde{\nabla} s = \nabla s = \partial s$ for any scalar s (i.e. covariant derivatives related to both metrics act on scalars in the same way), the overall equation of motion can therefore be written

$$\begin{aligned} -V_{,\phi} + A_{,\phi} A^3 \tilde{T} = &- \left(1 + A^3 \tilde{T} A_{,X}\right) \Box\phi \\ &+ 2A^2 \tilde{T} \left(A A_{,X\phi} + 3 A_{,X} A_{,\phi}\right) X \\ &+ A^2 \tilde{T} \left(A A_{,XX} + 3 A_{,X}^2\right) \Pi \\ &- A^3 A_{,X} \partial^\alpha \phi \partial_\alpha \tilde{T}. \end{aligned} \tag{5.19}$$

We now proceed to simplify these expressions by considering a uniform matter source, making the following key assumption about the system under consideration

- The stress-energy tensor describes a pressureless, non-relativistic fluid ($\tilde{T}^0_0 = -\tilde{\rho}$) with all other stress-energy tensor components vanishing.

Later on (in section 5.5) we will also assume a static, uniform source in the matter frame ($\tilde{\nabla}_\alpha \tilde{T}^{\mu\nu} = 0$ inside the source). Combining this with the first assumption, this is equivalent to assuming $\tilde{\nabla}_\alpha \tilde{\rho} = \partial_\alpha \tilde{\rho} = 0$ (again, inside the source). This means $\mathcal{J}_4 = 0$ everywhere except in the transition between source and surroundings. For this reason we will keep the full $\mathcal{J}_4$ term when evaluating radial profiles across this boundary and not make any further assumptions about $\mathcal{J}_4$ for the time being. Note that we also do *not* assume a static profile for ϕ ($\partial_0 \phi = 0$), which one may want to impose for further simplification in a late-universe system like the solar

[3] Note that, at any rate, the Vainshtein-like coupling referred to here would be a derivative coupling which leads to a screening, i.e. a classical renormalization of the kinetic energy, that depends locally on $T^{\mu\nu}$. This means that there will be no screening effect even a small distance away from the source. This is in contrast to Vainshtein screening sourced by derivative self-interactions of the scalar ϕ, e.g. a $X\Box\phi$ term in the action, which will lead to screening inside a Vainshtein radius r_V that can extend beyond the source itself, subject to introducing an appropriate coupling between matter and ϕ such as the linear $\phi\tilde{T}$.

system, which has had time to settle. Modeling matter as a pressureless, non-relativistic fluid results in

$$\begin{aligned}
\mathcal{J}_1 &= A_{,X}\hat{\rho}\Box\phi \\
\mathcal{J}_2 &= -2\hat{\rho}\left(A_{,X\phi} + 3A^{-1}A_{,X}A_{,\phi}\right)X \\
\mathcal{J}_3 &= -\hat{\rho}\left(A_{,XX} + 3A^{-1}A_{,X}^2\right)\Pi \\
\mathcal{J}_4 &= A^3 A_{,X}\partial^\alpha\phi\partial_\alpha\left(A^{-3}\hat{\rho}\right),
\end{aligned} \tag{5.20}$$

where we have substituted $A^3\tilde{T} = -\hat{\rho}$ (or equivalently $\hat{\rho} = A^3\tilde{\rho}$), since $\hat{\rho}$ is a conserved quantity in the Einstein frame. Explicitly working out $\mathcal{J}_4$, which will be relevant when modeling the transition across matter boundaries, one finds

$$\mathcal{J}_4 = 6A^{-1}A_{,X}A_{,\phi}\hat{\rho}X + 3A^{-1}A_{,X}^2\hat{\rho}\Pi + A_{,X}\partial^\alpha\phi\partial_\alpha\hat{\rho}. \tag{5.21}$$

This considerably simplifies the equations of motion, revealing additional symmetries that arise due to the functional form of A in combination with requiring a non-relativistic, pressureless fluid as the matter source. Consequently the equation of motion for ϕ may be written as

$$V_{,\phi} + A_{,\phi}\hat{\rho} = \left(1 - \hat{\rho}A_{,X}\right)\Box\phi + 2\hat{\rho}A_{,X\phi}X + \hat{\rho}A_{,XX}\Pi - A_{,X}\partial^\alpha\phi\partial_\alpha\hat{\rho}. \tag{5.22}$$

An instructive way to think of the physical properties of this system is to explicitly write it as a Klein-Gordon equation with an effective potential that only depends on ϕ and "friction terms" that encode the dependence on higher derivatives of ϕ. If we are purely interested in the profile inside the source (and hence ignore $\partial_\alpha\hat{\rho}$), this means we can write

$$\begin{aligned}
V_{eff,\phi}(\phi) &= \Box\phi + \mathcal{F}_1\left(\phi, X, \hat{\rho}\right)X + \mathcal{F}_2\left(\phi, X, \hat{\rho}\right)\Pi \\
&= \Box\phi + \text{“friction terms”}.
\end{aligned} \tag{5.23}$$

To make the form of the effective potential explicit, one can Taylor-expand the conformal factor A in powers of X, writing

$$A(\phi, X) = A^{(0)}(\phi) + A^{(1)}(\phi)X + \mathcal{O}(X^2), \tag{5.24}$$

which allows us to write down the effective potential for ϕ as

$$V_{eff,\phi}(\phi) = \left(V_{,\phi} + \hat{\rho}A^{(0)}_{,\phi}\right)\left(1 - A^{(1)}\hat{\rho}\right)^{-1}. \tag{5.25}$$

5.3.3 Phenomenology and comments

Given some original potential of the runaway form $V(\phi)$, the chameleon mechanism generates an environmentally dependent effective potential that gives ϕ a large mass in high density regions. Here "high density" and "large mass" are essentially references to solar system constraints [26] on the presence of a fifth force mediated by a scalar degree of freedom. As such, any theory with a runaway $V(\phi)$ that successfully implements the chameleon mechanism at the very least has to tick two boxes. Firstly it needs to give rise to an environmentally-dependent effective potential which has a minimum. And secondly the mass of small oscillations around that minimum has to be large enough to satisfy fifth force constraints. With this in mind let us investigate the effective potential described by (5.22) and (5.25). Note that in section 5.4 we will give explicit examples illustrating each of the effects outlined here in detail.

Position of minimum: For a finite density $\hat{\rho}$ and conformal factor A (more precisely, $A^{(1)}$), (5.25) shows that a potential minimum requires

$$V_{,\phi}\left(\phi_{\min}\right) + \hat{\rho} A^{(0)}_{,\phi}\left(\phi_{min}\right) = 0, \tag{5.26}$$

which is identical to the condition for the minimal conformal chameleon discussed in section 5.2. In other words, the position of the minimum, ϕ_{min}, is not altered by the introduction of an X-dependent conformal factor. This also shows that a conformal factor which does not depend on ϕ itself, but only on derivatives of ϕ, (i.e. $A^{(0)}_{,\phi} = 0$, as is the case for $A = A(X)$) cannot generate an effective potential with a minimum, if the original $V(\phi)$ does not already possess a minimum itself. The ϕ-dependence of A is consequently essential to obtaining a successful implementation of the chameleon mechanism, if starting with a runaway potential (e.g. $V(\phi) \sim \phi^{-n}$). A potential $V(\phi)$ with very small mass m_V^2 can however be uplifted by a pure derivative conformal factor (e.g. $A(X)$), leading to an effective potential with mass $m^2 \gg m_V^2$. Below and in the next section we will give explicit examples illustrating this behavior.

Effective mass: A derivative-dependent conformal factor affects the curvature of the effective potential. This means derivative chameleons generically come equipped with a new mass-altering mechanism. The effective potential, and hence the mass of the field, is classically renormalized by introducing higher order invariants into the conformal relation as follows

$$\begin{aligned} m^2 = V_{eff,\phi\phi}(\phi_{min}) &= \left(V_{,\phi\phi}(\phi_{min}) + \hat{\rho} A^{(0)}_{,\phi\phi}(\phi_{min})\right)\left(1 - A^{(1)}(\phi_{min})\hat{\rho}\right)^{-1} \\ &= m^2_{standard}\left(1 - A^{(1)}(\phi_{min})\hat{\rho}\right)^{-1}, \end{aligned} \tag{5.27}$$

where $m^2_{standard}$ denotes the effective mass for small oscillations around the minimum for a theory with identical conformal factor in the limit $A(\phi, X \to 0)$. In other words, $m^2_{standard}$ is the effective

mass for the theory in the limit where higher-derivative contributions can be neglected. This mass-altering mechanism can be separated into three branches, which we will discuss now.

Mass-lifting, ghost-like instabilities and anti-chameleons: Equation (5.27) shows that the effective mass m^2 is enlarged when $0 < A^{(1)}(\phi_{min})\hat{\rho} < 1$. This is interesting since it suggests that derivative chameleon models provide an additional mass-lifting mechanism, potentially alleviating the fine-tuning involved in obtaining a sufficiently large mass in dense environments for standard chameleon models. However, care must be taken when considering mass-lifting solutions for the following reason. Suppose we consider a conformal factor A such that a mass-lifting mechanism is in place, i.e. $0 < A^{(1)}(\phi_{min})\hat{\rho}_1 < 1$ for some given energy density $\hat{\rho}_1$. Now, assuming A has no density-dependence itself, one can solve for a (larger) critical density $\hat{\rho}_{crit}$ above which the solution becomes unstable. The effective potential switches sign since $1 - A^{(1)}(\phi_{min})$ becomes negative, so that ϕ_{min} becomes ϕ_{max} and $V_{eff,\phi\phi}(\phi_{min})$ turns negative. Thus we are left with a negative "mass term", signaling instabilities, and the solution becomes ghost-like. Figure 5.1 illustrates these different regimes. That different energy densities $\hat{\rho}$ will interpolate between stable mass-lifting and ghost-like solutions can also be seen from the relevant part of the equation of motion

$$V_{,\phi} + A_{,\phi}\hat{\rho} \quad = \quad (1 - \hat{\rho}A_{,X})\,\Box\phi + ..., \tag{5.28}$$

where, if $\hat{\rho}A_{,X} > 1$, this can be traced back to an action with the "wrong" sign for ϕ's kinetic term.[4]

There appear to be two obvious solutions to this instability problem.[5] Firstly one could consider making $A(\phi, X)$ a function of the energy density $\hat{\rho}$ as well. However, especially given that the matter stress-energy tensor $\tilde{T}^{\mu\nu}$ *is* a variation of the matter Lagrangian with respect to the metric $\tilde{g}_{\mu\nu} = A^2(\phi, X)g_{\mu\nu}$, it is not clear what such an iterative dependence on $\tilde{T}^{\mu\nu}$ would mean. Nevertheless it will be an interesting task for the future to think about whether there is some convincing way of implementing such a dependence. In any case, from a purely phenomenological point of view, a density dependent A allows us to have a stable, derivative dependent mass-lifting mechanism for all $\hat{\rho}$. Secondly one may choose a conformal factor such that $\hat{\rho}A_{,X} \lesssim 1$ up to the density cutoff of the theory. Note, however, that the effective chameleon mass m^2 is proportional to $(\hat{\rho}_{crit} - \hat{\rho})^{-1}$, where we have defined a critical density $\hat{\rho}_{crit}$ such that $A^{(1)}(\phi_{min})\hat{\rho}_{crit} = 1$. Introducing a cutoff at $\hat{\rho}_{crit}$ will then render derivative-dependent effects

[4] Jumping ahead slightly, also note that, if $A_{,X}$ is a function of ϕ and the field is not approximately constant $\phi \sim \phi_{min}$ inside the source (the so-called "thick-shell regime - see section 5.5), one needs to be aware that $\hat{\rho}A_{,X}(\phi)$ being smaller than unity for some initial $\phi_i(r = 0)$ no longer guarantees that this remains true for all values of ϕ taken inside the source.

[5] Should we decide to bite the bullet and accept the existence of ghost-like solutions above some density-scale $\hat{\rho}$, such an approach will also face major challenges when confronted with high energy/density early universe physics.

on m^2 suppressed by that same cutoff scale

$$m^2 = \frac{m^2_{standard}}{(\hat{\rho}_{cutoff} - \hat{\rho})\, A^{(1)}(\phi_{min})} = m^2_{standard}\left(1 - \frac{\hat{\rho}}{\hat{\rho}_{cutoff}}\right)^{-1}. \tag{5.29}$$

If the cutoff is low enough, significant derivative-dependent effects on m^2 may still be obtained, but for a high cutoff density such as the Planck density ρ_P they will be strongly suppressed.

The third branch of mass-altering solutions corresponds to the case when $A^{(1)}(\phi_{min}) < 0$? Then the effective mass of ϕ is reduced, counter-acting chameleon screening effects

$$m^2 = \frac{m^2_{standard}}{1 + \left|A^{(1)}(\phi_{min})\hat{\rho}\right|}. \tag{5.30}$$

This branch is free of instabilities and provides a robust mechanism to suppress mass terms in models with derivative conformal factors A, since the derivative dependence reduces the curvature of V_{eff}. This suggests that the simplest $A^2(\phi, X)$ models, where A is independent of $\hat{\rho}$ and no ghost-like instabilities arise for any $\hat{\rho}$, are anti-chameleon models in the sense that the effective mass m^2 is reduced compared with the non-derivative chameleon limit $A(\phi, X \to 0)$.

Thin shell regime: The computation of V_{eff} and its minima/maxima and masses above was oblivious to "friction" terms in (5.23) (by definition, since the effective potential is a function of ϕ and not its derivatives). But such terms automatically arise as a consequence of a higher order conformal coupling. Here "friction terms" is a reference to terms other than $\Box\phi$ that have a derivative dependence on ϕ, e.g. $\mathcal{J}_2$ and $\mathcal{J}_3$ which encode the dependence on X and Π. While not influencing V_{eff}, these terms do impact the dynamics of ϕ, warranting further investigation.

One particularly interesting consequence is a modification of the so-called thin-shell regime. In typical chameleon theories the chameleon-charge of large, massive objects can be Yukawa-screened with only a thin-shell on the outside of the object contributing to the exterior ϕ-profile [15, 16]. This is essential for e.g. avoiding unacceptably large effects on planetary orbits due to a chameleonic fifth force. Now the presence of additional friction terms modifies the gradient of ϕ inside the source and one therefore expects a modification of the thin-shell effect as well. We will discuss this in detail in section 5.5, where we investigate radial solutions around massive sources, focusing on new-found phenomenology due to derivative conformal factors.

Equivalence principle (violations): A final comment on possible equivalence principle violations. By construction, chameleon models (both standard as well as the derivative generalization considered here) respect the weak equivalence principle. In the non-derivative case, the extra degree of freedom ϕ locally influences the dynamics, but it does not discriminate between different test masses/types of matter due to the universal coupling of $A^2(\phi)$ to all matter fields Ψ_i (field-dependent couplings are discussed in [146]). In the derivative case, an additional degree

of freedom X enters. But since all matter still universally couples to $A^2(\phi, X)g_{\mu\nu}$, all test masses locally experience the same gravitational force (note that ϕ is viewed as a gravitational scalar here). In both cases ϕ's profile of course does not remain constant across space-time, as it depends on the ambient density. This means the strong equivalence principle is trivially violated (as is, in fact, the Einstein equivalence principle - cf. with e.g. the Horndeski constructions in [147] where the strong, but not the Einstein equivalence principle is broken).

The main point we wish to make here is that a prima facie worry one might have when considering $A^2(\phi, X)$, namely that a derivative-dependent coupling to matter violates even the weak equivalence principle, is not justified. This is simply due to the fact that there is no dependence of the gravitational coupling on the momentum of matter (Ψ_i) test masses, but only a derivative coupling to the gravitational scalar ϕ. Nevertheless this is an area that warrants further investigation, since known violations of the strong equivalence principle in chameleon models (cf. [148, 63, 149, 150, 151]) will be modified by introducing a derivative dependence. Computing this in detail should enable further disentangling of derivative vs. non-derivative chameleons.

5.4 Effective potentials and the chameleon mass

In this section we illustrate how chameleon mechanisms arise in derivative theories with a number of explicit examples. We focus on the mass-altering mechanism for different conformal factors.

5.4.1 Example I: Taylor-expanding $A(\phi, X)$

We begin by reminding ourselves of (5.24) and Taylor-expanding the conformal factor

$$A(\phi, X) = A^{(0)}(\phi) + A^{(1)}(\phi)X + \mathcal{O}(X^2), \tag{5.31}$$

where the associated mass of oscillations around an effective minimum is given by (5.25). We will ignore higher orders in X and, as a first simple example, focus on the zeroth order contribution in ϕ to $A^{(1)}$, treating it as a constant. As such the conformal factor takes on the form

$$A(\phi, X) = A^{(0)}(\phi) + k_X X. \tag{5.32}$$

The resulting potential and effective mass can be computed to give

$$V_{eff}(\phi) = \frac{V(\phi) + \hat{\rho}A^{(0)}(\phi)}{1 - k_x\hat{\rho}}, \tag{5.33}$$

$$m^2 = \frac{V_{,\phi\phi}(\phi_{min}) + \hat{\rho}A^{(0)}_{,\phi\phi}(\phi_{min})}{1 - k_X\hat{\rho}} = \frac{m^2_{standard}}{1 - k_X\hat{\rho}}. \tag{5.34}$$

This illustrates the point made in the previous section about the existence of a mass-raising and a mass-lowering branch for derivative chameleon models. In the simple setup considered here, if k_X is negative, the chameleon mass m^2 is reduced compared with the standard $A = A(\phi)$ theory. The fine-tuning necessary in order to get a sufficiently large chameleon mass is therefore made more severe in this case, since an additional mass-lowering mechanism is at work. For positive k_X a mass-lifting mechanism operates, increasing m^2. However, for $k_X\hat{\rho} = 1$ the mass diverges and, once $k_X\hat{\rho}$ exceeds unity, the "mass" turns negative and the solution ghost-like. As discussed in the previous section, this is cause for concern, since given some gravitational theory the conformal factor is fully specified (and hence k_X is fixed in this case). As such, sources with densities above $\hat{\rho}_{crit} = 1/k_X$ essentially see an unstable inverted potential $-V_{eff}$ inside the source. The bottom left graph of figure 5.1 illustrates this point by plotting the dependence of the effective potential on k_X for some given source density $\hat{\rho}_1$. This amounts to considering different normalizations of the effective potential. On the other hand, the bottom right plot shows how varying $\hat{\rho}$ for some given, positive and fixed k_X affects the solution. Note that having restricted to positive k_X means there are no mass-lowering solutions present in this plot. As expected from (5.33), changing $\hat{\rho}$ modifies both the position of the minimum as well as as changing the potential's normalization and hence curvature/mass via the $1 - k_x\hat{\rho}$ term.

As an aside, notice that, if one does allow the conformal factor to depend on energy density $\hat{\rho}$ (for a discussion of this approach see the previous section), then one can construct a solution which remains ghost-free inside sources with arbitrarily large energy densities $\hat{\rho}$. E.g. we may impose

$$A(\phi, X) = e^{k_1\phi} + \hat{\rho}^{-1}\left(1 - e^{k_2\phi}\right)X + \mathcal{O}(X^2), \tag{5.35}$$

which results in an effective mass given by

$$m^2 = e^{k_2\phi_{min}}\left(V_{,\phi\phi}(\phi_{min}) + k_1^2\hat{\rho}e^{k_1\phi_{min}}\right). \tag{5.36}$$

In other words, in this particular example k_2 allows tuning the mass m^2 arbitrarily. The potential is stable for all ϕ and the mass can be altered by modifying the X-dependence of the conformal factor A^2.[6]

5.4.2 Example II: A separable $A(\phi, X)$

In the Taylor-expanded picture laid out in the previous section, considering a separable conformal factor $A(\phi, X)$ amounts to setting $A^{(0)}(\phi) = A^{(1)}(\phi)$. An interesting feature compared with the k_X case discussed above is that the normalization of the potential now also becomes a function

[6] If one is concerned that additional singularities are introduced by the $\hat{\rho}^{-1}$ dependence of A here, notice that a condition such as $A^{(1)}\hat{\rho} \sim 1 - e^{-\hat{\rho}}$ will also ensure a ghost-free mass-lifting regime.

of ϕ. To see this explicitly we write the conformal factor as

$$A(\phi, X) = B(\phi)C(X), \tag{5.37}$$

which means the effective potential satisfies

$$V_{eff,\phi}(\phi) = \left(V_{,\phi} + \hat{\rho}B_{,\phi}C^{(0)}\right)\left(1 - BC^{(1)}\hat{\rho}\right)^{-1}, \tag{5.38}$$

where C^i refers to the i-th component in a Taylor-series expansion of $C(X)$ in powers of X^i. Since at the minimum of the potential $V_{eff,\phi}(\phi_{min}) = 0$, we obtain an effective chameleon mass for oscillations around ϕ_{min} of

$$m^2 = \frac{V_{,\phi\phi}(\phi_{min}) - \frac{1}{2}\hat{\rho}C^{(0)}B_{,\phi\phi}(\phi_{min})}{1 - \frac{1}{2}k_2B(\phi_{min})C^{(1)}\hat{\rho}}. \tag{5.39}$$

Viewing the effective potential as a renormalized version of the corresponding standard (i.e. non-derivative) chameleon setup with $C = 1$, the effect of the $B(\phi_{min})$ term in the denominator can be described as making this renormalization ϕ-dependent. The top right plot in figure 5.1 shows the effective potential for a separable conformal factor of the form $A(\phi, X) = \mathrm{Exp}\left[k_1\phi + k_2X\right]$. Note that we require a positive k_2, because otherwise the conformal factor diverges as $\partial_\mu\phi \to 0$, i.e. no stable, static solution is possible. Now the effective mass can be written as

$$m^2 = \frac{V_{,\phi\phi}(\phi_{min}) + \hat{\rho}k_1^2e^{k_1\phi_{min}}}{1 - k_2e^{k_1\phi_{min}}\hat{\rho}} = m^2_{standard}\left(1 - k_2e^{k_1\phi_{min}}\hat{\rho}\right)^{-1}. \tag{5.40}$$

The solution again separates into the three branches discussed. For positive $B(\phi)$ and when $0 < k_2B\hat{\rho} < 1$ the effective potential has a well behaved minimum protected by an infinite potential barrier at ϕ_{crit}, where ϕ_{crit} satisfies $k_2B(\phi_{crit})\hat{\rho} = 1$ for a given $\hat{\rho}$. Equation (5.40) then shows that m^2 is enhanced by the derivative-dependence of $A(\phi, X)$. For values of $\phi > \phi_{crit}$ there do exist unstable regions of parameter space. However, a particle, which starts at an initial field value ϕ_i for which $k_2B(\phi_i)\hat{\rho} < 1$, is protected from entering the unstable region. A more physically insightful understanding of the inherent instabilities in these models can be gained by phrasing the same argument in terms of a critical density. In particular we note that, given a conformal factor that fixes k_2, solutions with positive $k_2B(\phi)$ inside a source with energy density such that $\hat{\rho} > (k_2B)^{-1}$ are unstable. Finally, for negative $k_2B(\phi)C^{(1)}$ a stable, mass-lowering solution is obtained similar to the one discussed in the previous section. The top right plot of figure 5.1 illustrates these three branches and one can see the generation of an effective minimum and how the mass m^2 can be tuned by varying k_2 (for a fixed $\hat{\rho}$) at the expense of lowering the critical field value ϕ_{crit}.

