6-2012

Effective Field Theories in Cosmology

Riccardo Penco

Syracuse University

Abstract

During the last 20 years, a large amount of detailed cosmological observations have promoted cosmology to the rank of high-precision science. Remarkably, all the observations currently available can be accounted for by assuming that (*i*) the universe is approximately homogeneous and isotropic on large scales, (*ii*) gravitational interactions are described by General Relativity with a non-vanishing cosmological constant and (*iii*) 85% of the matter content of the universe is in the form of dark matter, a presently unknown type of matter which interacts with ordinary matter only gravitationally. Current theoretical efforts are focused on gaining a deeper understanding of the small departures from perfect homogeneity and isotropy observed in our universe, the nature of dark matter and the physical origin of the cosmological constant. Effective field theory methods provide a natural framework to try to address such outstanding questions. For instance, such methods have been extensively used to study alternative theories of gravity which mimic a non-vanishing cosmological constant and to build models of the early universe which generate the observed anisotropies and inhomogeneities through a period of accelerated cosmic expansion. In this thesis, we study effective field theories of gravity which violate some basic tenets of General Relativity such as Lorentz invariance and the weak equivalence principle. We also employ effective field theory methods to explore the imprint that high energy physics can leave on the small departures from homogeneity and isotropy generated in the early universe.

Effective Field Theories in Cosmology

by

Riccardo Penco
Laurea Specialistica in Fisica, Università degli Studi di Trieste, 2006
Laurea in Fisica, Università degli Studi di Trieste, 2003

DISSERTATION

Submitted in partial fulfillment of the requirements for the
degree of Doctor of Philosophy in Physics
in the Graduate School of Syracuse University

June 2012

To Paloma.

Thank you for sharing with me

every step of this journey.

Acknowledgements

The work included in this dissertation is the results of six years of work, and I am extremely grateful to all the people who made this possible.

First of all, I wish to thank my advisor, *Cristian Armendariz-Picon*, for his patient guidance and support throughout my Ph.D. studies. He encouraged me to work and think independently and he helped me to broaden my knowledge by getting exposed to many different topics in cosmology and particle physics. He is a great teacher and he always pushed me to present research material in a simpler and more transparent way than I would have otherwise.

I would like to thank *Scott Watson*, first and foremost for making me believe that I could do this. His continuous encouragement and faith in me was a great gift, and I am extremely grateful for all he taught me during the past two years.

I am indebted to *Mark Trodden*, who instilled in me the importance of collaboration, clear communication and academic integrity, and to *Carl Rosenzweig*, who taught me about teaching and encouraged me to be an open-minded researcher. Thank you also to *Ennio Gozzi*, who first exhorted me to pursue a Ph.D. in the United States.

Thank you to all my colleagues in the cosmology group—*Alessandra Silvestri, Eric West, Michele Fontanini, Jayanth Neelakanta, Hardik Panjwani, Richard Galvez* and *Ogan Ozsoy*—for sharing with me the daily life of a graduate student as well as for teaching me a great deal of physics during long and interesting conversations. A special thanks goes to *Mario Martone*, a true friend with a highly contagious passion for physics and life.

Things have not always been easy during these past few years, and I would like to thank all the people that made it possible for me to be with my family when it was most important for me to be there. In particular, I will never forget the compassionate spirit and understanding of *Mark Bowick, Cristina Marchetti* and my advisor *Cristian Armendariz-Picon*.

I also wish to express my gratitude to all the people in the Department of Physics, and especially *Penny Davis, Diane Sanderson* and *Cindy Urtz* for their help and support during these past few years. Life can be hectic, but they always find the time for a kind word or a smile, and for that I am very grateful.

Finally, I would like to thank the most important people in my life. Dear *Mom* and *Dad*, thank you for always believing in me and for encouraging me to fly away from the nest: *vi voglio tanto bene.* Dear Paloma, thank you for flying with me through the good and bad weather, and for making me a better man with your love and support: *je t'aime tellement.*

Syracuse, May 2012

Contents

Chapter 1

Introduction

Einstein's equations of General Relativity (GR), supplemented with the plausible assumption that our planet does not occupy a privileged position in the universe, generically lead to the prediction that the universe is not stationary. This prediction was confirmed in 1929 by Hubble's discovery that the universe is indeed expanding [1]. Such a discovery suggests that the universe must have been much denser in the past, and therefore its energy density much higher. This observation is the foundation of the Hot Big Bang theory.

The main predictions of the Hot Big Bang theory follow essentially from our understanding of particle physics applied to an expanding universe. For instance, the Cosmic Microwave Background (CMB) radiation we observe today was produced when electrons and nuclei combined to form neutral atoms, allowing the photons to propagate freely [2, 3]. This process took place when the energy density ρ of the universe was such that $\rho^{1/4} \sim 10^{-10}$ GeV and was determined by atomic physics. In turn, nuclei were formed when the universe was much denser and the typical energy scale was $\rho^{1/4} \sim 10^{-4}$ GeV. The formation of light nuclei is accurately described by nuclear physics and determines the current abundance of light elements in the universe [4–6].

The idea that the features of the universe we observe today are a reflection of the laws governing particle physics at particular energy scales is very powerful and drives research in modern theoretical cosmology. In particular, it is believed that current observations may offer at least three more windows on particle physics at different energy scales.

The first window is provided by the small temperature fluctuations observed in the CMB [7, 8]. These fluctuations are believed to have the same origin of the large scale structures we observe in the universe [9, 10] and could have been determined during a period of accelerated expansion known as *inflation* which took place at energy scales as high as $\rho^{1/4} \sim 10^{16}$ GeV. Thus, primordial perturbations offer a unique opportunity to test energy scales that would otherwise be too high to be tested with particle accelerators.

The second window follows from precise observations of galaxies [11], galaxy clusters [12–14] and large-scale structures [9, 10] which indicate that the vast majority of matter in the universe is in the form *dark matter*—a presently unknown form of matter which seems to interact with ordinary matter only gravitationally.[1] Given the current abundance of dark matter in the universe, it is plausible [15] that dark matter might be a manifestation of physics taking place at energy scales $\rho^{1/4} \sim 10^3$ GeV. This energy scale is of particular interest because it is currently being investigated at the Large Hadron Collider.

Finally, the last window on particle physics comes from the groundbreaking discovery that the universe is currently undergoing a phase of accelerated expansion [16, 17]. Although this result can be accounted for by a non-vanishing cosmological constant, such an explanation would require a remarkable fine-tuning of parameters, as we will point out in Sec. 2.3.1. For this reason, it was suggested that the phenomenon of cosmic acceleration may be due to a modification of the gravitational sector which becomes relevant at energy scales as low as $\rho^{1/4} \sim 10^{-42}$ GeV [18]. Thus, by studying cosmic acceleration we may be able to probe gravitational interactions at extremely small energy scales.

Given the preponderant role which particle physics plays in our current exploration of the universe, it is not surprising that, over time, techniques that were originally developed in particle physics have been also fruitfully applied in cosmology. A good

[1] CMB observations also provide indirect evidence for the existence of dark matter.

example of such techniques is provided by effective field theories (EFTs). The main advantage of EFTs is that they provide an isolated description of the relevant physics at any given energy scale. As such, EFTs are an ideal tool to study the energy scales we have access to in cosmology. In this thesis, we will make an extensive use EFT methods to study modifications of GR as well as features of the primordial fluctuations generated during inflation.

More specifically, this dissertation is organized as follows. In chapter 2, we introduce some basics EFT concepts which are extensively used in the rest of this thesis. We give a brief overview of the standard cosmological model, some of its limitations and some of the proposals put forward to address them. We focus in particular on possible modifications of GR, and we argue that these alternative theories require either additional degrees of freedom besides the graviton or violations of Lorentz symmetry. We also review the mechanism which leads to the generation of primordial fluctuations during inflation.

In chapter 3, we study theories of gravitation which admit violations of Lorentz symmetry. By extending the coset construction of Callan, Coleman, Wess and Zumino [19], we develop a systematic approach to low-energy theories of gravity which are locally invariant only under a subgroup of the Lorentz group. We illustrate our formalism by considering the explicit case of a theory invariant under local rotations.

In chapter 4, we consider what is arguably the simplest alternative to GR which preserves Lorentz symmetry, namely a gravitational theory with an additional scalar degree of freedom. We show how this class of theories inevitably leads to violations of the weak equivalence principle, which however are likely to be too small to be detectable.

In chapter 5 we examine the imprint that high energy physics might leave on the statistical properties of the small departures from homogeneity and isotropy produced during single-field inflation. We find that high energy physics significantly affects the spectrum of perturbations only when the physical size a fluctuation becomes $\gtrsim M_{\rm p} \times 10^5$. This value is likely to lie well beyond the regime of validity of the EFT, suggesting that for all practical purposes high energy physics would have a negligible impact on the spectrum of primordial perturbations. We conclude in chapter 6 by summarizing the main results of this dissertation and outlining some future directions.

Throughout this dissertation, we made a deliberate use of footnotes to include technical details that can be omitted on a first reading. For completeness, we conclude this introduction with a list of the conventions adopted in this thesis as well as a list of abbreviations.

Appendix 1.A List of Conventions

Throughout this dissertation, we have adopted the following conventions:

- Greek indices $\mu, \nu, \lambda, ...$ label the components of tensors with respect to the coordinate basis and take values $0, 1, 2, 3$.
- Latin indices $a, b, c, ...$ label the components of tensors with respect to an arbitrary orthonormal basis and take values $0, 1, 2, 3$.
- Latin indices $i, j, k, ...$ run over the spatial coordinates and take the values $1, 2, 3$.
- Repeated indices are summed over.
- The metric has a $(-, +, +, +)$ signature.
- The Riemann tensor is defined as $R^{\lambda}{}_{\mu\sigma\nu} \equiv \partial_\sigma \Gamma^\lambda_{\mu\nu} - \partial_\nu \Gamma^\lambda_{\mu\sigma} + \Gamma^\lambda_{\sigma\rho}\Gamma^\rho_{\mu\nu} - \Gamma^\lambda_{\nu\rho}\Gamma^\rho_{\mu\sigma}$.
- The Ricci tensor is defined as $R_{\mu\nu} \equiv R^{\lambda}{}_{\mu\lambda\nu}$.
- The energy momentum tensor $T_{\mu\nu}$ is related to the matter action S_m by

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

- We work in units such that $\hbar = c = k_B = 1$.
- We express energies in GeV, lengths in Mpc and times in yrs. The relation between these units of measure is 1 Mpc $\sim 3 \times 10^6$ yr $\sim 2 \times 10^{38}$ GeV^{-1}.
- We use the reduced Planck mass $M_{\rm p} = (8\pi G)^{-1/2} \approx 2 \times 10^{18}$ GeV.
- We denote cosmic time with t and conformal time with τ. The relation between these two time variables is $d\tau = dt/a(t)$, and I use the abbreviations $\dot{f} \equiv \partial_t f$ and $f' \equiv \partial_\tau f$.

Appendix 1.B List of Abbreviations

Throughout this dissertation, we have used the following abbreviations:

- CMB: Cosmic Microwave Background
- EFT: Effective Field Theory
- FRW: Friedmann-Robertson-Walker
- GR: General Relativity
- PNGB: Pseudo-Nambu-Goldstone Boson
- vev: vacuum expectation value

Chapter 2

Effective Field Theories and Modern Cosmology

2.1 Introduction

A fundamental feature of our world is that it is characterized by a variety of interesting phenomena occurring at different energy scales. At the core of the success of physical sciences is the basic realization that different phenomena can be modeled independently as long as they take place at different energy scales. Effective Field Theories (EFTs) provide a quantitative description of this qualitative statement. Initially developed in the context of particle physics and condensed matter, EFT methods are also particularly suited to the study of cosmology. After all, the universe is the system with the largest possible hierarchy of scales, ranging from the size of our observable patch of universe down to the Planck length.

The goal of this chapter is to provide the reader with an introduction to EFTs and their applications in cosmology. To this end, in Sec. 2.2 we will briefly review the basics of EFTs. Since scalar fields will play an important role in this thesis, the last part of this section will be devoted to EFTs involving scalar fields. In Sec. 2.3, we will review the standard cosmological model which embodies our current understanding of the universe. We will review some of the open problems which are driving current research in theoretical cosmology along with some of the solutions proposed in the literature. One of the possibilities often entertained is that, despite passing stringent observational constraints at solar system scales, General Relativity may fail to provide an accurate description of gravitational interactions at cosmological scales. For this reason, in Sec. 2.4 we will consider possible modifications of General Relativity from an EFT point of view. Many shortcomings of the standard cosmological model can be addressed by assuming that the early universe underwent a phase of quasi-exponential accelerated expansion known as *inflation*. In Sec. 2.5 we will review EFTs of inflation characterized by a single scalar field. We will pay particular attention to the properties of quantum fluctuations during inflation and we will review some of the challenges that need to be overcome in order to build a successful EFT of inflation.

2.2 Effective Field Theories

EFTs are a modern tool to exploit the simplification which arises whenever a system admits a large hierarchy of scales. In this section, we will introduce some basic EFT concepts that will be used in the rest of this thesis. For further details, we refer the reader to the many excellent reviews available in the literature [20–26]. A pedagogical introduction to EFT methods in the context of gravitational theories can also be found in [27, 28].

EFT methods are based on the fundamental assumption that, at any given energy scale, physical phenomena can be described using an *effective action* which provides an "isolated description of the important physics" [20]. The effective action is completely determined once we specify (i) the *symmetries* and (ii) the *field content* of the theory. The latter determines in turn the *particle content* once the theory is quantized. It is actually remarkable that many interesting physical predictions can follow from considerations based exclusively on symmetries and field content.

Any effective action in d space-time dimensions can be schematically written as

$$S = \int d^d x \sum_i g_i \mathcal{O}_i \tag{2.1}$$

where the g_i's are coupling constants and the $\mathcal{O}_i$'s are operators which (i) are invariant under the symmetries of the theory and (*ii*) depend on the fields and their derivatives at a single space-time point. The latter requirement is sufficient to ensure that the *principle of locality* is satisfied, i.e. that distant experiments yield uncorrelated results [29]. If the EFT is weakly coupled, then the typical size of quantum fluctuations is controlled by the free part of the action, which includes the terms that are quadratic in the fields and (usually) contain at most two derivatives. These are the terms that determine the mass dimension of the fields [26].

If a given term in the action has mass dimension $[\mathcal{O}_i] = \delta_i$, then we can introduce the dimensionless coupling $\bar{g}_i = g_i M^{\delta_i - d}$ where M is some mass scale known as *cut-off* (for reasons that will be soon clear) which is chosen in such a way that all $\bar{g}_i \lesssim 1$. We can then use simple dimensional analysis to estimate the contribution of each term in the effective action to a process characterized by an energy scale E larger than any other mass scale in the theory except M:

$$\int d^d x \, g_i \mathcal{O}_i \sim \bar{g}_i \left(\frac{E}{M} \right)^{\delta_i - d} . \tag{2.2}$$

It follows that the operators $\mathcal{O}_i$ can be divided into three distinct categories. The terms that have a mass dimension $\delta_i > d$ give a contribution to the action which becomes less and less important at low energies, and for this reason such terms are called *irrelevant*. The terms with $\delta_i = d$ give a contribution which is equally important at all energies and are known as *marginal*. Finally, the terms with $\delta_i < d$ give a contribution which becomes more and more important at low energies and are therefore called *relevant*. Equivalently, we can say that the low-energy phenomenology of an EFT is encoded in the relevant operators, while the small corrections due to physics at high energies are captured by the irrelevant operators.

Despite the fact that the effective action (2.1) contains an infinite number of terms $\mathcal{O}_i$, only a finite number of them have mass dimension equal or smaller than any given δ_i. This means that, as long as we are interested in carrying out perturbative low-energy calculations with a finite precision, only a finite number of terms and therefore of parameters g_i is needed. This ensures that EFTs are predictive at any given order in the E/M expansion.

Since the dimensionless couplings are all $\bar{g}_i \lesssim 1$, it should be clear from (2.2) that the corrections due to irrelevant operators become all equally important at energies $E \gtrsim M$ and therefore the validity of perturbation theory is cut off (hence the name) at energies of order M. As we will see in chapter 5, a more refined criterion is based on comparing the size of classical (tree-level) correlation functions of the fields with the perturbative corrections coming from quantum (loop) effects.[1] When energies are so large that quantum corrections become of the same order as the classical result, perturbation theory breaks down and analytical calculations become extremely difficult. When this happens, one can introduce a new EFT with a larger cut-off $M' > M$. In order for the new EFT to reproduce the results of the old one at energies $E < M$, the couplings $\bar{g}_i$ must change with the cut-off in a way which is captured by the *renormalization group equations*

$$M \frac{\partial \bar{g}_i}{\partial M} = \beta_i(\bar{g}). \tag{2.3}$$

The functions β_i are usually calculated as a perturbative expansion in the couplings $\bar{g}_i$ [30], although non-perturbative techniques have been developed as well [31].

Since the cut-off determines the largest energy at which perturbative calculations can be trusted, one might be tempted to avoid such limitation by working with an EFT with an infinite cut-off. However, it sometimes happens that the renormalization group equations cannot be integrated beyond a certain limiting scale M_{max}. In this

[1] A slightly technical remark is in order here. Loop corrections are generically divergent and need to be regulated in order to be handled properly. One way of regulating loop corrections consists in cutting off divergent integrals over momenta by introducing an arbitrarily large but finite scale Λ. The distinction between the regulator Λ and the cut-off M becomes blurred in the Wilson approach to EFTs, but these two concepts are independent of each other. This means that we do not need to use Λ as a regulator for an EFT with a cut-off M and we can instead choose to use dimensional regularization, i.e. to work in $d = 4 - \varepsilon$ space-time dimensions. There are indeed several advantages in doing so. For instance, dimensional regularization makes dimensional analysis much easier and allows for an easier determination of the which terms in the effective action should be considered at any given order in the E/M expansion. Moreover, in equation (2.3) we are assuming that the functions β_i do not depend on M, and this happens only if we use a mass-independent regularization procedure such as dimensional regularization.

case, two scenarios are possible:[2]

1. Heavier particles with a mass equal to or smaller than $M_{\max}$ come into play as the cut-off gets closer to $M_{\max}$. It is therefore necessary to include these heavy particles in the EFT if we want to consider a cut-off larger than their mass. These heavier particles will contribute to the renormalization group equations in such a way that the maximum integration scale $M_{\max}$ gets lifted to a larger value, possibly up to $M_{\max} = \infty$.

2. Another possibility is that the fields that appear in the effective action actually describe bound states and that at high energies these are no longer the correct degrees of freedom to use. An example of this scenario is provided by QCD: at low enough energies physical processes can be described using an effective action which involves only pions. However, as the energy increases the coupling between pions becomes stronger and stronger and eventually one needs to switch to QCD—a new EFT in which quarks are the appropriate (weakly coupled) degrees of freedom.

A sufficient condition[3] to ensure that the renormalization group equations (2.3) can be integrated up to $M = \infty$ is if the couplings $\bar{g}_i$ approach a fixed poin $\bar{g}^*$ such that $\beta(\bar{g}^*) = 0$ in the limit $M \to \infty$. A theory which satisfies this condition is called *asymptotically safe* [33] and is said to be "fundamental" because it gives a self-consistent description of physics at arbitrarily high energies. QCD is an example of a fundamental theory since it is not only asymptotically safe, but actually asymptotically *free* because the fixed point is at $\bar{g}^* = 0$.

---

[2]It is worth mentioning a third and definitely more speculative scenario known as *classicalization* [32]. In this scenario, scattering processes with an energy $E \gtrsim M_{\max}$ are still under control and they lead to the formation of classical field configurations.

[3]The rationale behind this condition can be intuitively understood by integrating the renormalization group equations to get

$$\int_{\bar{g}_i(M)}^{\bar{g}_i(M')} \frac{d\bar{g}_i}{\beta_i(\bar{g})} = \log \frac{M'}{M}.$$

If the couplings $\bar{g}_i$ approach a fixed point in the limit $M' \to \infty$, then both the LHS and the RHS of this equation will diverge.

An important concept to introduce is that of *naturalness* [34], which requires that the dimensionless coefficients $\bar{g}_i$ cannot be much smaller than one unless setting them to zero would increase the symmetry of the EFT. The expectation of naturalness comes from the fact that loop corrections will generically yield an order one contribution to $\bar{g}_i$ even if we start with $\bar{g}_i = 0$ at tree-level. We will explore this extensively in Chapter 4. If however the EFT acquires an additional symmetry when $\bar{g}_i = 0$, loop corrections to $\bar{g}_i$ which preserve such symmetry must vanish in the limit $\bar{g}_i \to 0$ and therefore must be proportional to $\bar{g}_i$ itself. Hence, if the tree-level value of $\bar{g}_i$ is much smaller than one, loop corrections are guaranteed to yield a negligible correction. This happens for instance in gauge theories, where gauge couplings can be much smaller than one because setting them to zero would turn the gauge bosons into free particles whose number is separately conserved [34]. Similarly, fermion masses can be much smaller than the cut-off because massless fermions enjoy an additional chiral symmetry $\psi \to e^{i\alpha\gamma^5}\psi$. Finally, it is presently believed that scalar fields can have a mass much smaller than the cut-off only in the presence of (*i*) spontaneously broken continuous symmetries or (*ii*) supersymmetry. Since we will make an extensive use of scalar fields in the rest of this dissertation, it is appropriate to conclude this section by briefly reviewing both these scenarios.

2.2.1 Pseudo-Nambu-Goldstone Bosons

Symmetries are spontaneously broken when the vacuum of the theory is not as symmetric as the theory itself. In the case of spontaneously broken *continuous* symmetries, Goldstone's theorem ensures the existence of one massless scalar particle—a Goldstone boson—for every symmetry broken by the vacuum.[4] If the theory does not have other massless particles, it is possible to consider a low-energy EFT which involves only the Goldstone bosons.

The prototypical example of such an EFT is a theory in which the global symmetry $SU(N)$ is spontaneously broken down to $SU(N-1)$ at the energy scale f. Such a

[4] This statement applies only to *internal* symmetries. In the case of *space-time* symmetries, the number of Goldstone bosons can also be smaller than the number of broken symmetries [35].

theory contains $2N-1$ Goldstone bosons π^a which transform according to a non-linear representation of $SU(N)$ [36]. The Goldstone bosons can be conveniently grouped as follows

$$\Sigma(x) = e^{i\pi_a(x)t^a/f}. \tag{2.4}$$

where the t^a's are the generators of the broken symmetries. This field parametrization is particularly convenient because it has very simple transformation properties under $SU(N)$, namely $\Sigma \to U\Sigma U^\dagger$ where $U \in SU(N)$. Then, the most generic low-energy effective action for the Goldstone bosons which is invariant under $SU(N)$ is

$$S = \int d^4x \, \frac{f^2}{4} \, \mathrm{Tr}\left[(\partial_\mu \Sigma)^\dagger \, (\partial^\mu \Sigma)\right] + \cdots , \tag{2.5}$$

$$= \int d^4x \left\{ \frac{1}{2}(\partial\pi)^2 + \frac{1}{24f^2}\left[(\partial\pi^2)^2 - 4\pi^2(\partial\pi)^2\right] + \cdots \right\} \tag{2.6}$$

where the dots in (2.5) stand for terms with higher derivatives which become negligible in the low-energy limit $E \ll f$.

The action (2.6) clearly shows that the Goldstone bosons are massless because there are no quadratic terms without derivatives. In fact, a small mass term for the π_a's would explicitly break $SU(N)$ and therefore it would be natural because setting it to zero would restore the $SU(N)$ symmetry. A Goldstone boson with a small mass term is known as *pseudo-Nambu-Goldstone boson* (PNGB).