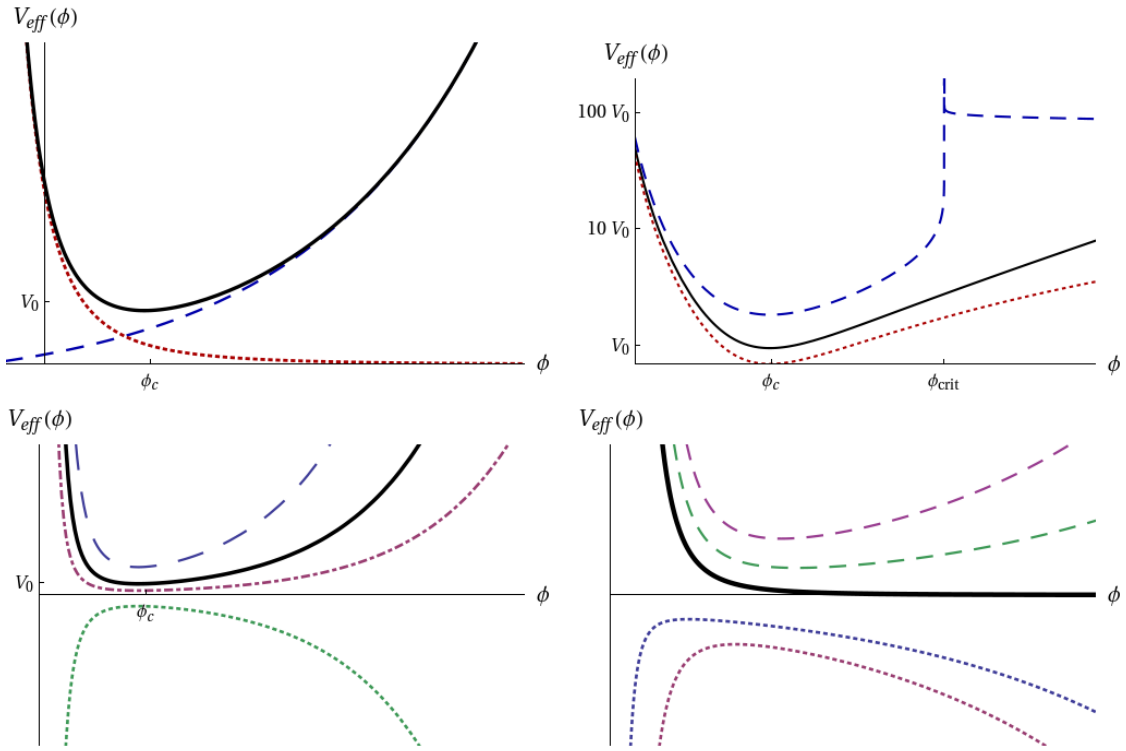


Figure 5.1: **Top Left**: Plot showing the effective potential $V_{eff}(\phi)$ (solid line) for a separable conformal factor $A(\phi, X) = B(\phi)C(X) = e^{k_1\phi + k_2 X}$ (dashed line) and a runaway potential $V(\phi) \sim \phi^{-3}$ (dotted line) in arbitrary units. **Top Right**: Logarithmic plot showing $V_{eff}(\phi)$ for different choices of k_2 (negative, zero and positive from bottom to top with $\hat{\rho} = 1$ in arbitrary units). This shows how the amplitude, and hence mass, of the chameleon field grows as k_2 is enlarged, while the stable region of parameter space is reduced. The dashed line shows how the pole in the denominator of V_{eff} creates a discontinuity at ϕ_{crit} and the effective potential assumes a runaway form for values of $\phi \gtrsim \phi_{crit}$. In other words, the solution transitions between the mass-lifting and ghost-like branches at ϕ_{crit}. Note that the value taken by ϕ_{crit} is a function of $\hat{\rho}$ and the effective potential. **Bottom Left**: Plot showing the effective potential for a conformal factor as in equation (5.32) and for different values of the parameter k_X and fixed $\hat{\rho}$: The solid line shows the non-derivative chameleon with $k_X = 0$, dashed lines show the mass-lifting branch with $0 < k_X\hat{\rho} < 1$, dot-dashed lines show the mass lowering branch with $k_X < 0$ and dotted lines lie in the unstable region with $k_X\hat{\rho} > 1$. **Bottom Right**: Analogous plot varying $\hat{\rho}$ for some fixed positive k_X (hence no mass-lowering branch is present here). Varying $\hat{\rho}$ changes the position of the minimum as well as the normalization of V_{eff} and hence its curvature and mass m^2. $\hat{\rho}$ is zero for the solid line and below/above $\hat{\rho}_{crit}$ for dashed/dotted lines respectively.

5.4.3 Example III: A purely derivative conformal factor $A(X)$

Suppose we have a purely derivative conformal factor that does not depend on the field value of ϕ itself. This is interesting, since it means the theory has a shift symmetry $\phi \to \phi + c$, potentially protecting it from a number of quantum corrections[7] (assuming this is at most softly broken by the potential $V(\phi)$ - cf. [152], where this point is discussed in an inflationary setting). As such, let us consider a conformal factor of the following form

$$\tilde{g}_{\mu\nu} = A^2(X) g_{\mu\nu} = \left(1 + A^{(1)} X + \mathcal{O}(X^2)\right) g_{\mu\nu}, \tag{5.41}$$

where $A^{(0)}$ has been appropriately normalized (i.e. there is no fundamental reason why it should be unity) and where all A^i are now constants and therefore *not* functions of ϕ. The reason to require $A^{(0)} \neq 0$ is that otherwise any coupling between gravity and matter vanishes once $\partial_\mu \phi = 0$, e.g. once ϕ has settled into its minimum. The full action can then be written as

$$\mathcal{S} = \int d^4x \sqrt{g} \left(\frac{M^2}{2} R + X - V(\phi) \right) + \mathcal{S}_m \left(A^2(X) g_{\mu\nu}, \Psi_i \right). \tag{5.42}$$

Now we have already seen that a conformal factor, which does not depend on the field value ϕ itself, cannot create an effective potential with a minimum, if $V(\phi)$ is of the runaway form (so it does not have a mass term). But what if dark energy is in fact sourced by a very light, but not massless, cosmological scalar? Then $V(\phi)$ does already have a (very small) mass term and hence a minimum. Let us focus on a particularly simple toy model and consider a power-law potential of the form

$$V(\phi) = \frac{1}{2} \left(\frac{\phi}{\phi_0} \right)^2, \tag{5.43}$$

i.e. a simple mass term for $m_\phi = \phi_0^{-1}$, which allows us to tune the mass of the field by choosing ϕ_0. In order for this field to act as dark energy on large scales, it has to be extremely light $m_\phi = \phi_0^{-1} \lesssim H_0 \sim 10^{-33} eV$. Now the effective equation of motion for ϕ is

$$V_{,\phi} = (1 - \hat{\rho} A_{,X}) \Box\phi + \hat{\rho} A_{,XX} \Pi - A_{,X} \partial^\alpha \phi \nabla_\alpha \hat{\rho}. \tag{5.44}$$

and we can write down an effective potential $V_{eff}(\phi) = V(\phi) \left(1 - A^{(1)} \hat{\rho}\right)^{-1}$. In dense environments the effective mass of the chameleon field consequently goes as

$$m^2 = \frac{V_{,\phi\phi}}{1 - A^{(1)} \hat{\rho}} = \left(\phi_0^2 - \phi_0^2 A^{(1)} \hat{\rho} \right)^{-1}, \tag{5.45}$$

[7] This statement is of course not as strong as for a theory with an additional Galilean shift symmetry (cf. [29]) where an effective non-renormalization theorem can be derived [35, 36]. Note, however, that even there analogous concerns enter due to the non shift-symmetric matter coupling $\phi \tilde{T}$, cf. discussions in [63, 64].

which can be very large subject to $0 < 1 - A^{(1)}\hat{\rho} \ll 1$. This outlines how a mass-lifting mechanism caused by a purely X-dependent conformal factor can give rise to a viable chameleon-type solution. An otherwise extremely light cosmological scalar field then acquires a large mass in dense environments.

We emphasize that the point made here is independent of the exact form of the potential. Given any potential for a non-massless field ϕ, i.e. a potential with a minimum at ϕ_c, the mechanism outlined here will raise the field's mass in a density-dependent manner. We can therefore apply this purely derivative mass-lifting mechanism to any potential $V(\phi)$ for a field with a small mass $m_\phi \neq 0$, which reproduces the desired dark energy behavior on cosmological scales.

However, this approach faces a major obstacle. The field's mass has to be enlarged by several orders of magnitude if a light cosmological scalar ϕ is to escape detection on solar system scales. This means $A^{(1)}\hat{\rho} \sim 1$ for a large range of densities $\hat{\rho}$, so that $A^{(1)}$ has to depend on $\hat{\rho}$ in order to suppress the otherwise strong dependence of m^2 on $\hat{\rho}$ and in order to prevent ghost-like instabilities from developing. We refer to section 5.3 and the previous examples in this section for a discussion of the possibility of such a density-dependent conformal factor. Also note that the mass-lowering branch with $A^{(1)} < 0$, albeit perhaps less interesting phenomenologically, does not face this problem just as discussed in the previous examples.

5.5 Radial solutions, the thin-shell effect and "friction terms"

In this section we investigate the static, radial ϕ-profile in and around a spherically symmetric body with uniform density $\hat{\rho}$ - a good approximation for the profile around the sun or earth, for example.[8] Expressing the general equation of motion (5.22) for such a profile one finds

$$\begin{aligned} V_{,\phi} + A_{,\phi}\hat{\rho} &= (1 - \hat{\rho}A_{,X})\left(\frac{d^2\phi}{dr^2} + \frac{2}{r}\frac{d\phi}{dr}\right) - \hat{\rho}A_{,X\phi}\left(\frac{d\phi}{dr}\right)^2 \\ &+ \hat{\rho}A_{,XX}\frac{d^2\phi}{dr^2}\left(\frac{d\phi}{dr}\right)^2 - A_{,X}\partial^\alpha\phi\partial_\alpha\hat{\rho}, \end{aligned} \tag{5.46}$$

where $\hat{\rho}$ and ϕ are functions of r. We now modify the approach taken in [16] in that we still assume

$$\hat{\rho}(r) = \begin{cases} \hat{\rho}_c & \text{for } r \ll R \\ \hat{\rho}_\infty & \text{for } r \gg R. \end{cases} \tag{5.47}$$

where R is the radius of the spherical object, $\hat{\rho}_c$ is its density and $\hat{\rho}_\infty$ is the density of the surroundings. With an eye on computing the gradient term $\partial_\alpha\hat{\rho}(r)$ in (5.46), we model the

[8] For a time-dependent chameleon field in the case of a radially pulsating mass, see [153].

boundary between source and surroundings by a very sharp, but smooth, transition between $\hat{\rho}_c$ and $\hat{\rho}_\infty$. The particular template we shall adopt is

$$\hat{\rho}(r) = \frac{1}{2}(\hat{\rho}_c - \hat{\rho}_\infty)(1 - \text{Tanh}\,[s(r-R)]) + \hat{\rho}_\infty, \tag{5.48}$$

taking $s \gg 1$, so that $\hat{\rho}$ effectively remains at its asymptotic values except for a sharp transition around $r = R$. A unique solution for (5.46) requires specifying two boundary conditions. Again following [16], we take these to be $d\phi(r=0)/dr = 0$, so that the solution is non-singular at the origin, and $\phi \to \phi_\infty$ as $r \to \infty$, which ensures that the ϕ-mediated force between two test bodies vanishes as $r \to \infty$.

In what follows we will first recap how this setup gives rise to a thin-shell effect for the minimal conformal chameleon (i.e. the standard case). Then we discuss how radial solutions and thin-shell behavior are modified by introducing derivative conformal factors, paying special attention to the effect of "friction terms" in (5.46) (i.e. X and Π dependent terms as in $\mathcal{J}_2$ and $\mathcal{J}_3$ (5.20)).

5.5.1 The thin-shell effect for standard chameleons

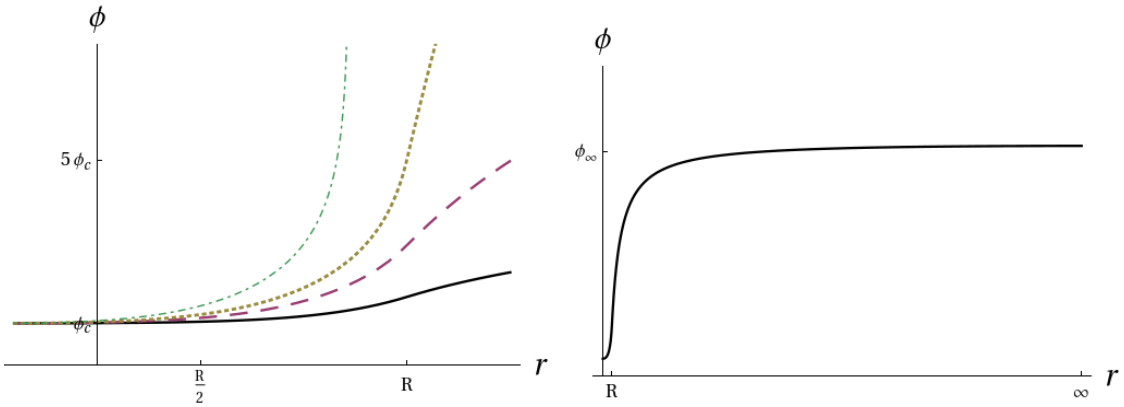


Figure 5.2: **Left**: The radial solution for a standard chameleon with conformal factor $A(\phi) = e^{k_1\phi}$ and different values for the parameter $k_1 = 1, 1.05, 1.1, 1.2$ from bottom to top. The initial $\phi_i(r=0)$ remains fixed for all k_1 here (note that the position of the extremum ϕ_c also depends on A and hence k_1). **Right**: Plot showing how a suitably chosen ϕ_i leads to a radial solution that asymptotes to ϕ_∞ as $r \to \infty$. $k_1 = 1$ here.

Any given physical system will come equipped with a specific $V(\phi)$, $\hat{\rho}$, R etc. This will then allow us to compute the corresponding solution to (5.46), in particular fixing the initial field-value ϕ_i at $r = 0$. Alternatively one can explore the system's solutions by choosing some ϕ_i and finding the corresponding region in $(\hat{\rho}, R, ...)$ parameter space [16].[9]

[9]This amounts to solving (5.46) as a classical mechanics problem with ϕ being a position- and r being a time-coordinate. Note that the opposite signs for temporal and spatial dimensions in $g_{\mu\nu}$'s signature mean that, for the radial solution we are considering here, the particle whose "position" is described by ϕ "moves" on the inverted, and hence unstable, potential $-V_{eff}$.

Now for the minimal conformal chameleon, as described in section 5.2, the radial equation of motion is simply

$$V_{,\phi} + A_{,\phi}\hat{\rho} \;=\; \left(\frac{d^2\phi}{dr^2} + \frac{2}{r}\frac{d\phi}{dr}\right). \tag{5.49}$$

Outside the source, i.e. at a radius $r > R$, the solution for ϕ then assumes the form [16]

$$\begin{aligned} \phi(r) &\sim \left(\frac{\Delta R}{R}\right)\frac{M_c e^{-m_\infty(r-R)}}{r}, \\ \frac{\Delta R}{R} &\sim \frac{\phi_\infty - \phi_c}{\Phi_c} \sim \frac{R - R_{roll}}{R} \ll 1, \end{aligned} \tag{5.50}$$

where R is the radius of the source, M_c is its mass, ϕ_∞ and m_∞ are the field value and mass of the field as $r \to \infty$, ϕ_c is the local extremum of the ϕ-potential inside the source[10] and Φ_c is the Newtonian potential at the surface of the source. We have also assumed that $\Delta R/R$, the so-called thin-shell suppression factor, is small. An intuitive picture for this setup is the following: The field is released at "time" $r = 0$ from ϕ_i. Inside the source the field remains stuck near ϕ_i with its dynamics dominated by the $2/r \cdot d\phi/dr$ friction term. Eventually, at $r = R_{roll}$, the field ϕ starts rolling and hence developing gradient terms. Only contributions from the region between R_{roll} and R, i.e. the region where gradient terms develop, are felt by the exterior profile. Since $\Delta R/R \ll 1$, the chameleon force from a large massive body on a test mass is therefore thin-shell suppressed. This is essential in explaining why e.g. planetary orbits are not affected by a chameleonic force. In the language set out above the thin-shell regime corresponds to $\phi_i - \phi_c \ll \phi_c$.

In contrast, the radial solution for the unsuppressed "thick-shell" regime, where $\phi_i \gtrsim \phi_c$, is given by

$$\phi(r) \;\sim\; \frac{M_c e^{-m_\infty(r-R)}}{r}. \tag{5.51}$$

Here the field starts significantly displaced from ϕ_c and consequently starts rolling almost immediately. Gradient terms develop inside the source and there is no thin-shell suppression (hence the absence of the $\Delta R/R$ factor). It is intuitively clear that the presence of further derivative terms in the general solution (5.22) will modify this behavior, since it will affect gradient terms in ϕ. We now move on to explore such "friction" terms and their effect on the thin-shell mechanism.[11]

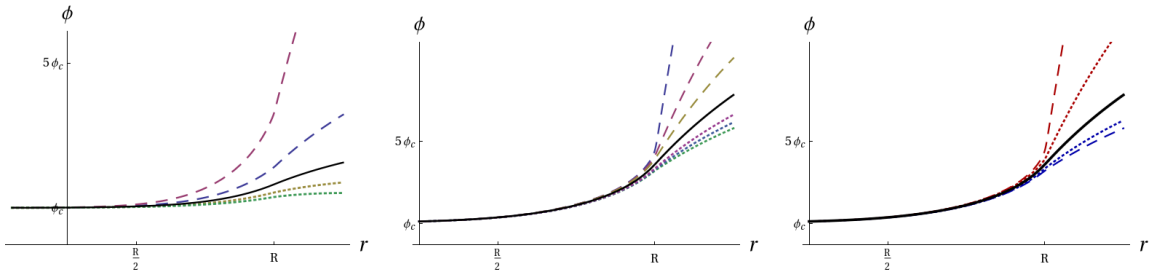


Figure 5.3: **Left**: Plot of the radial solution for a conformal factor as in (5.32) with $k_X = 0$ (solid line), k_X increasingly positive (dashed lines - k_X increases as one moves out from the solid line) and k_X increasingly negative (dotted lines - k_X decreases as one moves out from the solid line). **Middle**: Plot of the radial solution for a conformal factor as in (5.54). Again the solid line denotes $k_X = 0$. Dashed lines are for fixed positive k_X and outwardly increasing k_2, while dotted lines are for negative k_X, again with outwardly increasing k_2. In order to illustrate the effect of derivative contributions more clearly the solution is shown for an initial displacement $\phi_i - \phi_c \sim \mathcal{O}(\phi_c/10)$, i.e. an intermediate region between thin- and thick-shell screening. Notice that the most significant effect comes from the $\partial\hat{\rho}$ term at $r = R$, which produces a "kink" in the radial solution (actually smooth due to the density profile (5.48)). **Right**: Analogous plot with the solid line describing $k_X = 0$. k_X is positive above and negative below the solid line. Here dashed lines represent the solution including the friction term that depends on $A_{,X\phi}$ (cf equation (5.55)), while dotted lines ignore this contribution. The additional "friction term" therefore drives the solution further away from the non-derivative limit.