2.2.2 Supersymmetry

In supersymmetric theories, each fermonic field has a bosonic counterpart with the same mass. In particular, any spin 1/2 field is associated with a complex scalar field. A small mass for these scalars is then natural because the mass of the associated fermions is protected by chiral symmetry as mentioned above. There is however a price to pay: the constraints imposed by supersymmetry are so stringent that, even in the presence of gravitational interactions, the potential for the scalar fields must take the form [37]

$$V = e^{K/M_\mathrm{p}^2}\left[\left(\frac{\partial W}{\partial\varphi^n} + \frac{W}{M_\mathrm{p}^2}\frac{\partial K}{\partial\varphi^n}\right)^\dagger \left(\frac{\partial^2 K}{\partial\varphi_n^\dagger \partial\varphi_m}\right)^{-1} \left(\frac{\partial W}{\partial\varphi^m} + \frac{W}{M_\mathrm{p}^2}\frac{\partial K}{\partial\varphi^m}\right) - 3\frac{|W|^2}{M_\mathrm{p}^2}\right] \tag{2.7}$$

where W depends only on φ_n (not on $\varphi_n^\dagger$) and $K = \varphi_n^\dagger \varphi^n + \mathcal{O}(1/M_{\rm p}^2)$. As we will discuss in Sec. 2.5.3, this constraint will turn out to be a major obstacle toward the construction of successful models of inflation based on supersymmetry.

2.3 The Standard Cosmological Model

Our current understanding of the Universe is based on the assumption that gravitational interactions at cosmological scales are well described by Einstein's equations of General Relativity (GR):

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{T_{\mu\nu}}{M_{\rm p}^2} + \Lambda g_{\mu\nu}. \tag{2.8}$$

Moreover, precise observations of the Cosmic Microwave Background (CMB) and the Large Scale Structures respectively indicate that the universe appears to be isotropic and homogeneous at scales $\gtrsim$ 100 Mpc. For this reason, it is plausible that the large-scale behavior of space-time can be approximately described by the Friedmann-Robertson-Walker (FRW) metric

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = -dt^2 + a(t)^2 \left[\frac{dr^2}{1 - kr^2} + r^2 \left(d\theta^2 + \sin^2\theta \, d\varphi^2 \right) \right], \tag{2.9}$$

which is the most general metric invariant under spatial translations and rotations. The index $k = 0, \pm 1$ determines the intrinsic curvature of the hypersurfaces of constant time, while the scale factor $a(t)$ determines how the proper distance between two free-falling observers changes over time. Isotropy and homogeneity also constrain the energy-momentum tensor $T_{\mu\nu}$ appearing on the right hand side of equation (2.8), which must take the form $T^\mu{}_\nu = \mathrm{diag}\{-\rho(t), p(t), p(t), p(t)\}$ where ρ is the total energy density and p is the total pressure. Therefore, under the assumptions of homogeneity and isotropy Einstein's equations (2.8) reduce to the following system of differential equations:

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{\rho}{3M_{\rm p}^2} + \frac{\Lambda}{3} \tag{2.10a}$$

$$\frac{\ddot{a}}{a} = -\frac{\rho + 3p}{6M_{\rm p}^2} + \frac{\Lambda}{3} \tag{2.10b}$$

These equations can also be rewritten as

$$\Omega_m + \Omega_\Lambda + \Omega_k = 1 \tag{2.11a}$$

$$q(1-\Omega_m) = -\left(\Omega_\Lambda + \frac{\dot{\Omega}_m}{2H}\right) \tag{2.11b}$$

where we have defined the Hubble parameter $H = \dot{a}/a$, the deceleration parameter $q = -\ddot{a}/(aH^2)$ and the fractional densities of matter, $\Omega_m = \rho/(3M_{\rm p}^2 H^2)$, cosmological constant, $\Omega_\Lambda = \Lambda/(3H^2)$ and spatial curvature $\Omega_k = -k/(3a^2H^2)$. In order to cast Einstein's equations in the form (2.11) we have also used the conservation of the energy-momentum tensor, $\nabla_\mu T^{\mu\nu} = 0$. It is useful to break Ω_m into three distinct components,

$$\Omega_m = \Omega_r + \Omega_b + \Omega_{dm}, \tag{2.12}$$

by introducing the fractional densities of relativistic species, Ω_r, of non-relativistic baryons, Ω_b, and of dark matter Ω_{dm}. Relativistic species include known particles such as photon and neutrinos, but could also include unknown relativistic particles. Baryons include all known particles which are too heavy to be relativistic. Finally, dark matter consists of non-relativistic particles which haven't been discovered yet and seem to interact with baryons and relativistic species only gravitationally.

One of the most remarkable achievements of modern observational cosmology is the measurement of the current values of the Hubble parameter, the deceleration parameter and all the fractional densities appearing in equations (2.11) and (2.12). The results of these measurements can be summarized as follows:

$$H_0 \approx 2\times 10^{-42}\,\text{GeV} \tag{2.13a}$$

$$q_0 \approx -0.6 \tag{2.13b}$$

$$\Omega_{\Lambda,0} \approx 0.73 \tag{2.13c}$$

$$\Omega_{dm,0} \approx 0.23 \tag{2.13d}$$

$$\Omega_{b,0} \approx 0.04 \tag{2.13e}$$

$$\Omega_{r,0} \approx 10^{-5} \tag{2.13f}$$

$$\Omega_{k,0} \lesssim 10^{-2}. \tag{2.13g}$$

The values of these parameters reveal a surprising picture of our universe. First of all, a negative value of the deceleration parameter implies that the universe is currently undergoing a phase of accelerated expansion. This defies the naïve expectation that the gravitational attraction among the universe's constituents should slow down the cosmic expansion. Secondly, equations (2.13e) and (2.13f) indicate that all the particles that have been discovered so far account for just a small fraction of the universe's energy budget: the expansion of the universe seems to be essentially determined by non-baryonic matter and the cosmological constant. Finally, notice that the standard cosmological model is characterized by a single scale, the Hubble parameter H_0, which determines both the current age as well the size of our observable universe:

$$H_0^{-1} \approx 1.4 \times 10^{10} \text{ yr} \approx 4 \times 10^3 \text{ Mpc}. \tag{2.14}$$

Despite the fact that a remarkable array of observations is consistently explained by the values (2.13), the standard cosmological model is unable to fully account for the properties of the constituents described by the fractional densities (2.13c) – (2.13g). For instance, the nature of dark matter is still presently unknown, although many dark matter candidates have been proposed in the literature. Even the baryonic sector is somewhat puzzling, as we still lack a convincing explanation for the excess of matter over anti-matter. In the next section, we will focus on some of the open problems that are more relevant for the work presented in this dissertation.

2.3.1 Some Open Problems

Ω_Λ: Cosmological Constant Problem

The results (2.13) indicate that almost 75% of the energy density of the universe is accounted for by a non-vanishing cosmological constant. If we assume that GR is an EFT valid up to a certain cut-off M, then arguments based on naturalness lead to the expectation that $M_\mathrm{p}^2\Lambda \sim M^4$ [38]. If we assume that GR remains valid up to the Planck scale, then we should have $\Lambda \sim M_\mathrm{p}^2$, in stark contrast with the observational result $\Omega_{\Lambda,0} \sim 1$ which implies

$$\Lambda \sim H_0^2 \sim M_{\rm p}^2 \times 10^{-120}. \tag{2.15}$$

Remarkably, a cosmological constant of this order of magnitude was first predicted by Weinberg based on anthropic considerations [39]. The huge discrepancy between the expected and observed values of Λ is known as the *cosmological constant problem.* It is important to stress that this problem is not a consequence of the (maybe overly) optimistic assumption that GR is valid up to energies $M \sim M_{\rm p}$. In fact, even a cut-off of the order of the mass of the electron $M \sim m_e$ would lead to a cosmological constant which is too large by more than thirty orders of magnitude!

Ω_k: Flatness Problem

By differentiating the Friedmann equation (2.11a) w.r.t. the scale factor and using the conservation of the energy momentum tensor we obtain a differential equation describing how the fractional density Ω_k changes with the scale factor:

$$\frac{d\Omega_k}{da} = \left(1 + \frac{3p}{\rho}\right) \frac{\Omega_k(1-\Omega_k)}{a}. \tag{2.16}$$

This differential equation admits three fixed points, namely $\Omega_k = 0, 1, \infty$. The stability properties of these fixed points are determined by the sign of the coefficient $1 + 3p/\rho$ and are summarized in Table 2.1 in the case of an expanding universe. Since the universe expansion was dominated by relativistic ($p = \rho/3$) or non-relativistic ($p = 0$) matter during the past 14 billion years, one would expect today's value of Ω_k to be much closer to 1 than the observed value (2.13g). Equivalently, observations seem to require an extremely fine-tuned and therefore unnatural initial value of Ω_k: this puzzle is known as the *flatness problem.*

Ω_r: Horizon Problem

As we mentioned in the previous section, the most compelling evidence for the isotropy of the universe at large scales comes from the fact that the CMB temperature is

$1+3p/\rho$	$\Omega_k = 0$	$\Omega_k = 1$	$\Omega_k = \infty$
> 0	repeller	attractor	repeller
< 0	attractor	repeller	attractor

Table 2.1: Stability properties of the fixed points of equation (2.16) for an expanding universe.

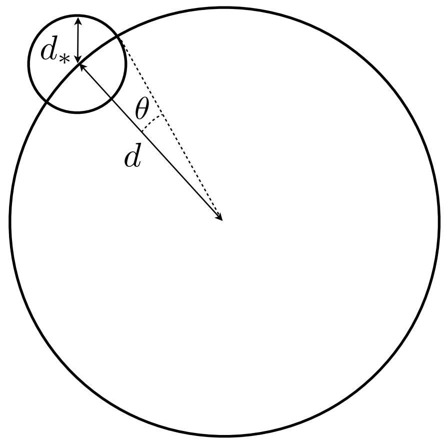


Figure 2.1: Particle horizon on the surface of last scattering.

uniform across the sky up to a part in 10^{-5}. This remarkable uniformity seems however to be at odds with the notion of causality in an expanding universe.

The characteristic scale of an FRW metric is the Hubble parameter H. Its inverse roughly determines the maximum proper distance that particles can travel during the time it takes for the universe to double in size. Regions of space-time that are separated by a proper distance greater than H^{-1} are therefore not in causal contact because the space in between is expanding faster than the particles can move. Equivalently, two regions are not in causal contact if they are separated by a comoving distance larger than the *comoving Hubble radius* $(aH)^{-1}$.

The CMB radiation decoupled from the baryon plasma when the universe was $t_d \sim 3 \times 10^5$ yr old, and at that time the comoving size of the causal regions was $(a_d H_d)^{-1}$. Taking into account that most of the expansion occurred when the universe was dominated by radiation, we can estimate the proper size of these regions today to be

$$d_* = \frac{a_0}{a_d H_d} \sim t_d \left(\frac{t_0}{t_d}\right)^{1/2} \sim 10 \text{ Mpc}. \tag{2.17}$$

Since decoupling, CMB photons have been traveling a much larger proper distance, roughly equal to $d = H_0^{-1} \sim 4 \times 10^3$ Mpc. Therefore, CMB photons can be thought of as originating from a sphere of proper radius d, as shown in Fig. 2.1. The solid angle of sky subtended by a causally connected region is approximately $\theta^2 \approx (d_*/d)^2 \sim 10^{-4}$, and therefore the CMB should consist of roughly $4\pi/\theta^2 \sim 10^5$ causally independent patches. How is it possible that 10^5 causally independent regions give rise to CMB radiation with almost exactly the same temperature? This puzzle is known as the *horizon problem.*

2.3.2 Possible Solutions

In what follows, I will briefly survey some possible solutions to the outstanding problems mentioned in the previous section.

Ω_Λ: Modified Gravity

As we pointed out in the previous section, the cosmological constant problem is essentially a naturalness problem. The smallness of the ratio $\Lambda/M_{\rm p}^2$ would not be a problem by itself if the value of the cosmological constant was protected by some symmetry. Unfortunately, the only symmetry which is presently known to protect the cosmological constant is supersymmetry, which however must be broken at energies $\gtrsim 10^3$ GeV. Thus, supersymmetry could at most account for a cosmological constant of the order of $(10^3 \text{ GeV})^4/M_{\rm p}^2 \sim M_{\rm p}^2 \times 10^{-62}$ that is still much larger than the observed value (2.15).

While it is true that any model which explains cosmic acceleration without resorting to a cosmological constant must still contain a small parameter (after all, the small ratio $\Lambda/M_{\rm p}^2$ must come from somewhere), the goal is to have a small parameter which is natural. In fact, the smallness of the ratio $m_e/M_{\rm p}$ between the mass of the

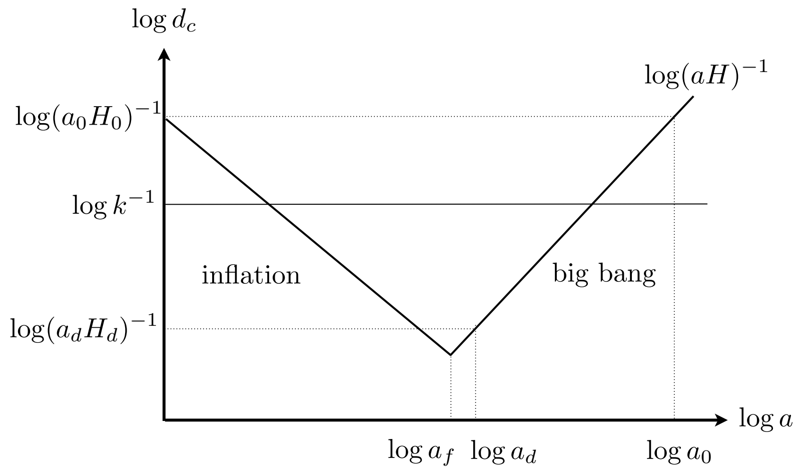


Figure 2.2: Evolution of the comoving Hubble radius during inflation and the Big Bang phase. The values of the scale factor at the end of inflation, at decoupling and today are denoted respectively with a_f, a_d and a_0. Similarly, H_d and H_0 stand for the Hubble parameter at decoupling at today. Finally, k represents the comoving wavenumber of the fluctuations we observe today.

electron and the Planck mass does not constitute a problem precisely because a small fermion mass is natural.

Any alternative explanation of cosmic acceleration necessarily requires a modification of Einstein's equations (2.8). A subtle distinction is often drawn between models that introduce additional fields besides the metric (*Dark Energy models*) and models that do not (*Modified gravity models*). However, we find this distinction artificial given that all these models can be thought of as modifications of the gravitational sector. For this reason, in what follows we will call any alternative to GR a modified gravity model. For a more comprehensive review of different approaches to cosmic acceleration, we refer the reader to [40].

Ω_r, Ω_k: Inflation or Contraction

Based on equation (2.16), a dynamical solution to the flatness problem requires a long enough period in the universe's history during which $\Omega_k = 0$ is a dynamical attractor. According to Table 2.1, this can happen only if (i) the universe is dominated by a source with $\rho + 3p < 0$, or (ii) the universe undergoes a period of contraction.

In the first case, Eq. (2.10b) implies that the universe must go through a phase of accelerated expansion known as *inflation*. During this period, the comoving Hubble radius $(aH)^{-1}$ decreases in time, since

$$\frac{d}{dt}(aH)^{-1} = -\frac{\ddot{a}}{(aH)^2} < 0. \tag{2.18}$$

This result is the main reason why inflation can also address the horizon problem. As illustrated in Figure 2.2, regions that were too large to be causally connected at the time of decoupling might have been smaller than the comoving Hubble radius during inflation provided the phase of accelerated expansion lasts long enough. Inflation is therefore able to solve the horizon problem if the comoving Hubble radius at the beginning and at the end of inflation are such that

$$N \equiv \log\frac{a_f H_f}{a_i H_i} \gtrsim \log\frac{a_d H_d}{a_0 H_0} \sim \log\frac{t_0}{t_d} \sim 10 \tag{2.19}$$

where we used the fact that $t_d \sim 3 \times 10^5$ yr and $t_0 \sim \times 10^{10}$ yr. This condition also ensures the resolution of the flatness problem [41].

As a bonus, inflation provides a causal mechanism to generate the primordial density fluctuations necessary to explain structure formation and CMB anisotropies [42–46]. As shown in Figure 2.2, inflation allows for the comoving wavelength of the fluctuations we observe today to be smaller than the comoving Hubble radius in the early universe. This means that physical process which took place during inflation within a causally connected region determined the properties of the fluctuations we observe today in our universe.

As we will see in section 2.5, the simplest models of inflation require just a single scalar degree of freedom. If we demand that this scalar has a luminal or sub-luminal speed of propagation, then it can be shown [47–50] that inflation is the only mechanism which (*i*) remains weakly coupled, (*ii*) is a dynamical attractor and (*iii*) gives rise to primordial perturbations in agreement with observations.

2.4 Effective Theories of Gravity

Einstein's equations (2.8) can be derived by varying the Einstein-Hilbert action:

$$S = \int \sqrt{-g} \left[\frac{M_{\rm p}^2}{2} (R - \Lambda) \right] + S_m[g_{\mu\nu}, \chi] \qquad (2.20)$$

where χ stands for an arbitrary number of matter fields. Given the observational evidence that $\Lambda \ll M_{\rm p}$, in what follows we will set $\Lambda = 0$. Then, the Minkowski metric $\eta_{\mu\nu}$ is the only vacuum solution of Einstein's equations. If we consider fluctuations around the Minkowski background, i.e. $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}/M_{\rm p}$, the Einstein-Hilbert action (2.20) reduces to

$$S = \int \left\{ (\partial h)^2 + \frac{h(\partial h)^2}{M_{\rm p}} + \frac{h^2(\partial h)^2}{M_{\rm p}^2} + \cdots + \frac{hT}{M_{\rm p}} + \cdots \right\}. \qquad (2.21)$$

where for simplicity we have suppressed all the indices.

At the classical level, the action (2.21) clearly shows that GR is a highly non-linear theory of a rank-2 tensor $h_{\mu\nu}$. In the linear approximation, the equations of motion for $h_{\mu\nu}$ can be written schematically as $\partial^2 h \sim T/M_{\rm p}$. Such approximation breaks down when the cubic and higher-order terms in Eq. (2.21) become as large as the quadratic term. Thus, classical non-linearities become important when $h \sim M_{\rm p}$. As an example, let us consider a static point-like source of mass m: the energy-momentum tensor is schematically $T \sim m\,\delta(\mathbf{r})$ and the linear equations can be solved using dimensional analysis arguments[5] to get $h \sim \frac{m}{M_{\rm p} r}$. Hence, classical non-linearities become important when $h \sim M_{\rm p}$ or, equivalently, at distances of the order of the Schwarzschild radius $r \sim m/M_{\rm p}^2$

From the point of view of quantum field theory, elementary excitations of rank-j tensors behave as particles with integer spin j. Therefore, the action (2.21) describes an effective theory for a self-interacting massless spin-2 particle. The Einstein-Hilbert action (2.20) contains all terms invariant under diffeomorphisms with at most two derivatives. However, based on the arguments of Sec. 2.2 we expect quantum effects to introduce an infinite number of higher-derivative terms, such as for instance

[5]We are neglecting for the moment the issue of gauge invariance, as it will not play any role in what follows.

$R^2, R_{\mu\nu}R^{\mu\nu}, \cdots$. These terms will in turn generate corrections to the action (2.21) of the form $(\partial^2 h)^2/M_\mathrm{p}^2$, which cannot be neglected when $\partial \sim M_\mathrm{p}$ or, equivalently, at distances[6] $r \sim 1/M_\mathrm{p}$. GR is therefore an EFT with a cut-off of the order of M_p.

It turns out that GR is actually the only consistent Lorentz invariant low-energy EFT of a self-interacting massless spin-2 particle [51]. Thus, any modification of gravity necessarily requires either (*i*) a modification of the particle content in the gravitational sector or (*ii*) violations of Lorentz invariance. In the rest of this section we will briefly review both scenarios.

2.4.1 Modifying the Particle Content

Any modification of GR which preserves Lorentz symmetry necessarily requires additional particles which couple to the energy-momentum tensor. Moreover, if we are interested in modifications which explain cosmic acceleration, then the additional particles need to mediate a force with a range which is at least $r \sim 1/H_0$. Such a long-range interaction can only be mediated by bosons, because fermions cannot form classical coherent states [52].[7] Moreover, massive bosons give rise to an interaction with a range of the order of the Compton wavelength of the particle, i.e. $r \sim 1/m$. For this reason, the additional bosons should be either massless or have an (effective) mass $m \lesssim H_0$.

As we mentioned earlier, bosons of integer spin j are associated with rank-j tensors. Any Lorentz-invariant coupling between the energy-momentum tensor $T^{\mu\nu}$ and a rank-1 or rank-j with $j \geqslant 3$ necessarily requires some derivatives [54] in order to contract all the indices in a Lorentz-invariant fashion. Therefore, such couplings will

[6]Incidentally, it is interesting to notice that $1/M_\mathrm{p}$ is typically much smaller than the length scale m/M_p^2 at which classical non-linear effects become relevant around point-like sources. The wide separation between these two scales is what ultimately justifies going beyond the linear approximation in classical contexts.

[7]Fermions can give rise to a composite boson by forming a condensate, but such boson would mediate a force which decays as $1/r^6$ at large distances [53] and is therefore negligible compared to the force mediated by the graviton which decays as $1/r^2$.

become negligible at large distances compared to the graviton coupling in Eq. (2.21).[8] It is also known that a low-energy EFT which contains more than one massless spin-2 particle is consistent only if such particles are not coupled to each other [56–58]. Hence, adding one or more massless spin-2 particles to the gravitational sector would result in a trivial modification of GR.[9] For this reason, the only Lorentz-invariant modifications of gravity that will be relevant to our purposes require the addition of one of more spin-0 particles to the gravitational sector.

Alternative theories of gravitation with additional scalar degrees of freedom have a very long history and it is impossible to do justice to all the models that have been proposed in the literature. In many models, the additional scalar degrees of freedom are not immediately apparent at the level of the action but they arise, for instance, from higher derivative terms (e.g. [18]) or additional spatial dimensions (e.g. [61]). For an extensive review of modified gravity models we refer the reader to [62].

The prototypical model is described by the action

$$S = \int \sqrt{-g}\left\{\frac{M_{\rm p}^2}{2}R - \frac{1}{2}\partial_\mu\varphi\,\partial^\mu\varphi - V(\varphi)\right\} + S_m[F^2(\varphi)g_{\mu\nu}, \chi], \tag{2.22}$$

and is a generalization of the theory originally proposed by Brans and Dicke [63]. Since we are interested in understanding the role played by the additional scalar, we will now consider the artificial limit $M_{\rm p} \to \infty$ in which the metric fluctuations decouple from the rest of the fields and $g_{\mu\nu}$ reduces to the Minkowski metric.

Because of the interaction between matter fields and φ, the stress-energy tensor of matter is not conserved. However, in the non-relativistic limit where the pressure of matter is negligible compared to its energy density, it is possible to define a conserved quantity $\tilde{\rho}$ which is related to the matter energy density ρ by $\tilde{\rho} = \rho/F(\varphi)$. Then, by varying the action (2.22) w.r.t. φ we see that the effective potential for φ is

$$V_{\rm eff}(\varphi) = V(\varphi) + F(\varphi)\tilde{\rho}. \tag{2.23}$$

[8]For massless particles with spin $j \geqslant 3$, there is also the additional problem that there are no self-interactions that can be written [55].

[9]See however [59] for an exception to this result when the spin 2 particles have a mass and [60] for a concrete example.