5.5.2 Radial solutions for derivative chameleons

Derivative solutions without "friction terms": How does the radial solution and thin-shell suppression change once we go beyond the simplest chameleon setups? We begin by considering a simple conformal factor of the form

$$A(\phi, X) = e^{k_1\phi} + k_x X, \tag{5.52}$$

i.e. a sum-separable $A(\phi, X) = B(\phi) + C(X)$ with purely linear dependence on X. This reduces (5.46) to

$$V_{,\phi} + A_{,\phi}\hat{\rho} \;=\; (1 - \hat{\rho}A_{,X})\left(\frac{d^2\phi}{dr^2} + \frac{2}{r}\frac{d\phi}{dr}\right) - A_{,X}\partial^\alpha\phi\partial_\alpha\hat{\rho}. \tag{5.53}$$

In contrast with the standard chameleon case there are therefore two new effects. Firstly the $(1 - \hat{\rho}A_{,X})$ term, which effectively renormalizes the potential and is responsible for the mass-changing mechanism discussed in section 5.4. And secondly the new gradient term in $\partial_\alpha\hat{\rho}$, which will only be significant very close to $r = R$. Importantly we have no X- or Π-dependent friction terms for a conformal factor of the form considered here.

We plot radial solutions for (5.53) in the left panel of figure 5.3 for different values of k_x. The

[10] Since we are effectively dealing with motion along $-V_{eff}$, ϕ_c is the same as ϕ_{min} in previous sections.

[11] Further discussions of standard thin-shell screening and spherically symmetric solutions can be found e.g. in [15, 16, 154, 155].

behavior observed can be understood as follows. For positive k_X, the $(1-\hat{\rho}A_{,X})$ term increases the curvature of the effective potential (and hence the mass of oscillations around the minimum). This means that the driving term $V_{eff,\phi}$ is enhanced and will overcome the $2/r \cdot d\phi/dr$ friction "earlier" in the evolution, i.e. it will reduce R_{roll}. In other words, in order to obtain the same thin-shell screened exterior solution one now needs to release the particle ϕ from even closer to the minimum value ϕ_c, so that $\phi_i^{derivative} \ll \phi_i^{standard}$. The $\partial_\alpha\hat{\rho}$ term in fact further increases this tendency, giving an additional positive "kick" to the gradient of ϕ. Note that this second effect is oblivious to the profile inside the source and as such does not modify the thin-shell condition. However, in order to reach the same boundary value ϕ_∞ as $r \to \infty$, both new terms require an initial ϕ_i closer to ϕ_c than in the non-derivative case.

For negative k_X the converse is true. The curvature and associated mass of the effective potential is reduced and thin-shell screening is enhanced, i.e. R_{roll} is pushed closer to R. This broadens the thin-shell screened parameter-space and reaching the boundary value ϕ_∞ as $r \to \infty$ requires an initial ϕ_i further away from ϕ_c than in the non-derivative case for negative k_X. Overall the new derivative dependent effects found here modify the parameter-space for thin-shell screened solutions, with a positive/negative k_X weakening/strengthening thin-shell screening respectively. As intuitively expected from modifying the curvature of the effective potential, the mass-lifting branch is therefore associated with a suppressed thin-shell mechanism, whereas the mass-lowering branch enhances thin-shell screening.

Derivative solutions with "friction terms": Here we finally wish to examine the effect of "friction terms" that depend on derivatives of ϕ. The toy model we adopt has a conformal factor

$$A(\phi, X) = e^{k_1\phi} + k_X e^{k_2\phi} X + \mathcal{O}(X^2), \tag{5.54}$$

where we will ignore the higher order corrections $\mathcal{O}(X^2)$. This means the radial solution (5.46) simplifies as Π-dependent contributions drop out (such contributions only come in for a conformal order with non-linear dependence on X). The equation of motion consequently becomes

$$V_{,\phi} + A_{,\phi}\hat{\rho} = (1-\hat{\rho}A_{,X})\left(\frac{d^2\phi}{dr^2} + \frac{2}{r}\frac{d\phi}{dr}\right) - \hat{\rho}A_{,X\phi}\left(\frac{d\phi}{dr}\right)^2 - A_{,X}\partial^\alpha\phi\partial_\alpha\hat{\rho}. \tag{5.55}$$

In the middle and right panel of figure 5.3 we plot solutions for constant k_1, but varying k_2 and k_X, essentially tuning the contribution of the $2\hat{\rho}A_{,X\phi}X$ term. One can see that a large k_2 in combination with the presence of the $\partial_\alpha\hat{\rho}$ term magnifies the "kick" at $r \sim R$. Perhaps more interestingly, for positive k_X the direct coupling of ϕ to its derivative in the conformal factor again counteracts the thin-shell effect, by driving the field away from near its minimum ϕ_c. As before this results in a reduced R_{roll}. Also as before the converse is true for negative k_X. The inclusion of the $2\hat{\rho}A_{,X\phi}X$ term therefore strengthens the tendencies observed above: The

mass-lifting branch becomes yet more fine-tuned in order to obtain a thin-shell screened solution, while for the mass-lowering branch the parameter-space corresponding to thin-shell screening is broadened.

This concludes our brief survey of how radial solutions around a massive source are modified by the introduction of derivative terms into the conformal factor. As a generic feature, we observe that a positive k_X, linear dependence on X, which leads to a mass-lifting mechanism, simultaneously tightens thin-shell screening constraints, requiring a value of ϕ_i closer to ϕ_c in order to maintain a nearly constant field value inside the source. The converse is true for negative k_X.

5.6 Disformal chameleons: A no-go theorem

Until here we have only been considering conformal couplings to matter. Here we wish to investigate whether switching on the disformal B^2 term (5.11) can result in interesting modifications to the chameleon mechanism (for other disformal dark energy models we refer to [156, 157, 158]). As such, matter now couples to the full disformal matter metric

$$\tilde{g}_{\mu\nu} = A^2(\phi, X)g_{\mu\nu} + B^2(\phi, X)\partial_\mu\phi\partial_\nu\phi. \tag{5.56}$$

Note that such a metric always has a well-defined inverse [24] given by

$$\tilde{g}^{\mu\nu} = A^{-2}(\phi, X)\left(g_{\mu\nu} - B^2(\phi, X)C^{-1}(\phi, X)\partial^\mu\phi\partial^\nu\phi\right), \tag{5.57}$$

subject to the condition that $A^2(\phi, X) \neq 0 \neq C(\phi, X)$, where $C = A^2(\phi, X) + B^2(\phi, X)X$. We then have

$$\mathcal{S} = \int d^4x\sqrt{g}\left(\frac{M^2}{2}R - \frac{1}{2}X - V(\phi)\right) + \mathcal{S}_m\left(A^2(\phi, X)g_{\mu\nu} + B^2(\phi, X)\partial_\mu\phi\partial_\nu\phi, \Psi_i\right). \tag{5.58}$$

Just as in the conformal case we consider a non-relativistic, pressureless source. In addition we here focus on static solutions for which $\partial_0\phi = 0$ (requiring spherical symmetry could reduce this further to $\phi = \phi(r)$). As such our assumption list now is

- The stress-energy tensor describes a pressureless, non-relativistic fluid ($\tilde{T}^0_0 = -\tilde{\rho}$) with all other stress-energy tensor components vanishing.
- ϕ's dynamics are described by a static solution ($\partial_0\phi = 0$).

As before we may also assume a static, uniform source in the matter frame ($\tilde{\nabla}_\alpha\tilde{T}^{\mu\nu} = 0$ inside the source), but this won't be necessary for the argument here. Instead we are solely interested

in whether a chameleon mechanism can arise inside a source as a consequence of disformal contributions here at all.

The equation of motion (5.14) then immediately gives rise to a no-go theorem. This is because all disformal (B^2-dependent) terms in fact involve a contraction of one of the following forms

$$\partial_\mu\phi\partial_\nu\phi\tilde{T}^{\mu\nu}, \qquad \partial_\mu\phi\tilde{T}^{\mu\nu}, \qquad \partial_\alpha\partial_\mu\phi\tilde{T}^{\mu\nu}, \qquad \partial_\mu\phi\nabla_\nu\tilde{T}^{\mu\nu}. \tag{5.59}$$

As such all disformal contributions vanish for a static solution around a non-relativistic, pressureless source, since the only non-zero component of the stress energy tensor is $\tilde{T}^{00}$, which vanishes when contracted with $\partial_0\phi$ for a static solution. Chameleonic effects from the disformal coupling are therefore suppressed when considering e.g. solutions around the earth. We do expect disformal effects to play an important role when computing relativistic or non-static corrections to this solution, however. Nevertheless, this shows that, with the mild assumptions implemented above, the dominant contribution to the chameleon profile necessarily has to come from a conformally coupled matter metric. Disformal effects here only come in at next order in relativistic corrections and for non-static profiles. Note that we have focused on chameleonic behavior here. As comparison with the discussion of $\mathcal{J}_5$ and $\mathcal{J}_6$ (5.18) in section 5.3 shows, Vainshtein-like screening can arise as a consequence of disformal couplings.[12]

5.7 Summary

The main results of this chapter can be summarized as follows.

- A derivative chameleon model described by (5.13), where all matter is universally and minimally coupled to a metric $\tilde{g}_{\mu\nu} = A^2(\phi, X)g_{\mu\nu}$, generically comes with a new mass-altering mechanism. This separates into three branches depending on the local energy density $\hat{\rho}$ and the conformal factor A: A mass-lifting ("screening") branch, where derivative effects lead to an additional enhancement of the effective chameleon mass compared with the non-derivative $A^2(\phi, X \to 0)$ limit. For high energy-densities above some $\hat{\rho}_{crit}$ this transitions into a ghost-like branch in which the effective potential becomes unstable. Thirdly there is a stable mass-lowering branch.

- This suggests that derivative chameleon models which are not plagued by ghost-like instabilities typically lower the effective chameleon mass. Implementing a mass-lifting mechanism while avoiding ghost-like instabilities for a range of different densities $\hat{\rho}$ requires some

[12] Once the metric $\tilde{g}_{\mu\nu}$ is promoted to a building block for further gravitational parts of the action as in [18] - e.g. a cosmological-constant-like term $\sqrt{\tilde{g}}\Lambda$, an extrinsic curvature $\tilde{K}_{\mu\nu}$ and a Ricci scalar $\tilde{R}$ for $\tilde{g}_{\mu\nu}$ - then a full galileon/Horndeski type solution can be obtained, making the Vainshtein screening in question more robust.

additional "engineering", e.g. introducing either an appropriate energy density cutoff or a density-dependent conformal factor.

- A very light, but massive, cosmological scalar can in principle be chameleon-screened by a purely derivative-dependent conformal factor $A^2(X)$ in the mass-lifting branch, which offers added protection from quantum corrections over $A^2(\phi)$ due to the presence of the shift symmetry $\phi \to \phi + c$.

- The position of the minimum ϕ_{min} of an effective chameleon potential V_{eff} cannot be affected by derivative dependent terms (X-dependent as well as for higher order terms). This is the reason why for $A^2(X)$ to yield chameleon screening a (small) mass term is needed, whereas $A^2(\phi, X)$ can produce chameleon screening even in the absence of an explicit mass-term. Consequently, if we start with a runaway potential, e.g. $V(\phi) \sim \phi^{-n}$ for some positive n, which does not contain any explicit mass term, this means a ϕ-dependent conformal factor is necessary in order to produce chameleon screening.

- Radial solutions and the thin-shell mechanism are modified in derivative chameleon models. The region of parameter-space exhibiting thin-shell screening is reduced in the mass-lifting branch and enhanced in the mass-lowering branch of solutions.

- Disformal contributions to the matter metric cannot source chameleon phenomenology for a static solution around a non-relativistic, pressureless source (they can source Vainshtein-screened solutions though).

To conclude, chameleon models provide a well-established framework for reconciling the presence of a light cosmological scalar that drives cosmic acceleration with small scale fifth force constraints. Here we have shown how chameleon-type setups can be generalized to derivative-dependent conformal couplings, suggesting that this naturally generates a further mass-changing mechanism. The mass-lifting branch offers the possibility to alleviate the fine-tuning involved in making a chameleon fit fifth force constraints and also comes with the exciting promise of making chameleon models more robust from a quantum perspective by endowing them with a shift symmetry. However, this branch also faces a ghost-problem which may be answered by introducing a $\hat{\rho}$-dependent conformal factor or a $\hat{\rho}$-dependent cutoff for the theory. For the mass-lowering branch no such problems exist, suggesting that a mass-suppressing mechanism is in fact a typical feature of derivative chameleon models. As such we hope that this work contributes to the enterprise of providing a wider survey of the ways in which chameleon-phenomenology might be implemented, realizing what types of challenges it has to face, and therefore putting such models on a firmer footing by showing what requirements stable chameleon models have to meet.

Chapter 6

Emergent Galileons

6.1 Introduction

In the previous chapters we saw how non-canonical scalar fields of the $P(X, \phi)$ type arise from disformal bimetric structures in inflationary settings and we also already explored the disformal metric relation in a chameleon setting, leading us to construct generalized derivative chameleon models. However, in the chameleonic setting we only found the conformal subgroup of the disformal relation to be of relevance. To close the cycle, here we investigate what type of effect disformal and indeed triformal metrics (1.7) have on modified gravity scenarios. In section 1.3.3 and indeed in chapter 5 we already pointed out that Vainshtein screening is associated with effective metrics of the type $\tilde{g}_{\mu\nu} = g_{\mu\nu} + \partial_\mu\pi\partial_\nu\pi$, i.e. of the disformal type and here we wish to develop this observation further.

In this context Galileons [29] and in particular their extensions/generalizations [33, 34][1] are a particularly intriguing specimen here, aiming to construct the most general ghost-free scalar-tensor theories. For further related constructions see [159, 160, 161, 162] In the standard Galileon theory this comes paired with a Galilean shift symmetry, which also leads to a non-renormalization theorem [35, 36] (and hence technical naturalness) for the Galileon terms. This coincides with purely second order equations of motion. In the generalized scenarios proposed by [33, 34], ghost-freedom is maintained by requiring equations of motion which are at most second order, but the absence of a shift symmetry now also allows first order terms in the equations of motion.

Our aim here then is to show that Galileons naturally arise when considering general bimetric

[1] Note that, while these theories are "extensions" from a model-building point of view, Horndeski's construction [33] historically significantly predates the Galileon model of [29].

scalar-tensor theories with the metric proposed at the very start of this thesis in 1.1.2.[2]

$$\tilde{g}_{\mu\nu} = g_{\mu\alpha}\left(\delta^{\alpha}_{\nu} + B\partial^{\alpha}\pi\partial_{\nu}\pi + C\partial^{\alpha}\partial_{\nu}\pi\right). \tag{6.1}$$

Constructing geometric invariants similarly to the probe brane construction [18] reviewed in section 1.3.3, this will afford us with a geometric understanding of Galileons at the level of an effective 4D theory without the need to wed our proposal to a particular UV-completion (be it higher-dimensional or otherwise).

The plan for the chapter is very straightforward. In section 6.2 we review the standard Galileon construction and introduce some notation by [34] which is instrumental in further computations as well as in establishing the equivalence between classes of theories (non-trivially) related by total derivatives. In section 6.3 we show how a disformally related metrics generate subset of Galileon terms, before showing in section 6.4 how the full set of Galileon terms follows from considering the full triformal relation. Interestingly, while the disformal relation can only generate two Galileon terms via the two Lovelock invariants available in 4D [30, 31], the triformal relation in fact generates all Galileon terms only via the single "cosmological constant" Lovelock term $\sqrt{\tilde{g}}$. Finally we summarize our finding and provide an outlook onto further research in section 6.5.

6.2 Galileons

We reviewed Galileon models [29] in section 1.3.3. Here we summarize some salient features of such models and introduce a more compact notation. This will be extremely useful when manipulating similar models in the following sections and also makes the relation between classically equivalent Lagrangians clearer (i.e. ones which are related by total derivatives and consequently produce identical equations of motion). But firstly let us remind ourselves of the DBI-Galileon probe brane Lagrangian [18] discussed in section 1.3.3. We have

$$\mathcal{L} = \sqrt{g}F\left(\tilde{g}_{\mu\nu}, \tilde{R}_{\alpha\beta\gamma\delta}, \tilde{K}_{\rho\sigma}, \tilde{\nabla}_{\epsilon}\right) + A\pi, \tag{6.2}$$

where the form of the Lagrangian is protected by a non-linearly realized 5D Poincare invariance. In particular we find that the Lagrangian can be written as a sum of Lovelock invariants [30, 31,

[2] We note that for the purposes of this chapter we have defined the disformal and triformal factors to be B and C and not B^2 and C^2 as in (1.7) in the introduction.

18, 65]

$$F_\lambda = -\lambda \int \mathrm{d}^4x\sqrt{-g} = -\lambda \int \mathrm{d}^4x\sqrt{1+(\partial\pi)^2} \tag{6.3}$$

$$F_K = -M_5^3 \int \mathrm{d}^4x\sqrt{-g}K = M_5^3 \int \mathrm{d}^4x \ \left([\Pi] - \gamma^2[\phi]\right) \tag{6.4}$$

$$F_R = \frac{M_4^2}{2} \int \mathrm{d}^4x\sqrt{-g}R = \frac{M_4^2}{2} \int \mathrm{d}^4x\, \gamma \left(([\Pi]^2 - [\Pi^2]) + 2\gamma^2 \left([\phi^2] - [\Pi][\phi]\right)\right) \tag{6.5}$$

$$\begin{aligned} F_{GB} &= -\beta\frac{M_5^3}{m^2} \int \mathrm{d}^4x\sqrt{-g}\,\mathcal{K}_{\mathrm{GB}} \\ &= \beta\frac{M_5^3}{m^2} \int \mathrm{d}^4x\,\gamma^2\Big(\frac{2}{3}\left([\Pi]^3 + 2[\Pi^3] - 3[\Pi][\Pi^2]\right) + 4\gamma^2([\Pi][\phi^2] - [\phi^3]) \\ &\qquad - 2\gamma^2([\Pi]^2 - [\Pi^2])[\phi]\Big)\,, \end{aligned} \tag{6.6}$$

where $\gamma = 1/\sqrt{1+(\partial\pi)^2}$ and we remember that $\Pi_{\mu\nu} = \partial_\mu\partial_\nu\pi$ and square brackets [...] denoting the trace operator (w.r.t. the Minkowski metric $\eta_{\mu\nu}$). Also $[\phi^n] \equiv \partial\pi\,.\,\Pi^n\,.\,\partial\pi\,.$ In this chapter we will focus on the non-relativistic limit of Galileon terms. In flat Minkowski space-time and in the non-relativistic limit $(\partial\pi)^2 \ll 1$ one then finds that the above Lovelock invariants reduce to

$$F_2 = S_\lambda^{NR} = -\frac{\lambda}{2} \int \mathrm{d}^4x\,(\partial\pi)^2 \tag{6.7}$$

$$F_3 = S_K^{NR} = \frac{M_5^3}{2} \int \mathrm{d}^4x\,(\partial\pi)^2\,\Box\pi \tag{6.8}$$

$$F_4 = S_R^{NR} = \frac{M_4^2}{4} \int \mathrm{d}^4x\,(\partial\pi)^2\ \left((\Box\pi)^2 - (\partial_\mu\partial_\nu\pi)^2\right) \tag{6.9}$$

$$F_5 = S_{GB}^{NR} = \beta\frac{M_5^3}{3m^2} \int \mathrm{d}^4x\,(\partial\pi)^2\left((\Box\pi)^3 + 2(\partial_\mu\partial_\nu\pi)^3 - 3\Box\pi(\partial_\mu\partial_\nu\pi)^2\right)\,. \tag{6.10}$$

Note that in the above we are ignoring the the tadpole term $F_1 = \pi$, which also satisfies the required symmetries and can therefore be added to any Galileon Lagrangian as well.