Although the justification for considering a modified gravity model such as (2.22) comes from the acceleration taking place at cosmological scales $\sim 10^3$ Mpc, one still needs to preserve agreement with current experimental constraints obtained at solar system distances $\lesssim 10^{-12}$ Mpc. The gravitational field outside a static source of mass m can be generically parametrized using the PPN expansion [64]

$$ds^2 \approx -\left[1 - \frac{m}{4\pi M_{\rm p}^2 r} + \frac{\beta}{2}\left(\frac{m}{4\pi M_{\rm p}^2 r}\right)^2\right] dt^2 + \left(1 + \gamma\frac{m}{4\pi M_{\rm p}^2 r}\right) dx_i dx^i, \qquad (2.24)$$

where $\gamma = \beta = 1$ in GR. Current experimental limits on these parameters are [64]

$$|\gamma - 1| \lesssim 10^{-5}, \qquad\qquad |\beta - 1| \lesssim 10^{-4} \qquad (2.25)$$

Agreement with these constraints is usually achieved by screening the effect of the scalar degree of freedom at small scales using one of the following mechanisms [65]:

- *Chameleon mechanism* [66, 67]: in this scenario the effective mass of the scalar mode is proportional to the local matter density. Then, the scalar mode becomes heavy close to the Earth surface and the effective range of the interaction becomes too short to affect experimental tests. A toy model which exhibits this behavior is provided by the action (2.22) with $V(\varphi) = \mu^{4+n}/\varphi^n$ and $F(\varphi) = 1 + \beta\varphi/M$. In this case, the effective potential (2.23) has a minimum at $\varphi_{\rm min} \sim \tilde{\rho}^{-1/(n+1)}$ and the effective mass of the fluctuations around this minimum is $m^2_{\rm min} \sim \tilde{\rho}^{(n+2)/(n+1)}$.

- *Symmetron mechanism* [68–70]: in this scenario the linear coupling between the scalar mode and the matter fields is proportional to the vacuum expectation value (vev) of the scalar field itself. The vev depends on the local density and it vanishes for large enough densities, leading to an extremely weak interaction. A toy model with this behavior is provided again by the action (2.22) where $V(\phi) = -\mu^2\varphi^2 + \lambda\varphi^4$ and $F(\varphi) = 1 + \varphi^2/M^2$. In this case, the effective potential (2.23) is such that the scalar field acquires a non-vanishing vev only when $\tilde{\rho} < \mu^2 M^2$. When the local density is larger than this critical value, the vev of the scalar field vanishes and so does the linear coupling between φ and the matter fields.

- *Vainshtein mechanism* [71]: in this scenario higher-derivative non-linear terms effectively increase the kinetic term of φ, leading to a small coupling between the canonically normalized scalar field and matter. A prototypical example of this mechanism can be obtained by replacing the potential $V(\varphi)$ with a derivative self-interaction of the form $\Box\varphi(\partial\varphi)^2/M^3$ in the action (2.22) and assuming that $F^2(\varphi) = 1 + \varphi/M_\mathrm{p}$. When $\Box\varphi/M^3 \gg 1$ then the non-linear interaction term becomes more important than the kinetic term. In the region surrounding a static point-like mass m, this happens for $r \lesssim 1/\Lambda$ with $\Lambda = (m/M_\mathrm{p})^{1/3}/M$, because $\varphi \sim m/(M_\mathrm{p} r)$.[10] Then, the canonically normalized scalar fluctuations around this background are $\delta\varphi_c \sim (\Lambda/r)^{3/2}\delta\varphi$ and the linear coupling with matter becomes very small: $(r/\Lambda)^{3/2}\delta\varphi_c T/M_\mathrm{p} \ll \delta\varphi_c T/M_\mathrm{p}$.

All these mechanisms can be used to ensure that the bounds (2.25) are satisfied and therefore that the motion of a *single* point-like particle outside a static, spherically symmetric mass m is indistinguishable from the one predicted by GR. However, much more stringent constraints follow from comparing the motion of *two* extended objects with different composition. In the zero-tide approximation, GR predicts that such objects will behave like point-like particles and will experience the same gravitational acceleration. This result is known as the *weak equivalence principle*. The current experimental limit on violations of the weak equivalence principle is [64]

$$\left|\frac{a_1 - a_2}{a_1 + a_2}\right| \sim 10^{-13}, \tag{2.26}$$

where a_1 and a_2 stand for the gravitational acceleration of two test masses.

In the case of the model (2.22), the form of the coupling between φ and matter is such that the motion of point-like particles will preserve the weak equivalence principle at the classical level. However, it is well known that extended objects with a gravitational binding energy comparable to the rest mass can grossly violate the equivalence principle: this phenomenon is known as the *Nordtvedt effect* [72]. Moreover, it was recently shown that models that employ the chameleon screening

[10] Notice that quantum corrections are still negligible at these scales because they become important only for $r \lesssim 1/M$.

mechanism can lead to substantial violations of the weak equivalence principle even with a very small gravitational binding energy [73].

Notice that the coupling between φ and matter in (2.22) was chosen *ad hoc* in order to preserve the weak equivalence principle for point-like particles at the classical level. In fact, all point-like particles will fall along the geodesics of the effective metric $\tilde{g}_{\mu\nu} = F^2(\varphi) g_{\mu\nu}$ and will experience the same gravitational acceleration. However, the structure of this coupling is not protected by any symmetry and therefore, from an EFT point of view, we expect quantum corrections to introduce couplings that violate the weak equivalence principle. We will explore this issue extensively in Chapter 4.

2.4.2 Modifying the Symmetries

In the previous section we considered modified theories of gravity which preserve the symmetries of GR. We will now pursue the alternative approach of modifying GR by breaking some of its symmetries. By doing so, we will often allow additional metric degrees of freedom to propagate and therefore modify the particle content of the gravitational sector as a byproduct.

In order to make all the symmetries of GR explicit, it is convenient to adopt the vierbein formalism and write the metric as follows:

$$g_{\mu\nu} = e_\mu{}^a e_\nu{}^b \eta_{ab}. \tag{2.27}$$

An extensive review of the vierbein formalism is provided for completeness in Appendix 2.A. In terms of the vierbein, it becomes clear that the symmetries of GR are diffeomorphisms, under which the vierbein transforms as

$$e_\mu{}^a(x) \to e'_\mu{}^a(x') = \frac{\partial x^\nu}{\partial x'^\mu} e_\nu{}^a(x), \tag{2.28}$$

and local Lorentz symmetry, which acts on the vierbein as follows

$$e_\mu{}^a(x) \to e'_\mu{}^a(x) = \Lambda^a{}_b(x)\, e_\mu{}^b(x). \tag{2.29}$$

On a Minkowski background, the vierbein acquire a vev $e_\mu{}^a = \delta^a_\mu$. This vev is responsible for the symmetry breaking pattern

$$\text{Diffeomorphisms} \;\times\; \text{Local Lorentz} \quad\longrightarrow\quad \text{Global Lorentz.}$$

Hence, we see that any modification of gravity which breaks either diffeomorphisms or Local Lorentz symmetry will necessarily lead to violations of global Lorentz symmetry on a Minkowski background.

Models that break diffeomorphisms have a very different phenomenology than the models which break local Lorentz symmetry, even when they preserve the same symmetries on a Minkowski background. A well studied example of a model which breaks diffeomorphisms while preserving rotations on a Minkowski background is provided by the ghost condensate [74]. In this model, a scalar field acquires a time dependent vev which breaks time diffeomorphisms without affecting spatial diffeomorphisms and local Lorentz symmetry. This leads to a gravitational sector with *three* propagating degrees of freedom corresponding to the two polarizations of the graviton and an additional massless scalar.

An example of an EFT of gravity which breaks local Lorentz symmetry while preserving rotations on a Minkowski background is provided by the Einstein-Aether theory [75]. In this theory, a vector field acquires a time-like vev which breaks local Lorentz symmetry down to local rotations. On a Minkowski background there are *five* propagating degrees of freedom and they are all massless: the two polarizations of the graviton, two additional vector modes and one scalar.

Although models which preserve rotational symmetry make it simpler to study cosmological solutions, we can also consider more generic patterns of symmetry breaking. This can be especially interesting given that experimental bounds on Lorentz breaking in the gravitational sector are much weaker than in the Standard Model sector [76]. For this reason, in chapter 3 we will explore the low-energy phenomenology of modified theories of gravity in which Lorentz symmetry is broken down to any of it subgroups.

2.5 Effective Theories of Inflation

As we mentioned in Section **??**, the flatness and horizon problems of the standard cosmological model can be solved if we assume that a phase of accelerated expansion—

inflation—took place in the early universe. In the simplest model of inflation, the energy-momentum budget of the universe is dominated by a single scalar field [77–79] described by the action

$$S = \sqrt{-g}\left\{\frac{M_{\rm p}^2}{2}R - \frac{1}{2}\partial_\mu\varphi\,\partial^\mu\varphi - V(\varphi) + \cdots\right\} \tag{2.30}$$

This is the most generic low-energy effective action that we can write down (up to field redefinitions) with at most two derivatives.[11] On A FRW background, the energy density and pressure of the scalar field are

$$\rho = \frac{\dot{\varphi}}{2} + V(\varphi), \qquad\qquad p = \frac{\dot{\varphi}}{2} - V(\varphi). \tag{2.31}$$

Therefore, the condition $\rho + 3p < 0$ necessary for inflation is satisfied if the potential energy of the scalar field is larger than its kinetic energy. The horizon and flatness problems are then solved if inflation lasts long enough that the requirement (2.19) is satisfied. As we will see in section 2.5.3, building an effective theory that produces a sufficient amount of inflation turns out to be quite challenging from a theoretical point of view.

At the background level, the isotropy and flatness of the universe are the only two observable consequences of inflation. In order to distinguish between different models of inflation it is therefore necessary to study the behavior of perturbations around the perfectly homogeneous and isotropic FRW background. It is convenient to classify the metric perturbations according to their transformation properties with respect to the background symmetries, i.e. spatial rotations and translations. This amounts to writing the perturbed line element as

$$\begin{aligned} ds^2 = &-(1+2\phi)dt^2 + 2a(\partial_i B - S_i)dx^i dt \\ &+ a^2\left[(1-2\psi)\delta_{ij} + 2\partial_i\partial_j E + 2\partial_{(i}F_{j)} + h_{ij}\right]dx^i dx^j \ . \end{aligned} \tag{2.32}$$

where $\partial^i S_i = \partial^i F_i = 0, \partial^i h_{ij} = 0$ and $\delta^{ij}h_{ij} = 0$. We will also decompose the scalar field into its background value $\bar{\varphi}$ and a perturbation $\delta\varphi$, i.e. $\varphi = \bar{\varphi} + \delta\varphi$. Thus, the spectrum of perturbations includes five scalars $(\phi, B, \psi, E, \delta\varphi)$, two transverse vectors

[11] The effect of higher-derivative irrelevant operators was considered in [80].

(S_i, F_i) and one traceless and transverse tensor (h_{ij}). The advantage of carrying out such a decomposition is that, at the linear level, perturbations which transform according to different representations of the background symmetries decouple and therefore can be studied separately. It turns out that vector perturbations are not sourced in the simple model of inflation we are considering. For this reason, in what follows we will focus our attention on scalar and tensor perturbations.