In order to keep notation concise we will follow [34] and define

$$\pi_\mu \equiv \partial_\mu\pi\,, \qquad \pi_\mu^\nu \equiv \partial_\mu\partial^\nu\pi. \tag{6.11}$$

Written in this way one can then establish the following notation, defining a shorthand for $[\phi^i] = \langle i\rangle$ and similarly a shorthand for $[\Pi^i] = [i]$ by

$$[i] \equiv \pi^{\mu_1}{}_{\mu_2}\pi^{\mu_2}{}_{\mu_3}\pi^{\mu_3}{}_{\mu_4}\cdots\pi^{\mu_i}{}_{\mu_1} \tag{6.12}$$

$$\langle i\rangle \equiv \pi_{\mu_1}\pi^{\mu_1}{}_{\mu_2}\pi^{\mu_2}{}_{\mu_3}\pi^{\mu_3}{}_{\mu_4}\cdots\pi^{\mu_i}{}_{\mu_{i+1}}\pi^{\mu_{i+1}}. \tag{6.13}$$

The simplest cases are therefore given by

$$[1] = \Box\pi\,, \qquad [2] = \pi^{\alpha}{}_{\beta}\pi^{\beta}{}_{\alpha} \tag{6.14}$$

$$\langle 1\rangle = \pi_{\alpha}\pi^{\alpha}{}_{\beta}\pi^{\beta}, \qquad \langle 2\rangle = \pi_{\alpha}\pi^{\alpha}_{\beta}\pi^{\beta}_{\sigma}\pi^{\sigma}. \tag{6.15}$$

It is also useful to set up notation at this point which will later on allow us to straightforwardly establish properties such as mass dimensions of various terms (e.g. in order to aid power-counting arguments). As such we denote the number of derivatives ∂, n, and the number of factors of the field π, m, by (n, m). In particular this means that the following holds

$$\bar{X}^n \rightarrow (2n, 2n) \tag{6.16}$$

$$[n] \rightarrow (2n, n) \tag{6.17}$$

$$\langle n\rangle \rightarrow (2n+2, n+2), \tag{6.18}$$

where we have defined $\bar{X} \equiv \partial_\mu\pi\partial^\mu\pi$ (in order to keep notation concise and in contrast to $X \equiv -1/2\bar{X}$). Having established this notation, the non-relativistic Galilean Lagrangians (6.7)-(6.10) can now be rewritten very concisely as

$$\mathcal{L}_2 = \bar{X} \tag{6.19}$$

$$\mathcal{L}_3 = \bar{X}[1] \tag{6.20}$$

$$\mathcal{L}_4 = \bar{X}\left([2] - [1]^2\right) \tag{6.21}$$

$$\mathcal{L}_5 = \bar{X}\left(3[1][2] - 2[3] - [1]^3\right). \tag{6.22}$$

A general Galileon Lagrangian can then be written as a linear superposition of these terms $\mathcal{L} = \sum_i c_i\mathcal{L}_i$.[3] Note that we have ignored the tadpole term $\mathcal{L}_1 = \pi$ in the above again and that there are no terms at higher derivative order (here meaning the number of derivatives in the term or analogously: the term's mass-dimension) than $\mathcal{L}_5$. The corresponding equations of motion are given by $\mathcal{E}_i = \frac{\delta\mathcal{L}_i}{\delta\pi} = 0$ and we find

$$\mathcal{E}_2 = [1] \tag{6.23}$$

$$\mathcal{E}_3 = [2] - [1]^2 \tag{6.24}$$

$$\mathcal{E}_4 = -2[3] + 3[1][2] - [1]^3 \tag{6.25}$$

$$\mathcal{E}_5 = 6[4] - 8[1][3] - 3[2]^2 + 6[1]^2[2] - [1]^4. \tag{6.26}$$

[3] We have therefore absorbed all "coupling constants" $\lambda, M_5, M_4, \ldots$ into the c_i.

Written in this form one can also spot the instructive pattern

$$\mathcal{L}_n = \bar{X}\mathcal{E}_{n-1}, \tag{6.27}$$

which is directly related to the fact that all Galileon Lagrangians can be constructed using Euler hierarchies [163, 164, 165], with the series terminating after $\mathcal{L}_5$ in four dimensions. Two further observations may be made here. Firstly note that $X\left(6[4]-8[1][3]-3[2]^2+6[1]^2[2]-[1]^4\right)$ is a total derivative, explicitly confirming that the series terminates after five Lagrangians are generated. Secondly, with Lagrangians as written in (6.27), it is straightforward to generalize the standard Galileon with purely second order equations of motion [29] to the most general scalar field theory with equations of motion of order two or lower. The result [34] is a theory with Lagrangians $\mathcal{L}_i = f(\pi, \bar{X})\mathcal{L}_i^{\text{Galileon}}$.

While there are no integration-by-parts ambiguities at the level of the equations of motion, the original Lagrangians $\mathcal{L}_i$ may of course be represented in different forms. The form chosen above can be related to other forms in the literature by noticing that the following combinations are total derivatives [166]

$$TD_3 = \bar{X}[1]+2\langle 1\rangle \tag{6.28}$$

$$TD_4 = 2\left([1]\langle 1\rangle-\langle 2\rangle\right)+\bar{X}\left([1]^2-[2]\right) \tag{6.29}$$

$$TD_5 = 2\left([1]^2-[2]\right)\langle 1\rangle-4\left([1]\langle 2\rangle-\langle 3\rangle\right)+\bar{X}\left([1]^3+2[3]-3[1][2]\right). \tag{6.30}$$

We will use these identities repeatedly in deriving the results presented in the following section.

6.3 Disformal Galileons

In the previous section we reviewed how Galileon terms can be constructed from geometric invariants by considering the effective theory on a 4-brane embedded in a fifth dimension. Of course this higher-dimensional motivation is attractive, but not ultimately necessary for considering Galileon-like theories. One may alternatively simply treat the Galileon Lagrangians as an effective 4D field theory, constructing the most general ghost-free scalar-tensor theory (and further requiring technical naturalness if we restrict ourselves to the standard Galileon model with purely second order equations of motion). Approaching the problem in this way, the geometric understanding is lost, but we also free ourselves from any commitment to a particular UV-completion. The approach promoted here is, so to speak, trying to have the best of both worlds: Obtain a geometric understanding in terms of (effective) metric structures at the level of a purely four-dimensional field theory without reference to possible UV-completions (higher-dimensional or otherwise). As we will see, this hybrid approach naturally generates Galileon solutions and

their various extensions, showing that very simple metric theories constructed from the general scalar-tensor metrics (1.8) recover known Galileon-like solutions in a very intuitive way.

In contrast to the 5D DBI probe brane construction alluded to above, in a purely 4D theory there are only two Lovelock invariants for any geometry specified by a given metric $g_{\mu\nu}$ [30, 31]

$$F_\lambda = -\lambda \int \mathrm{d}^4x\sqrt{-g} \tag{6.31}$$

$$F_R = \frac{M_4^2}{2} \int \mathrm{d}^4x\sqrt{-g}R, \tag{6.32}$$

where $R[g_{\mu\nu}]$, i.e. the Ricci scalar is formed with the metric $g_{\mu\nu}$. Here we wish to explore a purely 4D effective theory with these two simple Lovelock invariants and general disformal metric. Note that the remaining terms F_K and F_{GB}, which we encountered above, only arise as boundary terms in a five-dimensional theory [18], in particular making explicit reference to the extrinsic curvature $K_{\mu\nu}$, which satisfies

$$K_{\mu\nu} = -\frac{\partial_\mu\partial_\nu\pi}{\sqrt{1+(\partial\pi)^2}}. \tag{6.33}$$

As discussed in section 1.3.3, in the DBI probe brane case, the induced 4D metric is precisely of the disformal type with trivial disformal factor B: $\tilde{g}_{\mu\nu} = g_{\mu\nu} + \partial_\mu\pi\partial_\nu\pi$. This further motivates focusing on the disformal relation.

As such we reduce the triformal relation to the purely disformal case with $C = 0$ with trivial conformal factor $A = 1$

$$\tilde{g}_{\mu\nu} = g_{\mu\alpha}\left(\delta^\alpha_\nu + B\partial^\alpha\pi\partial_\nu\pi\right). \tag{6.34}$$

Focusing on the first Lovelock invariant (6.31) for the metric $\tilde{g}_{\mu\nu}$, we have to work out $\mathrm{Det}[\tilde{g}_{\mu\nu}]$. To this end we can use the relation

$$\mathrm{Det}\left[I+\alpha\right] = \mathrm{Exp}\left[\mathrm{Tr}\left[\mathrm{Ln}\left[I+\alpha\right]\right]\right] \tag{6.35}$$

where we can expand $\mathrm{Det}[\tilde{g}_{\mu\nu}] = \mathrm{Det}[g_{\mu\nu}]\mathrm{Det}[I+\alpha]$, where

$$\alpha = B\partial^\alpha\pi\partial_\nu\pi. \tag{6.36}$$

Fortuitously the disformal relation has the special property that $(\mathrm{Tr}[\alpha])^2 = \mathrm{Tr}[\alpha^2]$, allowing a

re-summation of terms in the determinant expansion. Specifically we have

$$\begin{aligned}
\tilde{g} &= g\text{Exp}\left[\text{Tr}\left[\alpha - \frac{\alpha^2}{2} + \frac{\alpha^3}{3} - \dots\right]\right] && (6.37)\\
&= g\text{Exp}\left[\text{Tr}\left[\alpha\right] - \text{Tr}\left[\frac{\alpha^2}{2}\right] + \text{Tr}\left[\frac{\alpha^3}{3}\right] - \dots\right] && (6.38)\\
&= g\text{Exp}\left[B\bar{X} - \frac{B\bar{X}^2}{2} + \frac{B\bar{X}^3}{3} - \dots\right] && (6.39)\\
&= g(1 + B\bar{X}). && (6.40)
\end{aligned}$$

This means the first Lovelock invariant corresponding to the general 4d effective metric $\tilde{g}_{\mu\nu}$ is

$$F_\lambda = -\lambda \int \mathrm{d}^4x\sqrt{-\tilde{g}} = -\lambda \int \mathrm{d}^4x\sqrt{-g}\sqrt{1 + B\bar{X}}. \tag{6.41}$$

Having specified the metric $\tilde{g}_{\mu\nu}$ we can also read off the expression for the second Lovelock invariant from (6.5)

$$F_R = \frac{M_4^2}{2}\int \mathrm{d}^4x\sqrt{-g}R = \frac{M_4^2}{2}\int \mathrm{d}^4x\,\gamma\left(\left([1]^2 - [2]\right) + 2\gamma^2\left(\langle 2\rangle - [1]\langle 1\rangle\right)\right) \tag{6.42}$$

where γ is the Lorentz factor introduced above and we notice that, for trivial disformal factor $B = 1$, this may also be written as

$$\gamma = \sqrt{\frac{g}{\tilde{g}}}. \tag{6.43}$$

$\mathcal{L}_3$ and $\mathcal{L}_5$ on the other hand can not be generated in this way from a disformal $\tilde{g}_{\mu\nu}$. As such a purely disformal Galileon has Lagrangian $\mathcal{L}_{\text{disformal}} = c_1\pi + c_2\mathcal{L}_2 + c_4\mathcal{L}_4$. In other words, the disformal relation generates a subset of Galileon solutions. However, as shown by [29], a solution with $c_3 = c_5 = 0$ cannot give rise to stable and spherically symmetric solutions with de Sitter asymptotics (i.e. self-accelerating solutions). Instead of attempting to salvage the purely disformal Galileon we have constructed in this section, however, we will now move on to show how the natural extension of the disformal relation which we constructed in the first chapter, the triformal relation (1.7), can produce the full Galileon solution.

6.4 Triformal Galileons

We begin by reminding ourselves of the triformal relation, which for the purposes of this chapter we write

$$\tilde{g}_{\mu\nu} = g_{\mu\alpha}\left(\delta^\alpha_\nu + B\partial^\alpha\pi\partial_\nu\pi + C\partial^\alpha\partial_\nu\pi\right). \tag{6.44}$$

We have chosen a trivial conformal factor $A = 1$ (as discussed in chapter 5, if $A \neq 1$ is chosen, we expect chameleon screening to arise on top of the Vainshtein-screened solutions which we are interested in in this chapter) and where the conformal and triformal factors $B(\phi, X, \Box\phi, \ldots)$ and $C(\phi, X, \Box\phi, \ldots)$ are in principle arbitrary functions of scalar coordinate invariants formed with ϕ and its derivatives. In the Galileon context investigated here, two interesting observations can be made straightaway. Firstly note that the triformal term $C\partial^\alpha\partial_\nu\pi$ somewhat mimics the extrinsic curvature term (6.33) in the higher-dimensional setup discussed above in that it introduces a dependence on $\partial^\alpha\partial_\nu\pi$. Secondly, had we been unaware of the existence of Galileon-like solutions, an obvious worry when presented with the triformal relation would have been that the triformal factor introduces higher-order derivatives, which then lead to the appearance of ghosts. Here, however, our aim is to show that the higher-order derivative terms introduced into the action by a triformal factor are precisely of the Galileon type, so that the equations of motion for π remain second-order and no Ostrogradski ghosts [32] are introduced.

The aim of this section now is the following. Consider the first Lovelock invariant again

$$F_\lambda = -\lambda \int \mathrm{d}^4x\sqrt{-\tilde{g}}. \tag{6.45}$$

Starting from the triformal metric (6.44), we once again work out the determinant of $\tilde{g}_{\mu\nu}$ by expanding it as

$$\mathrm{Det}\left[I + \alpha\right] = \mathrm{Exp}\left[\mathrm{Tr}\left[\mathrm{Ln}\left[I + \alpha\right]\right]\right] \tag{6.46}$$

where now

$$\alpha = B\partial^\alpha\pi\partial_\nu\pi + C\partial^\alpha\partial_\nu\pi. \tag{6.47}$$

What we wish to show is that, in the limit where we can treat α as small ($\alpha < I$)[4] and where we are considering non-relativistic solutions ($\bar{X} \ll 1$) for constant disformal and triformal factors B, C, the first Lovelock invariant is enough to generate all the Galileon terms. In the future we intend to generalize this results, dropping the assumptions of small α and considering fully relativistic regimes while allowing for arbitrary B and C, but for the time being let us concentrate on the simple case in question.

Once again we have

$$\tilde{g} = g\mathrm{Exp}\left[\mathrm{Tr}\left[\alpha - \frac{\alpha^2}{2} + \frac{\alpha^3}{3} - \ldots\right]\right] \tag{6.48}$$

$$= g\mathrm{Exp}\left[\mathrm{Tr}\left[\alpha\right] - \mathrm{Tr}\left[\frac{\alpha^2}{2}\right] + \mathrm{Tr}\left[\frac{\alpha^3}{3}\right] - \ldots\right] \tag{6.49}$$

[4] In other words, we are considering regimes with weak and intermediate ($\mathcal{O}(1)$) coupling here, not ones where we have a strongly coupled solution with $\tilde{g}_{\mu\nu} \sim (\partial_\mu\partial_\nu\pi)^n$.

However, even the simplified triformal relation we are considering here, no longer has the property that $(\mathrm{Tr}[\alpha])^2 = \mathrm{Tr}[\alpha^2]$. This means we will not be able to straightforwardly re-sum all terms in the determinant expansion. However, for $\alpha < I$ we can truncate the expansion eventually. We therefore begin by explicitly writing down all relevant terms up to $\mathcal{O}(\alpha^4)$.

$$T_2 = \mathrm{Tr}\,[\alpha] = B\bar{X} + C[1] \tag{6.50}$$

$$T_3 = \mathrm{Tr}\left[\frac{\alpha^2}{2}\right] = \frac{B^2}{2}\bar{X}^2 + \frac{C^2}{2}[2] + 2BC\langle 1\rangle \tag{6.51}$$

$$T_4 = \mathrm{Tr}\left[\frac{\alpha^3}{3}\right] = \frac{B^3}{3}\bar{X}^3 + \frac{C^3}{3}[3] + BC^2\langle 2\rangle + B^2C\bar{X}\langle 1\rangle \tag{6.52}$$

$$T_5 = \mathrm{Tr}\left[\frac{\alpha^4}{4}\right] = \frac{B^4}{4}\bar{X}^4 + \frac{C^4}{4}[4] + \frac{1}{2}B^2C^2\bar{X}\langle 2\rangle + B^2C\bar{X}^2\langle 1\rangle + BC^3\langle 3\rangle + B^2C^2\langle 1\rangle^2 \tag{6.53}$$

Earlier on we introduced the notation (n, m) to denote the number of derivatives ∂, n, and the number of factors of the field π, m, for a given term. We also found

$$\bar{X}^n \to (2n, 2n) \qquad [n] \to (2n, n) \qquad \langle n\rangle \to (2n+2, n+2). \tag{6.54}$$

Now, when expanding out the exponential in(6.48), we find that each trace operator proportional to α^i contributes at order $n = 2i$ in number of derivatives[5]. However, cross-terms between different trace operators of course also contribute so that for each order n we find contributions from the following operators

$$n = 2 \;:\; T_1 \tag{6.55}$$

$$n = 4 \;:\; T_2 \;\&\; T_1^2 \tag{6.56}$$

$$n = 6 \;:\; T_3 \;\&\; T_1^3 \;\&\; T_2T_1 \tag{6.57}$$

$$n = 8 \;:\; T_4 \;\&\; T_1^4 \;\&\; T_2^2 \;\&\; T_3T_1 \;\&\; T_1^2T_2. \tag{6.58}$$

In table 6.1 all the resulting terms are summarized, ordered by their dimensions (n, m) and with arising total derivative combinations arising for each n, m. When considering all terms arising from the triformal determinant expansion, after some algebra we find that interestingly only terms of the form $(2(i-1), i)$ do not vanish up to total derivatives. Of course this is exactly the form of the Galileon terms which we may label as $\mathcal{L}_i$ and where the tadpole term π satisfies $(0, 1)$ so that we may explicitly verify that all terms have the form $(2(i-1), i)$ by noting that non-relativistic Galileon terms (6.19) are of the form $(0, 1), (2, 2), (4, 3), (6, 4), (8, 5)$. And indeed when considering all terms in the triformal determinant expansion that do not vanish up to total

[5]This is equivalent to mass dimension, if we choose π to have dimension zero. Otherwise the mass dimension of a given term is given by $n + mD$, where D is the mass dimension of the field π.

(n,m)	Terms	Total derivative combinations
(2,1)	$[1]$	$[1]$
(2,2)	$\bar{X}$	–
(4,2)	$[2], [1]^2$	$[2]-[1]^2$
(4,3)	$\langle 1\rangle, \bar{X}[1]$	$\bar{X}[1]+2\langle 1\rangle$
(4,4)	$\bar{X}^2$	–
(6,3)	$[3], [2][1], [1]^3$	$[1]^3+2[3]-3[2][1]$
(6,4)	$\langle 2\rangle, \bar{X}[2], \bar{X}[1]^2, [1]\langle 1\rangle$	$2\left([1]\langle 1\rangle-\langle 2\rangle\right)+\bar{X}\left([1]^2-[2]\right)$
(6,5)	$\bar{X}\langle 1\rangle, \bar{X}^2[1]$	$4\bar{X}\langle 1\rangle+\bar{X}^2[1]$
(6,6)	$\bar{X}^3$	–
(8,4)	$[4], [3][1], [2]^2, [1]^2[2], [1]^4$	$[1]^4+8[1][3]+3[2]^2-6[1]^2[2]-6[4]$
(8,5)	$\langle 3\rangle, \langle 2\rangle[1], \langle 1\rangle[1]^2, \bar{X}[2][1], \bar{X}[1]^3, \bar{X}[3]$	$2\left([1]^2-[2]\right)\langle 1\rangle - 4\left([1]\langle 2\rangle-\langle 3\rangle\right) + \bar{X}\left([1]^3+2[3]-3[1][2]\right)$
(8,6)	$\langle 1\rangle^2, \bar{X}\langle 2\rangle, \bar{X}[1]\langle 1\rangle, \bar{X}^2[1]^2, \bar{X}^2[2]$	$\bar{X}^2[2]-\bar{X}^2[1]^2+2\bar{X}\langle 2\rangle-4\bar{X}[1]\langle 1\rangle$
(8,7)	$\bar{X}^3[1], \bar{X}^2\langle 1\rangle$	$3\bar{X}^2[2]+\bar{X}^3[1]$
(8,8)	$\bar{X}^4$	–

Table 6.1: This table summarizes the terms arising in the triformal determinant expansion, grouping terms by (n,m), where n denotes the number of derivatives ∂ and m denotes the number of factors of the field π for a given term. We list terms up to $n=8$. Only contributions from the (white background) $(2(i-1),i)$ terms do not vanish up to total derivatives.

derivatives, after some more algebra we may write the result very concisely as

$$\tilde{g}=g\left(1+\bar{X}\sum_i c_i\mathcal{E}_i+\mathcal{O}\left(\bar{X}^2\right)\right), \tag{6.59}$$

where $\mathcal{E}_i$ are the equations of motion resulting from each non-relativistic Galileon term (6.23). It is now straightforward to write down the first Lovelock invariant since in the non-relativistic limit we have

$$\int\sqrt{-\tilde{g}}=\int\sqrt{-g}\left(1+\bar{X}\sum_i c_i\mathcal{E}_i-\mathcal{O}\left(\bar{X}^2\right)\right), \tag{6.60}$$

where, given that we are working in the non-relativistic limit, we may now also drop terms of order $\bar{X}^2$ and higher. This is the main result of this chapter. Since, as we saw earlier on, the non-relativistic Galileon Lagrangians can be written $\mathcal{L}_n=\bar{X}\mathcal{E}_{n-1}$, this result is enough to generate

all Galileon terms. For the minimal triformal metric $\tilde{g}_{\mu\nu}$ (6.44) with constant coefficients B and C, in the non-relativistic limit the first Lovelock invariant therefore satisfies

$$F_\lambda^{NR} = -\lambda \int \mathrm{d}^4x\sqrt{-\tilde{g}} = -\lambda \int \mathrm{d}^4x\sqrt{-g}\left(1 + \sum_i c_i\mathcal{L}_i\right), \tag{6.61}$$

where the $\mathcal{L}_i$ label the non-relativistic Galileon terms

$$\mathcal{L}_2 = \bar{X} \tag{6.62}$$
$$\mathcal{L}_3 = \bar{X}[1] \tag{6.63}$$
$$\mathcal{L}_4 = \bar{X}\left([2] - [1]^2\right) \tag{6.64}$$
$$\mathcal{L}_5 = \bar{X}\left(3[1][2] - 2[3] - [1]^3\right). \tag{6.65}$$

Not only does the first Lovelock invariant therefore generate all Galileon terms when starting with a triformal metric relation, it also uniquely specifies the "coupling constants" c_i, where one finds $c_i \propto BC^{(i-1)}$ for $i \geq 1$, i.e. ignoring the tadpole term[6].

What about terms of the form (n, m) with $n > 8$? These terms vanish up to total derivatives in four dimensions. This may be seen by comparison with the result of [34], who show that Lagrangian terms of the form $\mathcal{L} = \mathcal{T}_{(2i)}^{\mu_1\ldots\mu_i\nu_1\ldots\nu_1}\pi_{\mu_1\nu_1}\ldots\pi_{\mu_i\nu_i}$ only generate field equations of at most order 2, i.e. Galileon-like solutions. Here $\mathcal{T}_{(2i)}$ is a $2i$ contravariant tensor that is a function of π and $\partial\pi$ only and is totally antisymmetric in its first i indices μ_i as well as separately in its final i indices ν_i. Comparison with equation (6.48) shows that all terms generated by the first Lovelock invariant from a triformal metric and which satisfy $(n, m) = (2(i-1), i)$ are evidently of this form (up to total derivatives). The fact that these terms are totally antisymmetric in their first i indices μ_i is then enough to conclude that these terms vanish for $n > 8$, since there are at least 5 indices i then, so antisymmetry paired with a 4D space-time will eliminate these terms.

Finally we once again stress that in deriving the results presented here we have worked in the weak to intermediate coupling limit with $\alpha < I$. We leave it to future work to show that the result remains true even in the strong coupling regime where $\alpha > I$, but we here wish to point out that there is excellent reason to believe this is indeed the case. Normally, if we do an expansion in α in increasing powers of α^n, this expansion will of course not converge for $\alpha > 1$. However, even though $\alpha > 1$ in the strong coupling regime, a power expansion up to order α^4 will, as we have shown in this section, recover all terms which are not total derivatives. In these circumstances we are therefore not worried whether the exact expression for remaining higher order terms converges, since these will be total derivatives and hence irrelevant classically anyway.

[6]Note that the tadpole term can in fact be generated once we introduce a non-trivial conformal factor $A(\pi)$.

Finally note that, by subtracting the equivalent Lovelock invariant for the Einstein metric $g_{\mu\nu}$ we can still engineer a theory with an effective strongly coupled metric $q_{\mu\nu} \sim \partial_\mu \partial_\nu \pi$ even though $\alpha < I$. In the future we plan to explicitly verify these claims by constructing example models.

6.5 Summary

In this chapter we have related Galileons [29] to bimetric scalar-tensor theories. We showed how Galileon-like solutions are generated simply from the first Lovelock invariant (or equivalently: from $\sqrt{\tilde{g}}$) for a triformal metric $\tilde{g}_{\mu\nu}$. This is very intriguing in that the full Galileon spectrum of terms can be generated by a single and very simple term in the Lagrangian now, which is simply motivated by considering a general bimetric scalar-tensor theory. This also equips us with a geometric understanding of Galileon theories at the level of an effective 4D field theory as was previously only available within the context of a particular UV-completion (the DBI Galileon [18]). Moreover, while the triformal relation does generate all Galileon terms[7] and gives rise to a Lagrangian $\mathcal{L} = \sum_i c_i \mathcal{L}_i$, the coefficients c_i are uniquely fixed by the disformal and triformal factors B, C in the metric relation. This means a particular subset of solutions will be particularly motivated from this geometric perspective.

This leads us straightaway to pressing questions which will be addressed in future work. It will be very interesting to generalize the solutions obtained to non-constant disformal and triformal factors B, C. As shown by [33, 34] the most general scalar-tensor theory with field equations of order 2 or lower has Lagrangians of the form $\mathcal{L}_i = f(\pi, \bar{X})\mathcal{L}_i^{\text{Galileon}}$. This strongly suggests that letting B and C depend on $(\pi, \bar{X})$ will generalize the purely second order Galileon solutions generated in this chapter to the more general solutions with $\mathcal{L} = f(\pi, \bar{X})\mathcal{L}^{\text{Galileon}}$. The same result also implies that letting the disformal and triformal factors depend on $\Box\pi$ or other higher order coordinate invariants will introduce ghost-like behavior. We also hope to investigate constraints from the existence of spherically symmetric and stable solutions with de Sitter asymptotics, which put constraints on the c_i and hence also on B, C. Finally the fact that all terms which are not of the $(2(i-1), i)$ form vanish in the derivative expansion in the previous two sections might hint at the existence of a symmetry protecting generalized Galileons. For the standard Galileon this has been identified with a Galilean shift symmetry in the effective 4D theory [29] which may stem from 5D Poincare invariance [18], if we choose to embed the theory in a higher-dimensional context. However, for the more general Galileon-like solutions with $\mathcal{L} = f(\pi, \bar{X})\mathcal{L}^{\text{Galileon}}$ to our knowledge no corresponding symmetry has been shown to exist explicitly so far, so it will be an interesting task for the future to investigate whether such a symmetry can be found.

[7] We also saw how the purely disformal subset of solutions can generate two of the Galileon terms, $\mathcal{L}_2$ and $\mathcal{L}_4$, but not the others.

In summary we hope to have demonstrated that triformal bimetric theories and Galileon solutions are intimately linked, with the minimal triformal model with constant B, C generating the standard Galileon with purely second-order equations of motion and consequently protection from non-renormalization theorems. The geometric understanding of Galileon theories both shifts the focus onto a particularly motivated subgroup of Galileon solutions, which we hope to investigate more in the future, and also will hopefully help us in gaining a better understanding of ghost-free scalar-tensor theories in the future.

Part IV

Conclusions

Chapter 7

Conclusions and Outlook

In this thesis we have attempted to show that the bimetric framework we have coined "disformal gravity" and its extensions are very powerful tools in improving our understanding of scalar-tensor theories, both in the context of primordial physics (inflationary or not) and of dark energy/modified gravity. How successful have we been? We would like to argue: very. Throughout this thesis we have seen that the disformal bimetric framework (and its triformal extension) provides a unified description of general single field/scalar-tensor models and allows us to systematically classify and construct such models. As such we have obtained a geometrical understanding of how such models arise in the context of generic 4D effective field theories. But perhaps this approach is really a calculational curiosity, only providing a, albeit useful, description of known models? As this thesis has shown, this is not the case. To this end let us briefly summarize the findings of the foregoing chapters.

- In chapter 1 we set the scene and showed how a general bimetric scalar-tensor theory can be constructed systematically. As a result we re-derived the disformal relation between metrics as initially proposed by [24] and found its natural extension, which we dubbed the "triformal relation". In an inflationary context, it was shown how disformal setups generate generic inflation models of the $P(X, \phi)$ type [19] and how their perturbative properties automatically give rise to an effective bimetric description [21]. In particular we pointed out a link between DBI-type solutions [22] and the minimal disformal model [23]. Finally, in a dark energy/modified gravity context, we showed how both chameleon- [15, 16] as well as Vainshtein-screened [54] solutions arise in generic scalar-tensor theories, emphasizing how these link to the conformal and purely disformal subgroups of the disformal relation respectively.

- In chapter 2 we investigated the perturbative properties of $P(X, \phi)$ models [19], as arising in a disformal context. In particular we focused on deriving the non-Gaussian phenomenology

of such models without assuming slow-roll conditions. It was found that a much broader region of parameter space is allowed by present observational constraints, when "fast-roll" effects are taken into account [9, 66, 1, 2]. This is the case because of a generic "fast-roll" suppression of the non-Gaussian amplitude we found, which occurs jointly with modified non-Gaussian shapes. We obtained bounds on slow-roll parameters from the induced running of the three-point correlation function and outlined how tensor modes may provide us with even stronger constraints in the future.

- In chapter 3 we made use of the fact that the disformal bimetric relation allows us to construct models with "superluminally" propagating perturbations without violating causality constraints e.g. by producing CCCs (closed causal curves). We reviewed the construction of such a disformal bimetric model and then investigated its non-Gaussian properties in an analogous fashion to the foregoing chapter. The shape found is predominantly equilateral and challenging to detect experimentally, but is uniquely defined in the large c_s ("superluminal") limit, thus making this model instantly falsifiable in future surveys [23, 3, 1].

- In chapter 4 we showed that disformally motivated $P(X, \phi)$ models also capture the dynamics of a large class of n-form models, which permit an effective scalar field description [27, 28, 5]. We derived dualities between various n-form descriptions. Focusing on a particular set of three-form models [27, 28], we obtained their two- and three-point correlation functions and investigated two particular example models with power-law and exponential three-form potentials respectively. Distinctive observational signatures are $n_s \sim 0.97$ irrespective of the potential used (assuming the number of e-folds $N \sim 60$) and generically equilateral non-Gaussianities. Interestingly no predominantly orthogonal amplitudes can be produced in the simplest setups with a power-law potential, reminiscent of the minimal DBI-type disformal setups discussed in chapter 1.

- In chapter 5 we investigated whether the disformal relation allows new ways of implementing the chameleon mechanism [15, 16]. We established a no-go theorem for disformal chameleons, but in the process also found a generalized conformal chameleon setup, which we named the "derivative chameleon" [4]. This setup includes a new mass-altering mechanism and we discussed how this chameleon mechanism might be implemented and how the chameleon mass as well as thin-shell screening effects are modified.

- In chapter 6 we finally investigated the full triformal relation and how it links to Vainshtein-screened dark energy/modified gravity solutions [6]. We showed how all Galileon terms [29] can be recovered from a single geometric invariant - the first Lovelock term $\sqrt{\bar{g}}$ [30, 31] - for a minimal triformal metric. We also conjectured how more general triformal setups

generalize the Galileon solutions. This is the first time Galileon type solutions have been constructed geometrically without the need to refer to higher-dimensional setups.

As such we believe that the bimetric disformal/triformal approach has both been extremely rewarding in pointing us to the existence of previously unknown models as well as providing us with a unified geometric understanding of generic scalar-tensor theories, both in inflationary and dark energy/modified gravity settings.

Several further lines of research suggest themselves following on from the findings discussed above. In particular we hope to investigate disformal models and their extensions in the following contexts

- Existing disformal theories assume an overall Riemannian geometry and metrics related by only one scalar degree of freedom [24]. Going beyond such restrictions should prove extremely rewarding, potentially giving further clues into how gravity can behave/be modified at large distances. In particular it may be interesting to generalize the bimetric relations constructed here to models with more than just one extra degree of freedom ϕ, subject to conditions such as ghost-freedom.

- Having constructed Galileon solutions [29] from triformal metrics in the dark energy/modified gravity context, a natural next step would be to implement these in an early universe setting as well. Following work on Galileon inflation [152, 161] we hope to investigate the particular subset of models motivated by triformal setups. Exploring their primordial signatures is an emphasis in this context.

- Surveys of the galaxy bispectrum and cluster counts are placing continuously improving constraints on primordial non-Gaussianity over a large range of scales. This opens up new possibilities to constrain models with running 3-point functions where amplitudes are enlarged on smaller scales. As we saw, $P(X, \phi)$ "fast-roll" models, as arising in a disformal setting, generically display this behavior. Further investigating large scale structure phenomenology for such models, specifically cluster counts and galaxy bispectra as well as future 21cm surveys, should consequently prove very rewarding.

- Constraining scale-dependent non-Gaussianities might allow identifying strong coupling scales for cosmological perturbations [167, 100, 75], i.e. determining scales at which e.g. the self-interactions of an inflaton become too large for a perturbative analysis to remain valid. Using observational data one may be able to identify potential strong coupling scales and hopefully extract generic signatures associated with strongly coupled single field models. Further exploring models such as the Gelaton scenario [45], where new degrees of freedom enter at the would-be strong coupling scale, should also provide new insights in this context.

- General slow-roll violations can be viewed as a combination of continuous slow-roll violations investigated in chapter 2 and temporary, feature-like, violations as described by the so-called "Generalized Slow Roll" approach [168]. Merging these two complementary approaches in order to develop a framework capable of analyzing generic types of fast-roll behavior in cosmology will be an exciting task for the future.

- Deriving the effective cubic action for general n-form theories, also when the extra physical *dof* explicitly break the duality to scalar field theories, should be very interesting. Studying three-point correlations in this way will allow identifying potentially new shapes of non-Gaussianity and investigating scale-dependence and the overall amplitude of the bispectrum. A long-term goal is the development of an effective field theory for n-form models permitting a homogeneous and isotropic (FLRW) cosmology. This should also prove useful in constraining and quantifying a potential breaking of overall homogeneity and isotropy (e.g. preferred directions due to a vector field). This approach goes hand in hand with generalizing the disformal relation to more additional degrees of freedom.

- Equivalence Principle violations are a natural consequence of emergent fifth forces that are the result of introducing a new scalar *dof* (see e.g. [148, 63, 149, 150, 151]). Examining this in the context of bimetric theories should allow constraining terms in the metric expansions e.g. the disformal and triformal factors B and C. It will therefore be interesting to further compute and investigate related effects on e.g. galactic halos, clustering behavior, galaxy luminosities etc. in this context (i.e. for both chameleon and Vainshtein-screened theories).

- In constructing Galileon-type theories, investigating the radiative stability and associated non-renormalization theorems for higher derivative terms has been a very powerful tool. But various avenues warrant further exploration. Perhaps there exist related symmetries protecting non-Galilean theories which still give rise to second order equations of motion and are ghost-free, but which do not impose a Galilean shift symmetry (see e.g. [160, 34]. Furthermore, in coupling Galilean scalars the prototypical coupling assumed has been a πT^{μ}_{μ} coupling. Studying generic bimetric couplings $\sqrt{\tilde{g}}\tilde{T}^{\mu}_{\mu}$ should prove useful in showing what (other) type of couplings are technically natural and and what classes of couplings are observationally acceptable. We hope to explore this avenue more in the future (see [63, 64] for some recent progress).

Much remains to be done, but we hope to have shown that the bimetric perspective has been a very instructive one in exploring single field/scalar-tensor effective field theories. When the theory of general relativity was formulated and explored in the early 20th century, it opened up new vistas of a geometric understanding of gravity and its associated phenomenology. In this spirit this thesis has aimed to review and extend geometrical interpretations of our current

best theories of primordial structure formation and late-time acceleration. Perhaps the main conclusion to draw then is that, when it comes to gaining a geometric understanding of gravity-related phenomena, there is still a lot left in the tank!

Part V

Appendices

Appendix A

Ostrogradski ghosts

Here we give a brief summary of Ostrogradski's theorem stating that equations of motion with higher than second-order derivatives exhibit instabilities. For simplicity we only show this for a single 1D point particle with position $q(t)$ here, closely following [169] - for details see [32].

Consider a Lagrangian $\mathcal{L}$ which is a function of $\ddot{q}$, i.e., $L(q, \dot{q}, \ddot{q})$. The corresponding equations of motion are

$$\frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} + \frac{\partial L}{\partial q} = 0 \ . \tag{A.1}$$

In principle (unless higher order terms cancel as in the Galileon scenario outlined in chapter 6 and section 1.3.3 of the introduction), this is a fourth order differential equation. A unique solution has to specify four initial conditions, e.g. $q(t_i)$, $\dot{q}(t_i)$, $\ddot{q}(t_i)$ and $\dddot{q}(t_i)$ at some initial time t_i. Casting this in Hamiltonian form requires four canonical variables, which we may choose to be

$$q_1 \equiv q \ , \quad q_2 \equiv \dot{q} \ , \quad \text{and} \quad p_1 \equiv \frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}} \ , \quad p_2 \equiv \frac{\partial L}{\partial \ddot{q}} \ . \tag{A.2}$$

If the Lagrangian is non-degenerate, i.e. $p_2 \equiv \partial L/\partial \ddot{q}$ can be inverted to determine $\ddot{q}$, we can Legendre-transform the Lagrangian w.r.t. coordinates $\dot{q} \equiv q_2$ and $\ddot{q}$ in order to find the Hamiltonian

$$H(q_1, q_2, p_1, p_2) = p_1 q_2 + p_2 \ddot{q}(q_1, q_2, p_2) - L\left(q_1, q_2, \ddot{q}(q_1, q_2, p_2)\right) \ . \tag{A.3}$$

Now a problem arises, because the Lagrangian is a function of only three variables, since it does not depend on $\partial \ddot{q}$. When expressing the Lagrangian in terms of canonical variables p_1 is therefore not needed to do so. As a result p_1 only enters the Hamiltonian linearly via the term $p_1 q_2$.

As a result H is unbounded from below and therefore describes an unstable system. If the *dof* described by the given equation of motion (q and its derivatives here) interacts with other *dof* at all, this instability means modes these other *dof* can be excited to arbitrarily high energies,

since there is no lower bound on the energy of the "q-system". This is particularly worrying if we would like to eventually quantize the theory. Then the vacuum becomes exponentially unstable since fluctuations can be excited at will.

Introducing even higher derivative terms into the equation of motion cannot change this conclusion. In fact, barring Galileon-like cancellations, the degree of the equations of motion is enhanced by 2 with each added higher derivative term. For a Lagrangian of degree $2 + n$, $n+1$ pairs of canonical variables are needed in order to write down the Hamiltonian, Expressing the Lagrangian in terms of canonical variables only requires the use of $n + 2$ of these. Hence n momenta appear only linearly via terms of the form $p_i\dot{q}_i(q_1, \ldots, q_{n+1}, p_{n+1})$. The Hamiltonian consequently has n unstable directions.

We finish with two points worth noting. Firstly, escaping the Ostrogradski instability while allowing higher derivative terms can in principle be achieved via two routes. Either L is degenerate or constraints are added to the system. If we add n constraints, these may be used to eliminate the n canonical variables which are associated with unstable directions. Secondly, note that having an equation of motion for a degree of freedom ϕ, which is second-order or lower in its derivatives, of course does not guarantee by itself that the theory is ghost-free.

Appendix B

Bekenstein's derivation of the disformal relationship

Bekenstein's derivation of the disformal relationship [24] is different to the one presented in chapter 1. For comparison we here give a brief summary of Bekenstein's argument. We begin by assuming a Finsler geometry, the most general geometry in which the line-element satisfies

$$ds^2 = f(x^\mu, dx^\nu) \qquad , \qquad f(x^\mu, \alpha dx^\nu) = \alpha^2 f(x^\mu, dx^\nu), \tag{B.1}$$

If $f(x^\mu, dx^\nu)$ is a simple quadratic in dx^μ then the geometry is Riemannian. It can be shown [24] that the line element may be written in terms of a so-called quasi-metric $\mathcal{G}_{\mu\nu}$, such that

$$ds^2 = \mathcal{G}_{\mu\nu} dx^\mu dx^\nu. \tag{B.2}$$

The important difference to a standard Riemannian theory is that $\mathcal{G}_{\mu\nu}$ can also depend on coordinate-increments dx^μ. In order to honor the spirit of covariance as well as in order to prevent the metric to be ill-defined locally (a dependence on dx^μ essentially amount to having a different metric depending on which way one is "moving", i.e. it is no longer enough to just specify position), Bekenstein requires $\mathcal{G}_{\mu\nu}$ as well as its inverse $\mathcal{G}^{\mu\nu}$ to be independent of dx^μ and to be a function of co-ordinate invariants only. In order to construct non-trivial co-ordinate invariants a new *dof* is needed, which is the reason why Bekenstein introduces a scalar ϕ. Invariants are then constructed out of ϕ and $\partial_\mu \phi$ as well as the Einstein metric $g_{\mu\nu}$. Higher derivative terms are precluded explicitly in Bekenstein's construction in order to trivially avoid higher derivative terms from appearing in the equations of motion (we have already seen that such terms can lead to Ostrogradski-type ghosts).