2.5.1 Scalar Perturbations

The analysis of scalar perturbations is complicated by the fact that not all five scalar perturbations ϕ, B, ψ, E and $\delta\varphi$ are physical. This follows from the fact that, under infinitesimal coordinate transformations $x^\mu \to x^\mu + \xi^\mu$, perturbations of a rank-n tensor $A_{\mu_1\cdots\mu_n}$ around its background value $\bar{A}_{\mu_1\cdots\mu_n}$ transform as

$$\delta A_{\mu_1\cdots\mu_n} \to \delta A_{\mu_1\cdots\mu_n} - \xi^\nu \partial_\nu \bar{A}_{\mu_1\cdots\mu_n} - \bar{A}_{\nu\cdots\mu_n}\partial_{\mu_1}\xi^\nu - \cdots - \bar{A}_{\mu_1\cdots\nu}\partial_{\mu_n}\xi^\nu. \tag{2.33}$$

As a consequence, if we decompose the infinitesimal quantity ξ^μ as $\xi^\mu = (\alpha, \partial^i\beta + \gamma^i)$ with $\partial_i\gamma^i = 0$, we find that scalar perturbations of the metric and the inflaton transform as follows:

$$\phi \to \phi - \dot{\alpha} \tag{2.34a}$$

$$B \to B + \frac{\alpha}{a} - a\dot{\beta} \tag{2.34b}$$

$$E \to E - \beta \tag{2.34c}$$

$$\psi \to \psi + H\alpha \tag{2.34d}$$

$$\delta\varphi \to \delta\varphi - \dot{\bar{\varphi}}\,\alpha. \tag{2.34e}$$

Individual scalars are therefore not invariant under coordinate transformations, but it is still possible to give a coordinate-independent description of scalar perturbations by considering some linear combinations that remain invariant under (2.34). An example of such a linear combination is the is the *comoving curvature perturbation* [81]:

$$\mathcal{R} = \psi - \frac{H}{\bar{\rho} + \bar{p}}\,\delta q \tag{2.35}$$

where δq is the scalar part of the $(0,i)$ component of the stress-energy tensor, i.e. $T^0{}_i = \partial_i \delta q$. During inflation $\delta q = -\dot{\bar{\varphi}}\,\delta\varphi$ and $\bar{\rho}+\bar{p} = \dot{\bar{\varphi}}^2$, and $\mathcal{R}$ measures the spatial curvature of the hyper-surfaces of constant φ.

A very important property of $\mathcal{R}$ is that its Fourier modes with $k/(aH) \ll 1$ do not evolve in time [81].[12] Since $(aH)^{-1}$ decreases during inflation, modes with a comoving wavelength smaller than $(aH)^{-1}$ at the beginning of inflation will evolve until $k \sim aH$ and then remain constant until their wavelength becomes again larger than $(aH)^{-1}$ during the big bang phase (see Figure 2.2). This means that the initial conditions for today's curvature perturbations with wavenumber k were essentially determined during inflation at the time t_* such that $k \sim a(t_*)H(t_*)$, which in common parlance is usually referred to as *horizon-crossing* time. As we will see, these initial conditions are a direct consequence of the quantum fluctuations occurring during inflation.

It turns out that the scalar sector of single-field inflationary model is completely characterized by the comoving curvature perturbation $\mathcal{R}$. This is because the transformations (2.34) can always be exploited to set two scalar perturbations to zero, and Einstein's equations impose two additional constraints on the scalar sector leaving us with a single propagating degree of freedom. The free action for $\mathcal{R}$ can be obtained by inserting the perturbed metric and scalar field into (2.30) and expanding up to quadratic order, which yields

$$S^{(2)}_{\mathcal{R}} = \int dt d^3x \, \frac{az^2}{2}\left[\dot{\mathcal{R}}^2 - \frac{(\partial_i \mathcal{R})^2}{a^2}\right]. \tag{2.36}$$

where we have introduced the quantity $z \equiv a\dot{\bar{\varphi}}/H$. Since the action (2.30) is invariant under diffeomorphisms, it is not surprising that the action for perturbations can be expressed solely in terms of the invariant quantity $\mathcal{R}$. The action (2.36) can be put into canonical form by switching to conformal time $d\tau = dt/a(t)$ and introducing the *Mukhanov-Sasaki variable* $v \equiv z\mathcal{R}$. This leads to the following action for v [82]:

$$S^{(2)}_{v} = \frac{1}{2}\int d\tau d^3x \left[(v')^2 - (\partial_i v)^2 + \frac{z''}{z}v^2\right], \tag{2.37}$$

[12] Strictly speaking, this is true as long as matter perturbations are adiabatic. See the end of this section for a definition of adiabaticity as well as a discussion about the current observational evidence supporting this assumption.

which is essentially the action for a free scalar field with a time-dependent effective mass $-z''/z$. We can now study quantum fluctuations of the canonically normalized field v using standard quantum field theory methods [83]. For instance, we can decompose the operator $\hat{v}$ into a superposition of creation and annihilation operators,

$$\hat{v}(\tau, \mathbf{r}) = \int \frac{d^3k}{(2\pi)^3} \left[\hat{a}_{\mathbf{k}} v_{\mathbf{k}}(\tau) e^{i\mathbf{k}\cdot\mathbf{r}} + \hat{a}^{\dagger}_{\mathbf{k}} v^{*}_{\mathbf{k}}(\tau) e^{-i\mathbf{k}\cdot\mathbf{r}} \right], \tag{2.38}$$

with $\hat{a}_{\mathbf{k}}$ and $\hat{a}^{\dagger}_{\mathbf{k}'}$ satisfying the algebra $[\hat{a}_{\mathbf{k}}, \hat{a}^{\dagger}_{\mathbf{k}'}] = (2\pi)^3\delta(\mathbf{k} - \mathbf{k}')$. The vacuum is defined as usual by the relation $\hat{a}_{\mathbf{k}}|0\rangle = 0$. The mode functions $v_{\mathbf{k}}(\tau)$ are univocally determined by the requirement that they reduce to the usual mode functions on a Minkowski background in the limit $k^2 \gg |z''/z|$:

$$\lim_{|\mathbf{k}|\to\infty} v_{\mathbf{k}}(\tau) \approx \frac{e^{-ik\tau}}{\sqrt{2k}}. \tag{2.39}$$

The rationale behind this requirement is that the cosmological expansion should not play any role at very small scales. The variance of comoving curvature perturbations is then related to the variance of $v_{\mathbf{k}}(\tau)$ by the following equation:

$$\langle \mathcal{R}_{\mathbf{k}} \mathcal{R}_{\mathbf{k}'} \rangle = \frac{\langle v_{\mathbf{k}} v_{\mathbf{k}'} \rangle}{z^2} = (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}') \frac{|v_{\mathbf{k}}(\tau)|^2}{z^2} \equiv (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}') P_{\mathcal{R}}(\tau, k), \tag{2.40}$$

where $\langle ... \rangle$ denotes a vacuum expectation value and the delta function follows from the fact that the background is invariant under spatial translations. Since the modes $\mathcal{R}_{\mathbf{k}}$ become constant after they cross the horizon, the initial variance of a mode with wavenumber k which “re-enters” the horizon in the late universe is equal to $P_{\mathcal{R}}(\tau_*, k)$. Equivalently, we can describe the variance of curvature perturbations using the *power spectrum*[13]

$$\Delta^2_{\mathcal{R}}(k) \equiv \frac{k^3}{2\pi} P_{\mathcal{R}}(\tau_*, k) \approx A_s \left(\frac{k}{k_0} \right)^{n_s - 1}. \tag{2.41}$$

Notice that the horizon-crossing time τ_* can be expressed in terms of k by inverting the relation $k = a(\tau_*)H(\tau_*)$. The RHS of equation (2.41) is a phenomenological parametrization used to compare theoretical predictions with CMB and large scale structures observations. In particular, current CMB observations [7] show that, for

[13] The normalization of the power spectrum is chosen in such a way that the real space variance is simply $\langle \mathcal{R}\mathcal{R} \rangle = \int_0^\infty \Delta^2_{\mathcal{R}}(k) d\log k$.

$k_0 = 2\times 10^{-3}$ Mpc^{-1}, the *amplitude* A_s of the power spectrum and the *spectral index* n_s are

$$A_s = 2\times 10^{-9}, \qquad n_s = 0.97. \tag{2.42}$$

This means that initial curvature perturbations are very small ($\mathcal{R}\sim 10^{-5}$) and almost scale-invariant, in the sense that $\Delta^2_{\mathcal{R}}(k)$ has a very mild dependence on the wavenumber k.

It is important to realize that an early phase of accelerated expansion will not automatically give rise to an almost scale-invariant spectrum of curvature perturbations. However, a sufficient condition for scale invariance is that the effective mass of v be approximately $z''/z \approx 2/\tau^2$. In this case, the initial conditions (2.39) imply that the wave modes are

$$v_{\mathbf{k}}(\tau) \approx \frac{e^{-ik\tau}}{\sqrt{2k}}\left(1-\frac{i}{k\tau}\right). \tag{2.43}$$

Since in the long-distance limit $k\to 0$ we have $\mathcal{R}_{\mathbf{k}} = v_{\mathbf{k}}/z \sim 1/(\tau z) \equiv$ const, we must have that $z = a\dot{\varphi}/H \sim 1/\tau$. It was shown in [47] that the only accelerating solution which satisfies this condition is an *almost* exponential expansion with $a\sim e^{Ht}\sim 1/(H\tau)$, $H\approx$ const and $\dot{\varphi}\approx$ const. The deviation from a perfect exponential expansion leads in turn to a deviation from perfect scale invariance and it remains small provided the following conditions are met:[14]

$$-\frac{\dot{H}}{H^2}\equiv \varepsilon \ll 1, \qquad \frac{\dot{\varepsilon}}{\varepsilon H}\equiv 2(\varepsilon-\eta)\ll 1. \tag{2.44}$$

Loosely speaking, the first requirement ensures that the Hubble parameter remains approximately constant during the time it takes for the universe to double in size, while the second requirement ensures that the phase of almost exponential expansion lasts "long enough". We can use Friedmann's equations (2.10) together with the conservation of energy-momentum and equations (2.31) to rewrite the conditions (2.44) as

$$\varepsilon = \frac{\dot{\varphi}^2}{2M_{\mathrm{p}}^2H^2}\sim M_{\mathrm{p}}^2\left(\frac{V_{,\varphi}}{V}\right)^2 \ll 1, \qquad \eta = -\frac{\ddot{\varphi}}{H\dot{\varphi}}\sim M_{\mathrm{p}}^2\left(\frac{V_{,\varphi\varphi}}{V}\right)\ll 1. \tag{2.45}$$

[14] More precisely, one should also require that $\dot{\eta}/(\eta H)\equiv 2(\eta-\xi)\ll 1$ and so on, but it turns out that these additional conditions are usually satisfied if those in Eq.(2.44) are satisfied.

Therefore, the potential V needs to be very flat (in Planck units) in order to generate an almost exponential expansion.[15] Moreover, if we identify $V_{,\varphi\varphi}$ with the effective mass of the scalar field, the requirement $\eta \ll 1$ is equivalent to demanding that this mass be much smaller than the Hubble parameter during inflation, because

$$V_{,\varphi\varphi} \ll V/M_{\rm p}^2 \approx \rho/M_{\rm p}^2 \sim H^2, \qquad (2.46)$$

where we used once again Friedmann's equations. An intuitive explanation of this result is that a non-negligible mass would introduce a characteristic scale in the power spectrum, thus spoiling scale invariance.

In the quasi-exponential limit (2.44), we can use the mode functions (2.43) to calculate explicitly the amplitude of the power spectrum and the spectral index. To leading order in ε and η, we get

$$A_s = \Delta_{\mathcal{R}}^2(k_0) = \frac{1}{8\pi^2}\frac{H_*^2}{M_{\rm p}^2}\frac{1}{\varepsilon_*} \qquad (2.47a)$$

$$n_s - 1 = \frac{d\log\Delta_{\mathcal{R}}^2}{d\log k} \approx \left(2\frac{d\log H_*}{dt_*} - \frac{d\log\varepsilon_*}{dt_*}\right)\frac{dt_*}{d\log k} \approx 2\eta_* - 4\varepsilon_*, \qquad (2.47b)$$

where we used the fact that $d\log k/dt_* \approx H_*$ and the subscript "$_*$" quantities evaluated at horizon-crossing. It is quite remarkable that the results (2.47) are valid for any potential V which satisfies the conditions (2.45).

To conclude this long section on scalar perturbations, let us mention two more properties of curvature perturbations. The first one is that curvature perturbations are gaussian to a very good approximation. Deviations from gaussianity are usually characterized by the phenomenological parameter f_{NL} which, for historical reasons, is defined by a relation with a schematic form $\langle\mathcal{R}^3\rangle \sim f_{NL}\langle\mathcal{R}^2\rangle^2$. The magnitude of this parameter can be estimated by studying the cubic corrections to the free action (2.36). Current observations [7] indicate that $|f_{NL}| \lesssim 10^2$, which corresponds to non-gaussianites with a relative size $|f_{NL}\mathcal{R}| < 10^{-3}$. This bound is more than satisfied by the simple model (2.30), which predicts $f_{NL} \sim \varepsilon_*^2 \ll 1$ [85].

[15]The simple model (2.30) could be generalized by considering a non-canonical kinetic term [84]. In this case, inflation can occur also with a very steep potential. In this scenario, the field v can have a speed of sound smaller than the speed of light.

A level of non-gaussianity much closer to the current observational bound can be produced in models with more than one scalar field (see e.g. [86]). However, these models usually generate *non-adiabatic* perturbations which lead to fluctuations in the ratio n_{dm}/n_γ between dark matter and photon number densities. The ratio between the power spectrums of $\mathcal{S} \equiv \delta\left(n_{dm}/n_\gamma\right)$ and of curvature perturbations is usually parametrized as

$$\frac{P_{\mathcal{S}}(k_0)}{P_{\mathcal{R}}(k_0)} \equiv \frac{\alpha(k_0)}{1-\alpha(k_0)} \tag{2.48}$$

and current CMB observations [7] lead to the bound $\alpha(k_0) < 0.1$ for $k_0 = 0.002\,\mathrm{Mpc}^{-1}$. Therefore, non-adiabatic modes, if present, must be fairly small compared to curvature perturbations, and this places a constraint on models with multiple scalars.

In summary, current observations indicate that primordial scalar perturbations are (i) scale-invariant, (ii) gaussian and (iii) adiabatic to a very good approximation. Models of inflation involving a single scalar field are able to account for such properties provided they lead to a quasi-exponential accelerated expansion.

2.5.2 Tensor perturbations

After an extensive analysis of scalar sector, we now move on to the study of tensor perturbations generated during inflation. This analysis turns out to be considerably simpler than the one carried out in the previous section, in part because tensor perturbations are invariant under coordinate transformations, but also because we can now borrow most of the formalism introduced above.

We can obtain the free action for tensor perturbations by inserting the perturbed metric into the action (2.30) and expanding up to quadratic order in h_{ij}. This procedure yields

$$S_h^{(2)} = \int dt d^3x\, \frac{M_{\mathrm{p}}^2 a^3}{8}\left[(\dot{h}_{ij})^2 - \frac{(\partial_k h_{ij})^2}{a^2}\right]. \tag{2.49}$$

Because of the constraints $\delta^{ij}h_{ij} = 0$ and $\partial^i h_{ij} = 0$, h_{ij} contains two degrees of freedom. It is therefore useful to express the Fourier modes of h_{ij} as a sum of two independent polarizations,

$$h_{ij}(t,\mathbf{r}) = \sum_{s=+,\times} \int \frac{d^3k}{(2\pi)^3} e^s_{ij}(k) h^s_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{2.50}$$

where $\delta^{ij} e^s_{ij} = 0$ and $k^i e^s_{ij} = 0$. We can now define the Fourier components of two canonically normalized field (one for each polarization) as $v^s_{\mathbf{k}} \equiv a M_{\rm p} h^s_{\mathbf{k}}/2$ and rewrite the action (2.49) as

$$S^{(2)}_v = \sum_{s=+,\times} \frac{1}{2} \int d\tau d^3x \left[(v'_s)^2 - (\partial_i v_s)^2 + \frac{a''}{a} v_s^2 \right]. \tag{2.51}$$

During a phase of almost exponential expansion the scale factor grows as $a \sim 1/(H\tau)$ and therefore $a''/a \approx 2/\tau^2$. From our previous discussion, we can already conclude that the spectrum of tensor perturbations will also be almost scale-invariant. We can check that explicitly by calculating the variance of tensor modes

$$\langle h^s_{\mathbf{k}} h^{s'}_{\mathbf{k}'} \rangle = \frac{4\langle v^s_{\mathbf{k}} v^{s'}_{\mathbf{k}'} \rangle}{a^2 M_{\rm p}^2} = (2\pi)^3 \delta(\mathbf{k}+\mathbf{k}') \delta^{ss'} \frac{4|v^s_{\mathbf{k}}(\tau)|^2}{a^2 M_{\rm p}^2} \equiv (2\pi)^3 \delta(\mathbf{k}+\mathbf{k}') \delta^{ss'} P_h(\tau, k) \tag{2.52}$$

where the modes $v^s_{\mathbf{k}}(\tau)$ are again approximately given by Eq. (2.43). Taking into account both polarizations, the power spectrum of primordial tensor perturbations is defined as follows:

$$\Delta^2_h(k) \equiv 2 \times \frac{k^3}{2\pi} P_h(\tau_*, k) \equiv A_t \left(\frac{k}{k_0} \right)^{n_t}. \tag{2.53}$$

Notice that, for historical reasons, the spectral index of tensor perturbations n_t is defined differently than the one for scalar perturbations. The simple model (2.30) therefore predicts the following values for the spectral index and the amplitude of tensor perturbations

$$A_t = \Delta^2_h(k_0) = \frac{2}{\pi^2} \frac{H_*^2}{M_{\rm p}^2} \tag{2.54}$$

$$n_t = \frac{d\log \Delta^2_h}{d\log k} = 2\, \frac{d\log H_*}{dt_*} \frac{dt_*}{d\log k} \approx -2\varepsilon_*. \tag{2.55}$$

Observations [7] constrain the ratio $r \equiv A_t/A_s$ between the amplitudes of tensor and scalar modes to be $r \lesssim 0.2$. Since the amplitude of scalar perturbations is known, this result can be easily turned into an upper bound on the energy scale of inflation. In fact, by using the fact that $H^2 \sim V/M_{\rm p}^2$ during inflation we obtain

$$V^{1/4} \sim \sqrt{M_{\rm p} H_*} \sim \left(\frac{r}{0.01}\right)^{1/4} \times 10^{16}\ \text{GeV}. \tag{2.56}$$

It is interesting to notice that the ratio $r/0.01$ also controls the change in φ occurring between the time when CMB scales cross the horizon and the end of inflation [87]:

$$\frac{\Delta\varphi}{M_{\rm p}} = \int_{t_{cmb}}^{t_f} dt\, \frac{\dot{\varphi}}{M_{\rm p}} \approx \sqrt{2\varepsilon_*} \log \frac{a_f}{a_{\rm cmb}} \sim \left(\frac{r}{0.01}\right)^{1/2}, \tag{2.57}$$

Therefore, in the context of the single-field model (2.30), a detection of tensor perturbations with a scalar-tensor ratio $r > 0.01$ would suggest that inflation occurred at energies close to the GUT scale $M_{GUT} \sim 10^{16}$ GeV and that the inflation value changed by a large amount in Planck units. This observation would have important implications from a theoretical point of view. In fact, in order for inflation to last despite a large change in the value of φ, the action for φ must be relatively insensitive to the actual value of φ, i.e. it must be approximately invariant under shift symmetry $\varphi \to \varphi + \text{const}$. For more details about the connection between field range and shift symmetry, we refer the reader to [88].

2.5.3 Open Problems

In the simple model of inflation discussed above we have $\dot{\varphi} \sim \sqrt{\varepsilon} H M_{\rm p}$ (see Eq. (2.45)). This means that effective corrections to the action (2.30) must be suppressed by a cut-off that cannot be smaller than $M \sim \sqrt{\varepsilon} M_{\rm p}$, otherwise higher-derivative terms with arbitrary powers of φ/M would dominate over the relevant operators shown in (2.30). Since we are interested in calculating the properties of fluctuations at horizon crossing, the energy scale of interest is $\sim H$. Hence, it follows from Eqs. (2.42) and (2.47a) that the expansion parameter for this EFT is $H/M \lesssim 6 \times 10^{-5}$ regardless of the value of the slow-roll parameter ε [80].

From an EFT point of view, one of the major obstacles on the path to a successful model of inflation stems from the requirement (2.45) that the mass of the inflation be sufficiently small compared to the Hubble scale during inflation. This requirement is much more stringent than the requirement that the mass of the scalar be smaller than the cut-off M, because the Hubble parameter H is already much smaller than M.

In section 2.1, we pointed out that the only two ways to have a scalar field with a mass naturally smaller than the cut-off involve either approximate global symmetries which are spontaneously broken or supersymmetry. In the first case, the inflaton is a PNGB which, in the simplest possible model [89, 90], is described by the Lagrangian

$$\mathcal{L}_\phi = -\frac{1}{2}\partial_\mu\varphi\,\partial^\mu\varphi - V_0\left[1-\cos(\varphi/f)\right]. \tag{2.58}$$

The problem with this approach is that the whole potential and not only the mass term is proportional to the small breaking scale $V_0^{1/4}$. Therefore, if we calculate the slow-roll parameters using equations (2.45) we obtain:

$$\varepsilon \sim \eta \sim \frac{M_{\rm p}^2}{f^2}. \tag{2.59}$$

Thus, the simplest model in which the inflation is a PNGB requires a super-Planckian decay constant f which seems at odds with the notion that quantum gravity effects should already become preponderant at the Planck scale. More complicated models that do not require super-Planckian decay constants have been considered in the literature (e.g. [91, 92]). However, these models are based on the assumption that the UV completion admits some global symmetries, and this clashes against the expectation that all global symmetries should be broken by quantum gravity effects at sufficiently high energies [93].

The other option that we have encountered to keep scalar fields light is supersymmetry. However, it turns out that supergravity corrections will generically introduce corrections to the inflation mass of the order of H [94]. This can be checked explicitly by using the supersymmetric potential (2.7) in the case of a single complex field φ. If we expand the potential in inverse powers of $M_{\rm p}$ and denote with V_0 the leading term responsible for driving inflation, then there are dimension 6 operators of the form

$$\frac{\mathcal{O}}{M_{\rm p}^2} \sim \frac{V_0}{M_{\rm p}^2}|\varphi|^2 \sim H^2|\varphi|^2 \tag{2.60}$$

that induce an $\mathcal{O}(1)$ correction to the eta parameter.[16]

[16]See however [95] for an interesting model in which accidental symmetries are combined with

An additional open problem stems from the fact that many models of inflation predict an accelerated phase that lasts much longer than what is necessary to solve the flatness and horizon problems. Based on the discussion in section 2.3.2, it follows that primordial perturbations with comoving wave numbers of cosmological interest had a physical wavenumber $k/a \gg M_{\rm p}$ at the beginning of inflation. Therefore, the linear analysis of perturbations carried out in this section is implicitly based on the assumption that perturbations remain weakly coupled at arbitrarily high-energies [96]. Such an assumption seems unwarranted, and this is known as the *trans-Planckian problem* [97] of inflationary cosmology. In Chapter 5, we will carefully explore this issue from an EFT perspective.

To conclude this section, we would like to mention the existence of a more pragmatic approach to inflationary model building. Given the difficulties encountered when trying to formulate natural EFT of inflation, we can choose to adopt an EFT with a cut-off $H \ll M \ll \sqrt{\varepsilon} M_{\rm p}$ [98, 99]. This theory will not be able to describe the background evolution, but can be used to study the properties of primordial perturbations produced during inflation. Since all observational constraints on inflationary models come from the study of perturbations, this seems like an acceptable compromise which does not affect the predictive power of inflation.