Unsurprisingly the result is a Riemannian metric, namely the disformal relation

$$\tilde{g}_{\mu\nu} = A^2(\phi, X) g_{\mu\nu} + B^2(\phi, X) \partial_\mu \phi \partial_\nu \phi. \tag{B.3}$$

We have already pointed out that this relationship is portrayed as yielding the most general metric $\tilde{g}_{\mu\nu}$ that is a function of $g_{\mu\nu}, \phi$ and its derivatives and which respects causality and the weak equivalence principle. Clearly this is only true subject to ignoring terms in $\partial^2\phi$ and higher. Also let us stress here that the disformal relation as it stands says nothing about causality or the weak equivalence principle. The fact that we have a metric formulation here with well defined past and future "light-cones" is enough in order to yield causal propagation for each individual metric. No closed causal curves (CCCs) exist [52]. In a bimetric theory an additional complication is the existence of a well-posed Cauchy problem, which requires the existence of a hyper-surface which is space-lie both with respect to $g_{\mu\nu}$ and $\tilde{g}_{\mu\nu}$. This puts a constraint on the conformal and disformal factors $A^2(\phi, X)$ and $B^2(\phi, X)$, but does not constrain the form of the disformal relation itself [52]. Similarly the weak equivalence principle is satisfied by virtue of having all matter universally couple to the same metric $\tilde{g}_{\mu\nu}$ and not because this metric has any particular form.

Appendix C

Details on the computation of the three-point function

C.1 The cubic action

In this appendix we give some further details showing how the non-Gaussian amplitudes discussed in chapter 2 and chapter 3 are computed explicitly. These amplitudes were first computed in [66] and we closely follow their presentation here.

We begin by recalling the cubic effective action derived in [37, 80] and valid for all $P(X,\phi)$ theories. This does not assume any slow-roll approximations to hold and is valid for arbitrary time-dependent speed of sound c_s:

$$\begin{aligned} S_3 &= {M_{Pl}}^2 \int \mathrm{d}t\mathrm{d}^3x \left\{ -a^3 \left[\Sigma \left(1 - \frac{1}{c_s^2} \right) + 2\lambda \right] \frac{\dot{\zeta}^3}{H^3} + \frac{a^3\epsilon}{c_s^4}(\epsilon - 3 + 3c_s^2)\zeta\dot{\zeta}^2 \right. \\ &+ \frac{a\epsilon}{c_s^2}(\epsilon - 2\epsilon_s + 1 - c_s^2)\zeta(\partial\zeta)^2 - 2a\frac{\epsilon}{c_s^2}\dot{\zeta}(\partial\zeta)(\partial\chi) \\ &+ \left. \frac{a^3\epsilon}{2c_s^2}\frac{d}{dt}\left(\frac{\eta}{c_s^2}\right)\zeta^2\dot{\zeta} + \frac{\epsilon}{2a}(\partial\zeta)(\partial\chi)\partial^2\chi + \frac{\epsilon}{4a}(\partial^2\zeta)(\partial\chi)^2 + 2f(\zeta)\left.\frac{\delta L}{\delta\zeta}\right|_1 \right\} . \end{aligned} \tag{C.1}$$

Dots here denote differentiation w.r.t. proper time t, ∂ is a spatial derivative and χ is shorthand notation for

$$\partial^2\chi = \frac{a^2\epsilon}{c_s^2}\dot{\zeta} . \tag{C.2}$$

For the final term in the cubic action (C.1), $2f(\zeta)\frac{\delta L}{\delta\zeta}|_1$, which denotes the variation of the quadratic action with respect to the perturbation ζ, we have

$$\left.\frac{\delta L}{\delta\zeta}\right|_1 = a\left(\frac{d\partial^2\chi}{dt} + H\partial^2\chi - \epsilon\partial^2\zeta\right) , \tag{C.3}$$

$$f(\zeta) = \frac{\eta}{4c_s^2}\zeta^2 + \frac{1}{c_s^2 H}\zeta\dot{\zeta} + \frac{1}{4a^2H^2}[-(\partial\zeta)(\partial\zeta) + \partial^{-2}(\partial_i\partial_j(\partial_i\zeta\partial_j\zeta))] + \frac{1}{2a^2H}[(\partial\zeta)(\partial\chi) - \partial^{-2}(\partial_i\partial_j(\partial_i\zeta\partial_j\chi))] \ , \tag{C.4}$$

where ∂^{-2} is the inverse Laplacian. Since $\frac{\delta L}{\delta\zeta}|_1$ is proportional to the linearized equations of motion, it can always be absorbed by a field redefinition

$$\zeta \to \zeta_n + f(\zeta_n) \ . \tag{C.5}$$

However, if $\eta = 0$ this simplifies considerably. This happens because only the ïňĄrst term in equation (C.4) contributes to the correlation function, as all other terms in $f(\zeta)$ contain one or more derivatives of ζ and consequently vanish outside the horizon, where ζ freezes in. For further details see [37, 80].

The standard calculation for the three-point function [89, 37, 80] can now be carried out, at first order in perturbation theory and in the interaction picture, leading to

$$\langle\zeta(t,\mathbf{k}_1)\zeta(t,\mathbf{k}_2)\zeta(t,\mathbf{k}_3)\rangle = -i\int_{t_0}^{t}\mathrm{d}t'\langle[\zeta(t,\mathbf{k}_1)\zeta(t,\mathbf{k}_2)\zeta(t,\mathbf{k}_3), H_{\rm int}(t')]\rangle \ , \tag{C.6}$$

where $H_{\rm int}$ is the Hamiltonian evaluated at third order in the perturbations and is directly derivable from (2.30) and vacuum expectation values are evaluated w.r.t. the interacting vacuum $|\Omega\rangle$. We recall from chapter 2 that, upon quantization, ζ can be expressed through creation and annihilation operators in the following way,

$$\zeta(y,\mathbf{k}) = u_k(y)a(\mathbf{k}) + u_k^*(y)a^\dagger(-\mathbf{k}) \qquad , \qquad [a(\mathbf{k}), a^\dagger(\mathbf{k}')] = (2\pi)^3\delta^3(\mathbf{k}-\mathbf{k}'). \tag{C.7}$$

where the mode functions $u_k(y)$ satisfy

$$u_k(y) = \frac{c_s^{1/2}}{aM_{\rm Pl}\sqrt{2\epsilon}}v_k(y) = \frac{c_s^{1/2}}{aM_{\rm Pl}2^{3/2}}\sqrt{\frac{\pi}{\epsilon}}\sqrt{-y}H_\nu^{(1)}(-ky). \tag{C.8}$$

With this information we can calculate the three point function for each term appearing in the action (2.30).

C.2 A useful integral

In computing the three-point function explicitly we frequently encounter integrals of the following type [66, 3].

$$\mathcal{C} = \int_{-\infty+i\varepsilon}^{y_{\rm end}}\mathrm{d}y\left(\frac{y}{y_{\rm end}}\right)^\gamma(-iy)^n e^{iKy} \, , \tag{C.9}$$

where y is sound horizon time as defined in (2.7). Only the imaginary part of this integral contributes to our calculation. $y_{\rm end} < 0$ denotes the end of the inflationary phase in sound horizon time and $K \equiv k_1 + k_2 + k_3$ as before. Note that for any observable mode $K|y_{\rm end}|$ is a small quantity. Focusing on scale-invariant solutions we have

$$\frac{H}{c_s^{1/2}} = \frac{H_{\rm end}}{\bar{c}_{s\,\rm end}^{1/2}} \quad (\gamma = 0)\,, \tag{C.10}$$

For $\gamma + n > -2$ the imaginary part of (C.9) is convergent as $y_{\rm end} \to 0$. Note especially the $i\epsilon$ term which is necessary in order to regularize the integral. We can now approximately extend the upper limit of integration to 0, which amounts to neglecting terms of higher order in $(k|y_{\rm end}|)$ One finds

$$\mathrm{Im}\,\mathcal{C} = -(K|y_{\rm end}|)^{-\gamma} \cos\frac{\gamma\pi}{2}\Gamma(1+\gamma+n)K^{-n-1}\,. \tag{C.11}$$

C.3 The individual three-point amplitude computation

The non-Gaussian amplitudes calculated in the way outlined above were given in (2.40). Here we illustrate their computation with an example, focusing on the contribution to the overall three-point function stemming from the $\zeta\dot{\zeta}^2$ term in (C.1). This can be expressed as

$$\begin{aligned}\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle_{\zeta\dot{\zeta}^2} \;=\; & i(2\pi)^3\delta^3(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)u_{k_1}(y_{\rm end})u_{k_2}(y_{\rm end})u_{k_3}(y_{\rm end}) \\ & \times\int_{-\infty+i\varepsilon}^{y_{\rm end}} \mathrm{d}y\frac{c_s}{a}\frac{a^3\epsilon}{c_s^4}(\epsilon-3+3c_s^2)u_{k_1}^*(y)\frac{du_{k_2}^*(y)}{dy}\frac{du_{k_3}^*(y)}{dy} + \text{perm.} + \text{c.c.},\end{aligned} \tag{C.12}$$

where *c.c.* denotes complex conjugates and we have applied the commutation relation (C.7) several times. Substituting in the mode functions and taking time-independent combinations outside of the integral this leads to

$$\begin{aligned}\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle_{\zeta\dot{\zeta}^2} \;=\; & i(2\pi)^3\delta^3(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\frac{H_{\rm end}^3(1-\epsilon-\epsilon_s)^4}{4^3\,\epsilon^2\,c_{s\,\rm end}^{3/2}}\frac{1}{\Pi_j k_j^3} \\ & \times\int_{-\infty+i\varepsilon}^{y_{\rm end}} \mathrm{d}y\frac{H}{c_s^{5/2}}(\epsilon-3+3c_s^2)(1-ik_1y)k_2^2k_3^2e^{iKy} + \text{perm.} + \text{c.c.}\end{aligned} \tag{C.13}$$

Background-dependent functions inside the integrand have one of two time-dependences, which both give rise to convergent integrals (C.11)

$$\frac{H}{c_s^{1/2}} = \frac{H_{\rm end}}{\bar{c}_{s\,\rm end}^{1/2}} \quad (\gamma = 0)\,, \qquad \frac{H}{c_s^{5/2}} = \frac{H_{\rm end}}{c_{s\,\rm end}^{5/2}}\left(\frac{y}{y_{\rm end}}\right)^{\alpha} \quad (\gamma = \alpha)\,, \tag{C.14}$$

where we recall that

$$\alpha = \frac{2\epsilon - \epsilon_s}{\epsilon_s + \epsilon - 1} = -\frac{4\epsilon}{1+\epsilon}, \tag{C.15}$$

where the second equality holds in the scale-invariant limit where $\epsilon_s = -2\epsilon$. Expressing the final answer in terms of quantities calculated at horizon-crossing (denoted by a bar) the amplitude found is

$$\begin{aligned}\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle_{\zeta\dot{\zeta}^2\ (\mathbf{I})} \;=\; & (2\pi)^3\delta^3(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\frac{\bar{H}^4(1+\epsilon)^4}{16\epsilon^2\,\bar{c}_s^4}\frac{1}{\Pi_j k_j^3}\frac{k_2^2k_3^2}{K} \\ & \times\left\{(\epsilon-3)\cos\frac{\alpha\pi}{2}\Gamma(1+\alpha)\left[1+(1+\alpha)\frac{k_1}{K}\right]+3\bar{c}_s^2\left[1+\frac{k_1}{K}\right]\right\}+\text{sym.}\end{aligned} \tag{C.16}$$

Appendix D

Non-Gaussian amplitudes - $\mathcal{O}(n_s - 1)$ corrections

The main results presented in this thesis and the computation in the previous appendix focused on the leading order non-Gaussian contribution corresponding to a scale-invariant spectral index $n_s = 1$. In this appendix we estimate the $\mathcal{O}(n_s - 1)$ corrections by approximating the Hankel function dependence of the "propagator" $u_k(y)$ in the limit $|ky| \ll 1$. Higher order propagator corrections have since been computed in [67, 69] and combined with the corrections derived by us in [1, 3] to complete the task of working out generic non-Gaussian amplitudes beyond the slow-roll paradigm. We refer to [69] for general slow-roll violating amplitudes taking higher-order propagator corrections into account, but here briefly summarize the estimate, which is significantly less cumbersome.

D.1 Mode functions $u_k(y)$

In section 2.2.1 we found that the mode functions for scalar perturbations are given by

$$u_k(y) = \frac{c_s^{1/2}}{aM_{\rm Pl}\sqrt{2\epsilon}} v_k(y) = \frac{c_s^{1/2}}{aM_{\rm Pl}2^{3/2}} \sqrt{\frac{\pi}{\epsilon}} \sqrt{-y} H_\nu^{(1)}(-ky). \tag{D.1}$$

We now expand at $|ky| \ll 1$ to obtain an approximate expression for the Hankel function $H_\nu^{(1)}(-ky)$

$$H_\nu^{(1)}(-ky) = -i \frac{2^\nu \Gamma(\nu)(-ky)^{-\nu}}{\pi} [1 + iky + \mathcal{O}(ky)^2] e^{-iky}\,. \tag{D.2}$$

which gives the following approximate expression for u_k,

$$u_k(y) \approx -i \frac{H(\epsilon + \epsilon_s - 1)}{2M_{\rm Pl}\sqrt{c_s k^3 \epsilon}} \left(\frac{-ky}{2}\right)^{3/2-\nu} (1 + iky) e^{-iky}, \tag{D.3}$$

where we have used $\Gamma(\nu \approx 3/2) \approx \sqrt{\pi}/2$ as scale invariance constrains $\nu - 3/2 \ll 1$. One can check that modes freeze out in the $y \to 0$ limit by comparing with (2.9) and noting that

$$\frac{H}{c_s^{1/2}} \sim (-y)^{\nu-3/2}\,. \tag{D.4}$$

Differentiating $u_k(y)$ with respect to y one also finds

$$u_k'(y) \approx -i\,\frac{H(\epsilon+\epsilon_s-1)}{2M_{\rm Pl}\sqrt{c_s k^3 \epsilon}}\left(\frac{-ky}{2}\right)^{3/2-\nu} k^2 y\; e^{-iky}. \tag{D.5}$$

D.2 The individual three-point amplitude computation

Here we show the computation of $\mathcal{A}_{\zeta\dot\zeta^2}$ in detail. All the other amplitudes are computed along the same lines. The corresponding three-point correlator is

$$\begin{aligned}\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle_{\zeta\dot\zeta^2} \;=\; & i(2\pi)^3\delta^3(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)u_{k_1}(y_{\rm end})u_{k_2}(y_{\rm end})u_{k_3}(y_{\rm end}) \\ & \times\int_{-\infty+i\varepsilon}^{y_{\rm end}}{\rm d}y\frac{c_s}{a}\frac{a^3\epsilon}{c_s^4}(\epsilon-3+3c_s^2)u_{k_1}^*(y)\frac{du_{k_2}^*(y)}{dy}\frac{du_{k_3}^*(y)}{dy}+{\rm perm.}+{\rm c.c.}\end{aligned} \tag{D.6}$$

The subscript "end" means that the quantity has to be evaluated at the end of the structure forming (scaling) phase. Making use of (2.9), (D.3) and (D.5) we take some time-independent combinations outside the integral and evaluate them at $y = y_{\rm end}$:

$$\begin{aligned}\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle_{\zeta\dot\zeta^2} \;=\; & i(2\pi)^3\delta^3(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\frac{H_{\rm end}^6(1-\epsilon-\epsilon_s)^6 2^{6\nu-9}}{4^3 {M_{Pl}}^4\,\epsilon^2\,c_{s\,\rm end}^3}\frac{(k_1k_2k_3)^{3-2\nu}}{\Pi_j k_j^3}|y_{\rm end}|^{6\left(\frac{3}{2}-\nu\right)} \\ & \times\int_{-\infty+i\varepsilon}^{y_{\rm end}}{\rm d}y(\epsilon-3+3c_s^2)\frac{a^2}{c_s^3}(1-ik_1y)k_2^2k_3^2y^2e^{iKy}+{\rm perm.}+{\rm c.c.}\,.\end{aligned} \tag{D.7}$$

Note that we have dropped a factor of $\Pi_j(1+ik_jy_{end})e^{-iKy_{end}}$ as this will be negligibly small for any observable modes. Evaluating the integral and re-expressing variables in terms of quantities calculated at sound horizon crossing (when, by convention, $y = K^{-1}$) we obtain

$$\begin{aligned}&\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle_{\zeta\dot\zeta^2} \;=\; (2\pi)^3\delta^3(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\frac{\bar H^4(\epsilon+\epsilon_s-1)^4 2^{6\nu-9}}{16{M_{Pl}}^4\epsilon^2\,\bar c_s^4}\frac{1}{\Pi_j k_j^3}\frac{(\Pi_j k_j^3)^{3-2\nu}}{K^{9-6\nu}} \\ &\times\frac{k_2^2k_3^2}{K}\left\{(\epsilon-3)\cos\frac{\alpha_2\pi}{2}\Gamma(1+\alpha_2)\left[1+(1+\alpha_2)\frac{k_1}{K}\right]+3\bar c_s^2\cos\frac{\alpha_1\pi}{2}\Gamma(1+\alpha_1)\left[1+(1+\alpha_1)\frac{k_1}{K}\right]\right\}+{\rm sym.}\end{aligned} \tag{D.8}$$

where we have defined

$$\alpha_1 = 3 - 2\nu = \frac{2\epsilon + \epsilon_s}{\epsilon_s + \epsilon - 1} \qquad ; \qquad \alpha_2 = \frac{2\epsilon - \epsilon_s}{\epsilon_s + \epsilon - 1}. \tag{D.9}$$

In evaluating this expression we have once again (cf. discussion in the previous appendix) made use of the convergent integral

$$\mathrm{Im}\,\mathcal{C} = -(K|y_{\rm end}|)^{-\gamma} \cos\frac{\gamma\pi}{2}\Gamma(1+\gamma+n)K^{-n-1}\,, \tag{D.10}$$

where the two types of common behavior encountered are now

$$\frac{a^2y^2}{c_s} = \frac{c_s}{(1-\epsilon-\epsilon_s)^2H^2} = \frac{c_{s\,\rm end}}{(1-\epsilon-\epsilon_s)^2H^2_{\rm end}}\left(\frac{y}{y_{\rm end}}\right)^{\alpha_1} \tag{D.11}$$

$$\frac{a^2y^2}{c_s^3} = \frac{1}{(1-\epsilon-\epsilon_s)^2H^2c_s} = \frac{1}{(1-\epsilon-\epsilon_s)^2H^2_{\rm end}c_{s\,\rm end}}\left(\frac{y}{y_{\rm end}}\right)^{\alpha_2}. \tag{D.12}$$

D.3 The full amplitudes

We recall that the 3-point correlation function may be conventionally expressed through the amplitude $\mathcal{A}$

$$\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle = (2\pi)^7\delta^3(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)P_\zeta{}^2\frac{1}{\Pi_j k_j^3}\mathcal{A}\,. \tag{D.13}$$

where, by convention, the power spectrum P_ζ in the above formula is calculated for the mode $K = k_1 + k_2 + k_3$. Computing the amplitude in this way for each term in the action (C.1) we can write

$$\mathcal{A} = \mathcal{A}_{\dot\zeta^3} + \mathcal{A}_{\zeta\dot\zeta^2} + \mathcal{A}_{\zeta(\partial\zeta)^2} + \mathcal{A}_{\dot\zeta\partial\zeta\partial\chi} + \mathcal{A}_{\epsilon^2}, \tag{D.14}$$

where $\mathcal{A}_{\epsilon^2}$ accounts for the $\partial\zeta\partial\chi\partial^2\chi$ and $(\partial^2\zeta)(\partial\chi)^2$ terms in the action (C.1). Computing each of the interaction vertices this yields

$$\begin{aligned}
\mathcal{A}_{\dot\zeta^3} &= \frac{1}{2\bar c_s^2}\left(\frac{k_1k_2k_3}{2K^3}\right)^{n_s-1}\left[(\epsilon+\epsilon_s-1)(f_X-1)\mathcal{I}_{\dot\zeta^3}(\alpha_2)+\bar c_s^2(\epsilon+\epsilon_s-1)\mathcal{I}_{\dot\zeta^3}(\alpha_1)\right]\ ;\\
\mathcal{A}_{\zeta\dot\zeta^2} &= \frac{1}{4\bar c_s^2}\left(\frac{k_1k_2k_3}{2K^3}\right)^{n_s-1}\left[(\epsilon-3)\mathcal{I}_{\zeta\dot\zeta^2}(\alpha_2)+3\bar c_s^2\mathcal{I}_{\zeta\dot\zeta^2}(\alpha_1)\right]\ ;\\
\mathcal{A}_{\zeta(\partial\zeta)^2} &= \frac{1}{8\bar c_s^2}\left(\frac{k_1k_2k_3}{2K^3}\right)^{n_s-1}\left[(\epsilon-2\epsilon_s+1)\mathcal{I}_{\zeta(\partial\zeta)^2}(\alpha_2)-\bar c_s^2\mathcal{I}_{\zeta(\partial\zeta)^2}(\alpha_1)\right]\ ;\\
\mathcal{A}_{\dot\zeta\partial\zeta\partial\chi} &= \frac{1}{4\bar c_s^2}\left(\frac{k_1k_2k_3}{2K^3}\right)^{n_s-1}\left[-\epsilon\,\mathcal{I}_{\dot\zeta\partial\zeta\partial\chi}(\alpha_2)\right]\ ;\\
\mathcal{A}_{\epsilon^2} &= \frac{1}{16\bar c_s^2}\left(\frac{k_1k_2k_3}{2K^3}\right)^{n_s-1}\left[\epsilon^2\,\mathcal{I}_{\epsilon^2}(\alpha_2)\right]\ ,
\end{aligned}\tag{D.15}$$

where

$$\begin{aligned}
\mathcal{I}_{\dot\zeta^3}(\alpha) &= \cos\frac{\alpha\pi}{2}\Gamma(3+\alpha)\frac{k_1^2k_2^2k_3^2}{K^3}\ ;\\
\mathcal{I}_{\zeta\dot\zeta^2}(\alpha) &= \cos\frac{\alpha\pi}{2}\Gamma(1+\alpha)\left[(2+\alpha)\frac{1}{K}\sum_{i<j}k_i^2k_j^2-(1+\alpha)\frac{1}{K^2}\sum_{i\neq j}k_i^2k_j^3\right]\ ;\\
\mathcal{I}_{\zeta(\partial\zeta)^2}(\alpha) &= -\cos\frac{\alpha\pi}{2}\Gamma(1+\alpha)\left(\sum_i k_i^2\right)\left[\frac{K}{\alpha-1}+\frac{1}{K}\sum_{i<j}k_ik_j+\frac{1+\alpha}{K^2}k_1k_2k_3\right]\\
&= \cos\frac{\alpha\pi}{2}\Gamma(1+\alpha)\left[\frac{1}{1-\alpha}\sum_j k_j^3+\frac{4+2\alpha}{K}\sum_{i<j}k_i^2k_j^2-\frac{2+2\alpha}{K^2}\sum_{i\neq j}k_i^2k_j^3\right.\\
&\quad\left.+\frac{\alpha}{(1-\alpha)}\sum_{i\neq j}k_ik_j^2-\alpha k_1k_2k_3\right]\ ;\\
\mathcal{I}_{\dot\zeta\partial\zeta\partial\chi}(\alpha) &= \cos\frac{\alpha\pi}{2}\Gamma(1+\alpha)\left[\sum_j k_j^3+\frac{\alpha-1}{2}\sum_{i\neq j}k_ik_j^2-2\frac{1+\alpha}{K^2}\sum_{i\neq j}k_i^2k_j^3-2\alpha k_1k_2k_3\right]\ ;\\
\mathcal{I}_{\epsilon^2}(\alpha) &= \cos\frac{\alpha\pi}{2}\Gamma(1+\alpha)(2+\alpha/2)\left[\sum_j k_j^3-\sum_{i\neq j}k_ik_j^2+2k_1k_2k_3\right]\ ,
\end{aligned}\tag{D.16}$$

where we remind ourselves that $\mathcal{A}_{\epsilon^2}$ accounts for the $\partial\zeta\partial\chi\partial^2\chi$ and $(\partial^2\zeta)(\partial\chi)^2$ terms in the action (2.30). $K=k_1+k_2+k_3$ is the sum of the lengths of the three wavevectors $\mathbf{k_1},\mathbf{k_2},\mathbf{k_3}$.

D.4 General expressions for $f_{\rm NL}$ and n_{NG}

Using the full amplitudes derived above, we may express the size of the (equilateral) non-Gaussian amplitude

$$f_{\rm NL}^{equi} = 30\frac{\mathcal{A}_{k_1=k_2=k_3}}{K^3}\,, \tag{D.17}$$

and its running with scale k

$$n_{\rm NG} - 1 \equiv {\rm d}\ln|f_{\rm NL}^{equi}|/{\rm d}\ln K. \tag{D.18}$$

Comparison with (D.15) and (D.16) shows that we can always schematically express

$$\ln|f_{\rm NL}^{equi}| = \ln|\mathcal{C}_1| + \ln|1 - \mathcal{C}_2\bar{c}_s^{-2}|. \tag{D.19}$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are functions of the slow-roll parameters $\mathcal{C}_i(\epsilon, n_s, f_X)$ or equivalently $\mathcal{C}_i(\epsilon, \epsilon_s, f_X)$. Specifically we find that

$$\begin{aligned}
&f_{\rm NL}^{general} = \\
&\frac{5\cdot 2^{-3-n_s}3^{-2-3n_s}}{c_s^2(1+\epsilon)(n_s-2)}\Big(36c_s^2(1+\epsilon)(-27+n_s(13+7n_s))\Gamma[n_s]\sin\left[\frac{n_s\pi}{2}\right] + \\
&16c_s^2(1+\epsilon)^2\Gamma[2+n_s]\sin\left[\frac{n_s\pi}{2}\right] + \big(-9\left(9\epsilon^3(n_s-2)(n_s-1) - 3\epsilon^2(38+5(n_s-7)n_s) + \right. \\
&\left. 8\epsilon\left(21-22n_s+8{n_s}^2\right) - 4(33+n_s(5n_s-31))\right)\Gamma\left[\frac{(3\epsilon-1)(n_s-2)}{1+\epsilon}\right] + \\
&\left. 16(1+\epsilon)^2(f_X-1)\Gamma\left[\frac{4-4\epsilon-n_s+3\epsilon n_s}{1+\epsilon}\right]\right)\sin\left[\frac{(\epsilon(8-3n_s)+n_s)\pi}{2(1+\epsilon)}\right]\Big).
\end{aligned} \tag{D.20}$$

and the functions $\mathcal{C}_i = \mathcal{C}_i(n_s, \epsilon, f_X)$ are given by

$$\mathcal{C}_1(n_s,\epsilon,f_X) \sim \frac{5(-243+n_s(121+67n_s+4\epsilon(1+n_s)))\Gamma[n_s]\sin\left[\frac{n_s\pi}{2}\right]}{2^{1+n_s}3^{2+3n_s}(n_s-2)}, \tag{D.21}$$

$$\begin{aligned}
\mathcal{C}_2(n_s,\epsilon,f_X) \sim -\frac{5\cdot 2^{-3-n_s}3^{-2-3n_s}}{(1+\epsilon)(n_s-2)}&\left(9\left(9\epsilon^3(n_s-2)(n_s-1) - 3\epsilon^2(38+5(n_s-7)n_s) + \right.\right. \\
&\left. 8\epsilon\left(21-22n_s+8{n_s}^2\right) - 4(33+n_s(5n_s-31))\right)\Gamma\left[\frac{(3\epsilon-1)(n_s-2)}{1+\epsilon}\right] - \\
&\left.\left. 16(1+\epsilon)^2(f_X-1)\Gamma\left[\frac{4-4\epsilon-n_s+3\epsilon n_s}{1+\epsilon}\right]\right)\sin\left[\frac{(\epsilon(8-3n_s)+n_s)\pi}{2(1+\epsilon)}\right].
\end{aligned} \tag{D.22}$$

As a result we can also see that the non-Gaussian tilt n_{NG} is given by

$$n_{NG} - 1 = \frac{-2\epsilon_s}{\epsilon_s+\epsilon-1}\frac{\mathcal{C}_2}{\bar{c}_s^2-\mathcal{C}_2}, \tag{D.23}$$

where the functions $\mathcal{C}_i = \mathcal{C}_i(n_s, \epsilon, f_X)$ are as specified above.

Appendix E

Covariant derivatives for derivative chameleons

Here we present some details of the calculation for a generalized equation of motion in conformal derivative chameleon setups, expanding on the results presented in section 5.3. We start with a metric relation

$$\tilde{g}_{\mu\nu} = A^2\left(\phi, X\right) g_{\mu\nu}. \tag{E.1}$$

The covariant derivatives for the two metrics are related by

$$\tilde{\nabla}_\alpha \omega_\beta = \nabla_\alpha \omega_\beta - \Gamma^\gamma_{\alpha\beta}\omega_\gamma. \tag{E.2}$$

For the Einstein frame covariant derivative ∇_α acting on the matter frame stress-energy tensor, one therefore finds

$$\nabla_\alpha \tilde{T}^{\mu\nu} = \tilde{\nabla}_\alpha \tilde{T}^{\mu\nu} - \Gamma^\mu_{\beta\alpha}\tilde{T}^{\beta\nu} - \Gamma^\nu_{\beta\alpha}\tilde{T}^{\mu\beta}. \tag{E.3}$$

In calculating the connection $\Gamma^\mu_{\beta\alpha}$ explicitly, the following relations will be useful

$$\begin{aligned}
\nabla_\alpha A &= A_{,\phi}\partial_\alpha\phi + A_{,X}\partial_\alpha X, \\
&= A_{,\phi}\partial_\alpha\phi - A_{,X}\partial_\alpha\partial_\mu\phi\partial^\mu\phi. \\
\nabla_\alpha A_{,X} &= A_{,X\phi}\partial_\alpha\phi + A_{,XX}\partial_\alpha X, \\
&= A_{,X\phi}\partial_\alpha\phi - A_{,XX}\partial_\alpha\partial_\mu\phi\partial^\mu\phi.
\end{aligned} \tag{E.4}$$

As such we can work out the connection, arriving at

$$\begin{aligned}
\Gamma^{\gamma}_{\alpha\beta} &= \frac{1}{2}\tilde{g}^{\gamma\delta}\left(\nabla_{\alpha}\tilde{g}_{\beta\delta} + \nabla_{\beta}\tilde{g}_{\alpha\delta} - \nabla_{\delta}\tilde{g}_{\beta\alpha}\right) \\
&= A^{-1}\left(2\delta^{\gamma}_{(\beta}\nabla_{\alpha)}A - g_{\alpha\beta}g^{\gamma\delta}\nabla_{\delta}A\right) \\
&= 2A_{,\phi}A^{-1}\delta^{\gamma}_{(\beta}\partial_{\alpha)}\phi - 2A_{,X}A^{-1}\delta^{\gamma}_{(\beta}\partial_{\alpha)}\partial_{\mu}\phi\partial^{\mu}\phi \\
&- g_{\alpha\beta}g^{\gamma\delta}\partial_{\delta}\phi A^{-1}A_{,\phi} + g_{\alpha\beta}g^{\gamma\delta}\partial_{\delta}\partial_{\mu}\phi\partial^{\mu}\phi A^{-1}A_{,X}. \qquad \text{(E.5)}
\end{aligned}$$

Part VI

Bibliography

Bibliography

[1] J. Noller and J. Magueijo, "Non-Gaussianity in single field models without slow-roll," Phys.Rev. **D83**, 103511 [arXiv:1102.0275] (2011).

[2] J. Noller, "Constraining fast-roll inflation," Proceedings for Rencontres de Moriond (Cosmology) [arXiv:1205.5796] (2012).

[3] J. Magueijo, J. Noller and F. Piazza, "Bimetric structure formation: non-Gaussian predictions," Phys.Rev. **D82**, 043521 [arXiv:1006.3216] (2010).

[4] J. Noller, "Derivative Chameleons," JCAP **1207**, 013 [arXiv:1203.6639] (2012).

[5] D. J. Mulryne, J. Noller and N. J. Nunes, "Three-form inflation and non-Gaussianity," [arXiv:1209.2156] (2012).

[6] J. Noller, "Emergent Galileons," in preparation .

[7] J. Noller and J. Magueijo, "Non-adiabatic primordial fluctuations," Class.Quant.Grav. **28**, 105008 [arXiv:0911.1907] (2011).

[8] J. Magueijo and J. Noller, "Primordial fluctuations without scalar fields," Phys.Rev. **D81**, 043509 [arXiv:0907.1772] (2010).

[9] E. Komatsu et al., "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation," Astrophys.J.Suppl. **192**, 18 [arXiv:1001.4538] (2011).

[10] A. H. Guth, "The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems," Phys.Rev. **D23**, 347–356 (1981).

[11] A. D. Linde, "A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems," Phys.Lett. **B108**, 389–393 (1982).

[12] A. Albrecht and P. J. Steinhardt, "Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking," Phys.Rev.Lett. **48**, 1220–1223 (1982).

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

[14] T. Clifton, P. G. Ferreira, A. Padilla and C. Skordis, "Modified Gravity and Cosmology," Phys.Rept. **513**, 1–189 [arXiv:1106.2476] (2012).

[15] J. Khoury and A. Weltman, "Chameleon fields: Awaiting surprises for tests of gravity in space," Phys.Rev.Lett. **93**, 171104 [arXiv:astro-ph/0309300] (2004).

[16] J. Khoury and A. Weltman, "Chameleon cosmology," Phys.Rev. **D69**, 044026 [arXiv:astro-ph/0309411] (2004).

[17] G. Dvali, G. Gabadadze and M. Porrati, "4-D gravity on a brane in 5-D Minkowski space," Phys.Lett. **B485**, 208–214 [arXiv:hep-th/0005016] (2000).

[18] C. de Rham and A. J. Tolley, "DBI and the Galileon reunited," JCAP **1005**, 015 [arXiv:1003.5917] (2010).

[19] C. Armendariz-Picon, T. Damour and V. F. Mukhanov, "k - inflation," Phys.Lett. **B458**, 209–218 [arXiv:hep-th/9904075] (1999).

[20] J. Garriga and V. F. Mukhanov, "Perturbations in k-inflation," Phys.Lett. **B458**, 219–225 [arXiv:hep-th/9904176] (1999).

[21] E. Babichev, V. Mukhanov and A. Vikman, "k-Essence, superluminal propagation, causality and emergent geometry," JHEP **0802**, 101 [arXiv:0708.0561] (2008).

[22] M. Alishahiha, E. Silverstein and D. Tong, "DBI in the sky," Phys.Rev. **D70**, 123505 [arXiv:hep-th/0404084] (2004).

[23] J. Magueijo, "Bimetric varying speed of light theories and primordial fluctuations," Phys.Rev. **D79**, 043525 [arXiv:0807.1689] (2009).

[24] J. D. Bekenstein, "The Relation between physical and gravitational geometry," Phys.Rev. **D48**, 3641–3647 [arXiv:gr-qc/9211017] (1993).

[25] J. Khoury, "Theories of Dark Energy with Screening Mechanisms," [arXiv:1011.5909] (2010).

[26] For reviews of such constraints see e.g. C.M. Will, *Theory and Experiment in Gravitational Physics*, 2nd Ed., (Basic Books/Perseus Group, New York, 1993); E. Fischbach and C. Talmadge, *The Search for Non-Newtonian Gravity*, (Springer-Verlag, New York, 1999); C.M. Will, Living Rev. Rel. 4, 4 (2001) .

[27] T. S. Koivisto and N. J. Nunes, "Three-form cosmology," Phys.Lett. **B685**, 105–109 [arXiv:0907.3883] (2010).

[28] T. S. Koivisto and N. J. Nunes, "Inflation and dark energy from three-forms," Phys.Rev. **D80**, 103509 [arXiv:0908.0920] (2009).

[29] A. Nicolis, R. Rattazzi and E. Trincherini, "The Galileon as a local modification of gravity," Phys.Rev. **D79**, 064036 [arXiv:0811.2197] (2009).

[30] D. Lovelock, "The Einstein tensor and its generalizations," J.Math.Phys. **12**, 498–501 (1971).

[31] D. Lovelock, "The four-dimensionality of space and the einstein tensor," J.Math.Phys. **13**, 874–876 (1972).

[32] M. Ostrogradski, Memoires de l'Academie Imperiale des Science de Saint-Petersbourg, 4:385, 1850 .

[33] G. Horndeski, "Second-order scalar-tensor field equations in a four-dimensional space," International Journal of Theoretical Physics , 10(6):363–384 (1974).

[34] C. Deffayet, X. Gao, D. Steer and G. Zahariade, "From k-essence to generalised Galileons," Phys.Rev. **D84**, 064039 [arXiv:1103.3260] (2011).

[35] M. A. Luty, M. Porrati and R. Rattazzi, "Strong interactions and stability in the DGP model," JHEP **0309**, 029 [arXiv:hep-th/0303116] (2003).

[36] M. Porrati and J. Rombouts, "Strong coupling vs. 4-D locality in induced gravity," Phys.Rev. **D69**, 122003 [arXiv:hep-th/0401211] (2004).

[37] D. Seery and J. E. Lidsey, "Primordial non-Gaussianities in single field inflation," JCAP **0506**, 003 [arXiv:astro-ph/0503692] (2005).

[38] W. H. Kinney, "Inflation: Flow, fixed points and observables to arbitrary order in slow roll," Phys.Rev. **D66**, 083508 [arXiv:astro-ph/0206032] (2002).

[39] N. Agarwal and R. Bean, "Cosmological constraints on general, single field inflation," Phys.Rev. **D79**, 023503 [arXiv:0809.2798] (2009).

[40] H. V. Peiris, D. Baumann, B. Friedman and A. Cooray, "Phenomenology of D-Brane Inflation with General Speed of Sound," Phys.Rev. **D76**, 103517 [arXiv:0706.1240] (2007).

[41] J. M. Bardeen, P. J. Steinhardt and M. S. Turner, "Spontaneous Creation of Almost Scale - Free Density Perturbations in an Inflationary Universe," Phys.Rev. **D28**, 679 (1983).

[42] D. Lyth, "Large Scale Energy Density Perturbations and Inflation," Phys.Rev. **D31**, 1792–1798 (1985).

[43] K. A. Malik and D. Wands, "Gauge invariant variables on cosmological hypersurfaces," [arXiv:gr-qc/9804046] (1998).

[44] K. A. Malik and D. Wands, "Cosmological perturbations," Phys.Rept. **475**, 1–51 [arXiv:0809.4944] (2009).

[45] A. J. Tolley and M. Wyman, "The Gelaton Scenario: Equilateral non-Gaussianity from multi-field dynamics," Phys.Rev. **D81**, 043502 [arXiv:0910.1853] (2010).

[46] J. Moffat, "Superluminary universe: A Possible solution to the initial value problem in cosmology," Int.J.Mod.Phys. **D2**, 351–366 [arXiv:gr-qc/9211020] (1993).

[47] A. Albrecht and J. Magueijo, "A Time varying speed of light as a solution to cosmological puzzles," Phys.Rev. **D59**, 043516 [arXiv:astro-ph/9811018] (1999).

[48] J. Magueijo, "New varying speed of light theories," Rept.Prog.Phys. **66**, 2025 [arXiv:astro-ph/0305457] (2003).

[49] J. Magueijo, "Speedy sound and cosmic structure," Phys.Rev.Lett. **100**, 231302 [arXiv:0803.0859] (2008).

[50] Y.-S. Piao, "Seeding Primordial Perturbations During a Decelerated Expansion," Phys.Rev. **D75**, 063517 [arXiv:gr-qc/0609071] (2007).

[51] B. A. Bassett, S. Liberati, C. Molina-Paris and M. Visser, "Geometrodynamics of variable speed of light cosmologies," Phys.Rev. **D62**, 103518 [arXiv:astro-ph/0001441] (2000).

[52] J. P. Bruneton, theories. gravity," Phys. Rev. D **75**, 085013 (2007) [arXiv:gr-qc/0607055]; motion and closed signal curves," arXiv:0706.1538 [astro-ph]; ARXIV:0706.1538;superluminal propagation, causality and emergent geometry," JHEP **0802**, 101 (2008) [arXiv:0708.0561 [hep-th]] .

[53] S. Weinberg, "Photons and gravitons in perturbation theory: Derivation of Maxwell's and Einstein's equations," Phys.Rev. **138**, B988–B1002 (1965).

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

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

[56] C. Deffayet, G. Dvali, G. Gabadadze and A. I. Vainshtein, "Nonperturbative continuity in graviton mass versus perturbative discontinuity," Phys.Rev. **D65**, 044026 [arXiv:hep-th/0106001] (2002).

[57] K. Hinterbichler and J. Khoury, "Symmetron Fields: Screening Long-Range Forces Through Local Symmetry Restoration," Phys.Rev.Lett. **104**, 231301 [arXiv:1001.4525] (2010).

[58] K. A. Olive and M. Pospelov, "Environmental dependence of masses and coupling constants," Phys.Rev. **D77**, 043524 [arXiv:0709.3825] (2008).

[59] M. Pietroni, "Dark energy condensation," Phys.Rev. **D72**, 043535 [arXiv:astro-ph/0505615] (2005).

[60] P. J. E. Peebles and B. Ratra, Astrophys. J. **325**, L17 (1988); C. Wetterich, Nucl. Phys. B **302**, 668 (1988); M. S. Turner and M. White, Phys. Rev. D **56**, 4439 (1997); R. R. Caldwell, R. Dave and P. J. Steinhardt, Phys. Rev. Lett. **80**, 1582 (1998); I. Zlatev, L. M. Wang and P. J. Steinhardt, Phys. Rev. Lett. **82**, 896 (1999) .

[61] K. Hinterbichler, J. Khoury and H. Nastase, "Towards a UV Completion for Chameleon Scalar Theories," JHEP **1103**, 061 [arXiv:1012.4462] (2011).

[62] C. de Rham, "Galileons in the Sky," [arXiv:1204.5492] (2012).

[63] L. Hui, A. Nicolis and C. Stubbs, "Equivalence Principle Implications of Modified Gravity Models," Phys.Rev. **D80**, 104002 [arXiv:0905.2966] (2009).

[64] C. Armendariz-Picon and R. Penco, "Quantum Equivalence Principle Violations in Scalar-Tensor Theories," Phys.Rev. **D85**, 044052 [arXiv:1108.6028] (2012).

[65] K. Van Acoleyen and J. Van Doorsselaere, "Galileons from Lovelock actions," Phys.Rev. **D83**, 084025 [arXiv:1102.0487] (2011).

[66] J. Khoury and F. Piazza, "Rapidly-Varying Speed of Sound, Scale Invariance and Non-Gaussian Signatures," JCAP **0907**, 026 [arXiv:0811.3633] (2009).

[67] C. Burrage, R. H. Ribeiro and D. Seery, "Large slow-roll corrections to the bispectrum of noncanonical inflation," JCAP **1107**, 032 [arXiv:1103.4126] (2011).

[68] T. Battefeld and J. Grieb, "Anatomy of bispectra in general single-field inflation – modal expansions," JCAP **1112**, 003 [arXiv:1110.1369] (2011).

[69] R. H. Ribeiro, "Inflationary signatures of single-field models beyond slow-roll," JCAP **1205**, 037 [arXiv:1202.4453] (2012).

[70] E. Komatsu, N. Afshordi, N. Bartolo, D. Baumann, J. Bond et al., "Non-Gaussianity as a Probe of the Physics of the Primordial Universe and the Astrophysics of the Low Redshift Universe," [arXiv:0902.4759] (2009).

[71] K. Koyama, "Non-Gaussianity of quantum fields during inflation," Class.Quant.Grav. **27**, 124001 [arXiv:1002.0600] (2010).

[72] E. Komatsu, D. N. Spergel and B. D. Wandelt, "Measuring primordial non-Gaussianity in the cosmic microwave background," Astrophys.J. **634**, 14–19 [arXiv:astro-ph/0305189] (2005).

[73] N. Bartolo, E. Komatsu, S. Matarrese and A. Riotto, "Non-Gaussianity from inflation: Theory and observations," Phys.Rept. **402**, 103–266 [arXiv:astro-ph/0406398] (2004).

[74] E. Komatsu, "Hunting for Primordial Non-Gaussianity in the Cosmic Microwave Background," Class.Quant.Grav. **27**, 124010 [arXiv:1003.6097] (2010).

[75] D. Baumann and D. Green, "Equilateral Non-Gaussianity and New Physics on the Horizon," JCAP **1109**, 014 [arXiv:1102.5343] (2011).

[76] E. D. Stewart and D. H. Lyth, "A More accurate analytic calculation of the spectrum of cosmological perturbations produced during inflation," Phys.Lett. **B302**, 171–175 [arXiv:gr-qc/9302019] (1993).

[77] X. Chen, "Primordial Non-Gaussianities from Inflation Models," Adv.Astron. **2010**, 638979 [arXiv:1002.1416] (2010).

[78] J. Martin and R. H. Brandenberger, "The TransPlanckian problem of inflationary cosmology," Phys.Rev. **D63**, 123501 [arXiv:hep-th/0005209] (2001).

[79] U. H. Danielsson, "A Note on inflation and transPlanckian physics," Phys.Rev. **D66**, 023511 [arXiv:hep-th/0203198] (2002).

[80] X. Chen, M.-x. Huang, S. Kachru and G. Shiu, "Observational signatures and non-Gaussianities of general single field inflation," JCAP **0701**, 002 [arXiv:hep-th/0605045] (2007).

[81] L. Lorenz, J. Martin and C. Ringeval, "Constraints on Kinetically Modified Inflation from WMAP5," Phys.Rev. **D78**, 063543 [arXiv:0807.2414] (2008).

[82] L. Senatore, K. M. Smith and M. Zaldarriaga, "Non-Gaussianities in Single Field Inflation and their Optimal Limits from the WMAP 5-year Data," JCAP **1001**, 028 [arXiv:0905.3746] (2010).

[83] S. Pandolfi, A. Cooray, E. Giusarma, E. W. Kolb, A. Melchiorri et al., "Harrison-Z'eldovich primordial spectrum is consistent with observations," Phys.Rev. **D81**, 123509 [arXiv:1003.4763] (2010).

[84] X. Chen, "Running non-Gaussianities in DBI inflation," Phys.Rev. **D72**, 123518 [arXiv:astro-ph/0507053] (2005).

[85] E. Sefusatti, M. Liguori, A. P. Yadav, M. G. Jackson and E. Pajer, "Constraining Running Non-Gaussianity," JCAP **0912**, 022 [arXiv:0906.0232] (2009).

[86] C. T. Byrnes, M. Gerstenlauer, S. Nurmi, G. Tasinato and D. Wands, "Scale-dependent non-Gaussianity probes inflationary physics," JCAP **1010**, 004 [arXiv:1007.4277] (2010).

[87] R. Easther and H. Peiris, "Implications of a Running Spectral Index for Slow Roll Inflation," JCAP **0609**, 010 [arXiv:astro-ph/0604214] (2006).

[88] C. Burgess, R. Easther, A. Mazumdar, D. F. Mota and T. Multamaki, "Multiple inflation, cosmic string networks and the string landscape," JHEP **0505**, 067 [arXiv:hep-th/0501125] (2005).

[89] J. M. Maldacena, "Non-Gaussian features of primordial fluctuations in single field inflationary models," JHEP **0305**, 013 [arXiv:astro-ph/0210603] (2003).

[90] A. Gruzinov, "Consistency relation for single scalar inflation," Phys.Rev. **D71**, 027301 [arXiv:astro-ph/0406129] (2005).

[91] D. Babich, P. Creminelli and M. Zaldarriaga, "The Shape of non-Gaussianities," JCAP **0408**, 009 [arXiv:astro-ph/0405356] (2004).

[92] J. Fergusson and E. Shellard, "The shape of primordial non-Gaussianity and the CMB bispectrum," Phys.Rev. **D80**, 043510 [arXiv:0812.3413] (2009).

[93] E. Komatsu and D. N. Spergel, "Acoustic signatures in the primary microwave background bispectrum," Phys.Rev. **D63**, 063002 [arXiv:astro-ph/0005036] (2001).

[94] P. D. Meerburg, J. P. van der Schaar and M. G. Jackson, "Bispectrum signatures of a modified vacuum in single field inflation with a small speed of sound," JCAP **1002**, 001 [arXiv:0910.4986] (2010).

[95] P. Creminelli, A. Nicolis, L. Senatore, M. Tegmark and M. Zaldarriaga, "Limits on non-gaussianities from wmap data," JCAP **0605**, 004 [arXiv:astro-ph/0509029] (2006).

[96] X. Chen, R. Easther and E. A. Lim, "Generation and Characterization of Large Non-Gaussianities in Single Field Inflation," JCAP **0804**, 010 [arXiv:0801.3295] (2008).

[97] R. Flauger, L. McAllister, E. Pajer, A. Westphal and G. Xu, "Oscillations in the CMB from Axion Monodromy Inflation," JCAP **1006**, 009 [arXiv:0907.2916] (2010).

[98] S. Hannestad, T. Haugbolle, P. R. Jarnhus and M. S. Sloth, "Non-Gaussianity from Axion Monodromy Inflation," JCAP **1006**, 001 [arXiv:0912.3527] (2010).

[99] P. Creminelli and M. Zaldarriaga, "Single field consistency relation for the 3-point function," JCAP **0410**, 006 [arXiv:astro-ph/0407059] (2004).

[100] D. Baumann, L. Senatore and M. Zaldarriaga, "Scale-Invariance and the Strong Coupling Problem," JCAP **1105**, 004 [arXiv:1101.3320] (2011).

[101] M. LoVerde, A. Miller, S. Shandera and L. Verde, "Effects of Scale-Dependent Non-Gaussianity on Cosmological Structures," JCAP **0804**, 014 [arXiv:0711.4126] (2008).

[102] M. Bruni, R. Crittenden, K. Koyama, R. Maartens, C. Pitrou et al., "Disentangling non-Gaussianity, bias and GR effects in the galaxy distribution," Phys.Rev. **D85**, 041301 [arXiv:1106.3999] (2012).

[103] J. E. Lidsey and I. Huston, "Gravitational wave constraints on Dirac-Born-Infeld inflation," JCAP **0707**, 002 [arXiv:0705.0240] (2007).

[104] R. Bean, D. J. Chung and G. Geshnizjani, "Reconstructing a general inflationary action," Phys.Rev. **D78**, 023517 [arXiv:0801.0742] (2008).

[105] M. Clayton and J. Moffat, "A Scalar - tensor cosmological model with dynamical light velocity," Phys.Lett. **B506**, 177–186 [arXiv:gr-qc/0101126] (2001).

[106] I. T. Drummond, "Bimetric gravity and [dark matter]," Phys.Rev. **D63**, 043503 [arXiv:astro-ph/0008234] (2001).

[107] D. Bessada, W. H. Kinney, D. Stojkovic and J. Wang, "Tachyacoustic Cosmology: An Alternative to Inflation," Phys.Rev. **D81**, 043510 [arXiv:0908.3898] (2010).

[108] E. Silverstein and D. Tong, "Scalar speed limits and cosmology: Acceleration from D-cceleration," Phys.Rev. **D70**, 103505 [arXiv:hep-th/0310221] (2004).

[109] V. F. Mukhanov and A. Vikman, "Enhancing the tensor-to-scalar ratio in simple inflation," JCAP **0602**, 004 [arXiv:astro-ph/0512066] (2006).

[110] N. Afshordi, D. J. Chung, M. Doran and G. Geshnizjani, "Cuscuton Cosmology: Dark Energy meets Modified Gravity," Phys.Rev. **D75**, 123509 [arXiv:astro-ph/0702002] (2007).

[111] N. Afshordi, D. J. Chung and G. Geshnizjani, "Cuscuton: A Causal Field Theory with an Infinite Speed of Sound," Phys.Rev. **D75**, 083513 [arXiv:hep-th/0609150] (2007).

[112] W. H. Kinney and K. Tzirakis, "Quantum modes in DBI inflation: exact solutions and constraints from vacuum selection," Phys.Rev. **D77**, 103517 [arXiv:0712.2043] (2008).

[113] A. Hamma, F. Markopoulou, I. Premont-Schwarz and S. Severini, "Lieb-Robinson bounds and the speed of light from topological order," Phys.Rev.Lett. **102**, 017204 [arXiv:0808.2495] (2009).

[114] J. Magueijo, "DSR as an explanation of cosmological structure," Class.Quant.Grav. **25**, 202002 [arXiv:0807.1854] (2008).

[115] P. Horava, "Quantum Gravity at a Lifshitz Point," Phys.Rev. **D79**, 084008 [arXiv:0901.3775] (2009).

[116] G. Amelino-Camelia, "Doubly special relativity," Nature **418**, 34–35 [arXiv:gr-qc/0207049] (2002).

[117] J. Magueijo and L. Smolin, "Lorentz invariance with an invariant energy scale," Phys.Rev.Lett. **88**, 190403 [arXiv:hep-th/0112090] (2002).

[118] J. Magueijo and L. Smolin, "Generalized Lorentz invariance with an invariant energy scale," Phys.Rev. **D67**, 044017 [arXiv:gr-qc/0207085] (2003).

[119] L. Ford, "Inflation driven by a vector field," Phys.Rev. **D40**, 967 (1989).

[120] T. Koivisto and D. F. Mota, "Vector Field Models of Inflation and Dark Energy," JCAP **0808**, 021 [arXiv:0805.4229] (2008).

[121] A. Golovnev, V. Mukhanov and V. Vanchurin, "Vector Inflation," JCAP **0806**, 009 [arXiv:0802.2068] (2008).

[122] B. Himmetoglu, C. R. Contaldi and M. Peloso, "Instability of anisotropic cosmological solutions supported by vector fields," Phys.Rev.Lett. **102**, 111301 [arXiv:0809.2779] (2009).

[123] M. Karciauskas, K. Dimopoulos and D. H. Lyth, "Anisotropic non-Gaussianity from vector field perturbations," Phys.Rev. **D80**, 023509 [arXiv:0812.0264] (2009).

[124] T. S. Koivisto, D. F. Mota and C. Pitrou, "Inflation from N-Forms and its stability," JHEP **0909**, 092 [arXiv:0903.4158] (2009).

[125] C. Germani and A. Kehagias, "P-nflation: generating cosmic Inflation with p-forms," JCAP **0903**, 028 [arXiv:0902.3667] (2009).

[126] T. Kobayashi and S. Yokoyama, "Gravitational waves from p-form inflation," JCAP **0905**, 004 [arXiv:0903.2769] (2009).

[127] A. De Felice, K. Karwan and P. Wongjun, "Stability of the 3-form field during inflation," [arXiv:1202.0896] (2012).

[128] M. Duff and P. van Nieuwenhuizen, "Quantum Inequivalence of Different Field Representations," Phys.Lett. **B94**, 179 (1980).

[129] X. Chen, M.-x. Huang and G. Shiu, "The Inflationary Trispectrum for Models with Large Non-Gaussianities," Phys.Rev. **D74**, 121301 [arXiv:hep-th/0610235] (2006).

[130] F. Arroja and K. Koyama, "Non-gaussianity from the trispectrum in general single field inflation," Phys.Rev. **D77**, 083517 [arXiv:0802.1167] (2008).

[131] S. Tsujikawa, "Modified gravity models of dark energy," Lect.Notes Phys. **800**, 99–145 [arXiv:1101.0191] (2010).

[132] S. S. Gubser and J. Khoury, "Scalar self-interactions loosen constraints from fifth force searches," Phys.Rev. **D70**, 104001 [arXiv:hep-ph/0405231] (2004).

[133] D. F. Mota and D. J. Shaw, "Evading Equivalence Principle Violations, Cosmological and other Experimental Constraints in Scalar Field Theories with a Strong Coupling to Matter," Phys.Rev. **D75**, 063501 [arXiv:hep-ph/0608078] (2007).

[134] P. Brax, C. van de Bruck and A.-C. Davis, "Compatibility of the chameleon-field model with fifth-force experiments, cosmology, and PVLAS and CAST results," Phys.Rev.Lett. **99**, 121103 [arXiv:hep-ph/0703243] (2007).

[135] P. Brax, C. van de Bruck, A.-C. Davis and D. J. Shaw, "f(R) Gravity and Chameleon Theories," Phys.Rev. **D78**, 104021 [arXiv:0806.3415] (2008).

[136] C. Burrage, A.-C. Davis and D. J. Shaw, "Detecting Chameleons: The Astronomical Polarization Produced by Chameleon-like Scalar Fields," Phys.Rev. **D79**, 044028 [arXiv:0809.1763] (2009).

[137] P. Brax, C. Burrage, A.-C. Davis, D. Seery and A. Weltman, "Collider constraints on interactions of dark energy with the Standard Model," JHEP **0909**, 128 [arXiv:0904.3002] (2009).

[138] A.-C. Davis, C. A. Schelpe and D. J. Shaw, "The Effect of a Chameleon Scalar Field on the Cosmic Microwave Background," Phys.Rev. **D80**, 064016 [arXiv:0907.2672] (2009).

[139] P. Brax, C. Burrage, A.-C. Davis, D. Seery and A. Weltman, "Higgs production as a probe of Chameleon Dark Energy," Phys.Rev. **D81**, 103524 [arXiv:0911.1267] (2010).

[140] P. Brax, C. van de Bruck, A. Davis, D. Shaw and D. Iannuzzi, "Tuning the Mass of Chameleon Fields in Casimir Force Experiments," Phys.Rev.Lett. **104**, 241101 [arXiv:1003.1605] (2010).

[141] R. Gannouji, B. Moraes, D. F. Mota, D. Polarski, S. Tsujikawa et al., "Chameleon dark energy models with characteristic signatures," Phys.Rev. **D82**, 124006 [arXiv:1010.3769] (2010).

[142] P. Brax, C. Burrage, A.-C. Davis, D. Seery and A. Weltman, "Anomalous coupling of scalars to gauge fields," Phys.Lett. **B699**, 5–9 [arXiv:1010.4536] (2011).

[143] Y. Li and W. Hu, "Chameleon Halo Modeling in f(R) Gravity," Phys.Rev. **D84**, 084033 [arXiv:1107.5120] (2011).

[144] D. S. Mak, E. Pierpaoli, F. Schmidt and N. Macellari, "Constraints on Modified Gravity from Sunyaev-Zeldovich Cluster Surveys," Phys.Rev. **D85**, 123513 [arXiv:1111.1004] (2012).

[145] P. Brax, A.-C. Davis, B. Li and H. A. Winther, "A Unified Description of Screened Modified Gravity," [arXiv:1203.4812] (2012).

[146] P. Brax, C. van de Bruck, D. F. Mota, N. J. Nunes and H. A. Winther, "Chameleons with Field Dependent Couplings," Phys.Rev. **D82**, 083503 [arXiv:1006.2796] (2010).

[147] C. Charmousis, E. J. Copeland, A. Padilla and P. M. Saffin, "General second order scalar-tensor theory, self tuning, and the Fab Four," Phys.Rev.Lett. **108**, 051101 [arXiv:1106.2000] (2012).

[148] F. C. Adams, "Stars In Other Universes: Stellar structure with different fundamental constants," JCAP **0808**, 010 [arXiv:0807.3697] (2008).

[149] P. Chang and L. Hui, "Stellar Structure and Tests of Modified Gravity," Astrophys.J. **732**, 25 [arXiv:1011.4107] (2011).

[150] A.-C. Davis, E. A. Lim, J. Sakstein and D. Shaw, "Modified Gravity Makes Galaxies Brighter," [arXiv:1102.5278] (2011).

[151] R. Pourhasan, N. Afshordi, R. Mann and A. Davis, "Chameleon Gravity, Electrostatics, and Kinematics in the Outer Galaxy," JCAP **1112**, 005 [arXiv:1109.0538] (2011).

[152] C. Burrage, C. de Rham, D. Seery and A. J. Tolley, "Galileon inflation," JCAP **1101**, 014 [arXiv:1009.2497] (2011).

[153] A. Silvestri, "Scalar radiation from Chameleon-shielded regions," Phys.Rev.Lett. **106**, 251101 [arXiv:1103.4013] (2011).

[154] T. Tamaki and S. Tsujikawa, "Revisiting chameleon gravity: Thin-shell and no-shell fields with appropriate boundary conditions," Phys.Rev. **D78**, 084028 [arXiv:0808.2284] (2008).

[155] P. Brax, R. Rosenfeld and D. Steer, "Spherical Collapse in Chameleon Models," JCAP **1008**, 033 [arXiv:1005.2051] (2010).

[156] J.-P. Bruneton and G. Esposito-Farese, "Field-theoretical formulations of MOND-like gravity," Phys.Rev. **D76**, 124012 [arXiv:0705.4043] (2007).

[157] T. S. Koivisto, "Disformal quintessence," [arXiv:0811.1957] (2008).

[158] M. Zumalacarregui, T. Koivisto, D. Mota and P. Ruiz-Lapuente, "Disformal Scalar Fields and the Dark Sector of the Universe," JCAP **1005**, 038 [arXiv:1004.2684] (2010).

[159] E. Babichev, C. Deffayet and R. Ziour, "k-Mouflage gravity," Int.J.Mod.Phys. **D18**, 2147–2154 [arXiv:0905.2943] (2009).

[160] C. Deffayet, O. Pujolas, I. Sawicki and A. Vikman, "Imperfect Dark Energy from Kinetic Gravity Braiding," JCAP **1010**, 026 [arXiv:1008.0048] (2010).

[161] T. Kobayashi, M. Yamaguchi and J. Yokoyama, "G-inflation: Inflation driven by the Galileon field," Phys.Rev.Lett. **105**, 231302 [arXiv:1008.0603] (2010).

[162] G. Gubitosi and E. V. Linder, "Purely Kinetic Coupled Gravity," Phys.Lett. **B703**, 113–118 [arXiv:1106.2815] (2011).

[163] D. Fairlie, J. Govaerts and A. Morozov, "Universal field equations with covariant solutions," Nucl.Phys. **B373**, 214–232 [arXiv:hep-th/9110022] (1992).

[164] D. Fairlie and J. Govaerts, "Euler hierarchies and universal equations," J.Math.Phys. **33**, 3543–3566 [arXiv:hep-th/9204074] (1992).

[165] D. Fairlie, "Comments on Galileons," J.Phys.A **A44**, 305201 [arXiv:1102.1594] (2011).

[166] C. Deffayet, G. Esposito-Farese and A. Vikman, "Covariant Galileon," Phys.Rev. **D79**, 084003 [arXiv:0901.1314] (2009).

[167] L. Leblond and S. Shandera, "Simple Bounds from the Perturbative Regime of Inflation," JCAP **0808**, 007 [arXiv:0802.2290] (2008).

[168] E. D. Stewart, "The Spectrum of density perturbations produced during inflation to leading order in a general slow roll approximation," Phys.Rev. **D65**, 103508 [arXiv:astro-ph/0110322] (2002).

[169] R. Durrer and R. Maartens, "Dark Energy and Dark Gravity," Gen.Rel.Grav. **40**, 301–328 [arXiv:0711.0077] (2008).