Appendix 2.A Vierbein Formalism

In any generally covariant theory defined on a spacetime manifold in which the metric has Lorentzian signature, and regardless of whether Lorentz invariance is broken or not, it is always possible to introduce a vierbein, an orthonormal set of forms $\hat{e}^{(a)} = e_\mu{}^a \, dx^\mu$ in the cotangent space of the spacetime manifold,

$$g^{\mu\nu} e_\mu{}^a \, e_\nu{}^b = \eta^{ab}. \tag{2.61}$$

Greek indices $\mu, \nu, \ldots$ now denote cotangent space indices in a coordinate basis, while latin indices $a, b, \ldots$ label the different vectors in the orthonormal set. Thus, the order

supersymmetry to keep the inflaton mass smaller than H.

of the vierbein indices is important. The first one is always a spacetime index, and the second one is always a Lorentz index. Spacetime indices are raised and lowered with the metric of spacetime, and Lorentz indices are raised and lowered with the Minkowski metric. Under coordinate transformations, the vierbein ${e_\mu}^a$ transforms like a vector,

$$\text{diff} : {e_\mu}^a(x) \mapsto {e'_\mu}^a(x') = \frac{\partial x^\nu}{\partial x'^\mu} {e_\nu}^a(x). \tag{2.62}$$

The freedom to choose a vierbein whose sixteen components satisfy the orthonormality condition (2.61) does not add anything to the original ten metric components if the theory remains invariant under the six parameter group of local Lorentz transformations,

$$g(x) : {e_\mu}^a(x) \mapsto {e'_\mu}^a(x) = {\Lambda^a}_b(g)\, {e_\mu}^b(x). \tag{2.63}$$

Note that this transformation does not affect the coordinates of the vectors, that is, the Lorentz group acts as an "internal" symmetry.

The derivatives of the vierbein do not transform covariantly under these local Lorentz transformations. We thus introduce the spin connection $\boldsymbol{\omega}_\mu$, which plays the role of the gauge field of the Lorentz group. Let l^k, $k = 1, \ldots 6$, denote the generators of the Lorentz group (in any representation), and let us define the components of the spin connection by

$$\boldsymbol{\omega}_\mu \equiv \omega_{\mu k}\, l^k, \tag{2.64}$$

which transforms like a one-form under general coordinate transformations,

$$\text{diff} : \boldsymbol{\omega}_\mu(x) \mapsto \boldsymbol{\omega}'_\mu(x') = \frac{\partial x^\nu}{\partial x'^\mu} \boldsymbol{\omega}_\nu(x). \tag{2.65}$$

In complete analogy with gauge field theories, let us assume that under local Lorentz transformations the spin connection transforms as[17]

$$g(x) : \boldsymbol{\omega}_\mu(x) \mapsto g\, \boldsymbol{\omega}_\mu(x)\, g^{-1} + g\, \partial_\mu g^{-1}. \tag{2.66}$$

In that case, it is then easy to verify that the covariant derivative

$$\nabla_\mu {e_\nu}^a = \partial_\mu {e_\nu}^a - {\Gamma^\rho}_{\nu\mu} {e_\rho}^a + \omega_{\mu k}\, {[l^k_{(4)}]^a}_b\, {e_\rho}^b \tag{2.67}$$

[17]To recover expressions fully analogous to those found in gauge theories, the reader should replace $\boldsymbol{\omega}_\mu$ by $-i\boldsymbol{\omega}_\mu$.

transforms covariantly both under coordinate and local Lorentz transformations. Here, $\Gamma^{\mu}{}_{\nu\rho}$ are the Christoffel symbols associated with the spacetime metric $g_{\mu\nu}$, and the $l^k_{(4)}$ are the Lorentz group generators in the fundamental representation, under which the vierbein transforms. In our convention, these matrices are purely imaginary. Similarly, given any matter field ψ that transforms as a scalar under diffeomorphisms, and in a representation of the Lorentz group with generators l_k, we can construct its covariant derivative by

$$\nabla_\mu \psi \equiv \partial_\mu \psi + \boldsymbol{\omega}_\mu \psi, \tag{2.68}$$

which also transforms covariantly both under diffeomorphisms and local Lorentz transformations.

In any generally covariant theory defined on a Riemannian spacetime manifold, the covariant derivative is compatible with the metric, that is, $\nabla_\mu g_{\nu\rho} = 0$. Moreover, because the Minkowski metric is invariant under Lorentz-transformations, its Lorentz-covariant derivative vanishes. Thus, differentiating equation (2.61) covariantly, and using Leibniz rule we obtain

$$\nabla_\nu e_\mu{}^a = 0. \tag{2.69}$$

Equation (2.67), in combination with equation (2.69) allows us to express the spin connection in terms of the vierbein,

$$\omega_{\mu k}[l^k_{(4)}]^a{}_b = \frac{1}{2}\left[e^{\nu a}(\partial_\mu e_{\nu b} - \partial_\nu e_{\mu b}) - e^\nu{}_b(\partial_\mu e_\nu{}^a - \partial_\nu e_\mu{}^a) - e^{\rho a} e^\sigma{}_b(\partial_\rho e_{\sigma c} - \partial_\sigma e_{\rho c}) e_\mu{}^c\right], \tag{2.70}$$

and it is readily verified that this connection indeed transforms as in equation (2.65).

Equation (2.70) is what sets gauge theories and gravity apart. In gauge theories, the gauge fields are "fundamental" fields on which the action functional depends. In gravity the spin connection can be expressed in terms of the vierbein, which constitute the fundamental fields in the gravitational sector. In particular, the metric can be also expressed in terms of the vierbein,

$$g_{\mu\nu} = e_{\mu a} e_\nu{}^a, \tag{2.71}$$

where, as in Subsection 3.2.3, $e^\mu{}_a$ is the (transposed) inverse of $e_\mu{}^a$, that is, $e_\mu{}^a\, e^\nu{}_a = \delta^\mu{}_\nu$. Because the covariant derivative of the vierbein vanishes by construction, one can use the vierbein to freely alter the transformation properties of any field under diffeomorphisms and Lorentz transformations. For instance, $\nabla_\mu A_a \equiv e^\nu{}_a \nabla_\mu A_\nu$, so one can use the vierbein to freely convert diffeomorphism vectors into Lorentz vectors and vice versa.

Since the spin connection transforms like a gauge field, the curvature tensor

$$\boldsymbol{R}_{\mu\nu} \equiv \partial_\mu \boldsymbol{\omega}_\nu - \partial_\nu \boldsymbol{\omega}_\mu + [\boldsymbol{\omega}_\mu, \boldsymbol{\omega}_\nu] \tag{2.72}$$

transforms like a two-form under general coordinate transformations, and in the adjoint representation under local Lorentz transformations $g(x) \in L_+^\uparrow$,

$$g(x) : \boldsymbol{R}_{\mu\nu} \mapsto g\, \boldsymbol{R}_{\mu\nu}\, g^{-1}. \tag{2.73}$$

This transformation law is particularly simple in the fundamental (form) representation of the Lorentz group. In that case, for fixed μ and ν the curvature $\mathbf{R}_{\mu\nu}$ is a matrix $[R_{\mu\nu}]_a{}^b$ whose elements transform according to

$$g(x) : [R_{\mu\nu}]^a{}_b \rightarrow [R'_{\mu\nu}]^a{}_b = \Lambda^a{}_c(g) \Lambda_b{}^d(g) [R_{\mu\nu}]^c{}_d. \tag{2.74}$$

Note that the curvature tensor is antisymmetric in the coordinate and Lorentz indices,

$$R_{\mu\nu ab} = -R_{\nu\mu ab} = -R_{\mu\nu ba}. \tag{2.75}$$

Recall that spacetime indices are raised and lowered with the spacetime metric $g_{\mu\nu}$, and Lorentz indices are raised and lowered with the Minkowski metric η_{ab}.

With these ingredients it is then possible to construct effective Lagrangians which are invariants under both general coordinate and Lorentz transformations. If we were dealing with an actual gauge theory, the appropriate kinetic term for the spin connection would be the curvature squared, but in the case of gravity the situation is slightly different. In fact, in this case the spin connection is not an independent field, but is determined instead by the vierbein. Since $\nabla_\mu e^\nu{}_a$ vanishes by construction, the only scalar invariant under coordinate transformations and local transformations which contains up to two derivatives of the vierbein is the Ricci scalar,

$$R \equiv e^{\mu\, a} e^{\nu\, b} R_{\mu\nu ab}. \tag{2.76}$$

Chapter 3

Lorentz-violating Theories of Gravity

3.1 Introduction

It is hard to overemphasize the central role that the Lorentz group plays in our present understanding of nature. The standard model of particle physics, for instance, consists of all renormalizable interactions invariant under Lorentz transformations and its internal symmetry gauge group, which act on the matter fields of the theory. While most standard model extensions alter either its field content or gauge group, they rarely drop Lorentz invariance [100]. Of course, such a reluctance has a well-established observational support. Elementary particles appear in (irreducible) representations of the Lorentz group, and their interactions seem to be well described by Lorentz-covariant laws. Lorentz-breaking operators in the standard model of particle physics were first considered by Colladay and Kostelecky [101], and Coleman and Glashow [102]. Experimental and observational constraints on such operators are so stringent [76] that it is safe to assume that any violation of Lorentz invariance in the standard model must be extremely small.

The status of the Lorentz group in theories of gravity is somewhat different. Because the group of diffeomorphisms does not admit spinor representations, in generally covariant theories the Lorentz group is introduced as a *local internal* symmetry. Thus, in gravitational theories one formally deals with two distinct groups of transformations: diffeomorphisms and local Lorentz transformations. Even in the context of generally covariant theories, it is thus natural to ask and inquire whether the gravitational interactions respect Lorentz invariance, and what constraints we can impose on any Lorentz-violating gravitational interactions. To date, experimental bounds still allow significant deviations from Lorentz invariance in gravitational interactions [64, 76, 103].

In this chapter we follow a general approach to study theories with broken Lorentz invariance, and address consequences that merely follow from the symmetry breaking pattern, regardless of any specific model of Lorentz symmetry breaking. Such a model-independent approach was first introduced by Weinberg to describe the spontaneous breakdown of chiral invariance in the strong interactions [104], and was subsequently generalized by Callan, Coleman, Wess and Zumino to the breaking of any internal symmetry group down to any of its subgroups [19, 105]. Their approach was further broadened to the case of spontaneous breaking of space-time symmetries [106–110] down to the Poincaré group. Here, we extend all these results to the case in which the Lorentz group itself is broken down to one of its subgroups.

Broken Lorentz symmetry has been mostly explored by means of particular models in which vector fields [111–117] or higher-rank tensors [118] develop a non-vanishing vacuum expectation value. Because the quantities that acquire a vacuum expectation value transform non-trivially under the Lorentz group, in these models the breaking is "spontaneous." This language is commonly accepted[1] for global symmetries, but for local symmetries it has been criticized on several grounds. It is often argued for instance that local symmetries are redundancies in the description of the system rather than actual symmetries [120]. In fact, as we shall show, under certain conditions local Lorentz invariance can be introduced and removed at will. Furthermore, it has been shown that under fairly generic assumptions local symmetries cannot be spontaneously broken [121], in the sense that the vacuum expectation value of a field that is not invariant under the local symmetry group always vanishes. For these

[1]See however [119] for an alternative point of view on this matter.

reasons we shall avoid using "spontaneous symmetry breaking" in the context of local symmetries, though this is the phrasing usually employed to describe the context in which the effective formalism that we employ here applies. To avoid confusions, when we speak of a theory with broken local Lorentz invariance we simply mean a theory which admits an action functional which is invariant under diffeomorphism and only a subgroup of local Lorentz transformations (we sharpen this definition in section 3.3.) Generically, the breaking of diffeomorphism invariance in non-trivial backgrounds can also be understood as Lorentz symmetry breaking [122–124], but this breaking does not fit our definition, is quite different from what we explore here, and indeed leads to quite different phenomenology.

These considerations are not a purely academic exercise, but also have important phenomenological implications. Motivated by cosmic acceleration, several authors have devoted substantial attention to massive theories of gravity [125–128] and other modifications [61, 74, 129], even though the distinction between modifications of gravity and theories with additional matter fields is often blurry. Within the last class, several groups have studied the cosmological dynamics induced by vector fields with non-zero expectation values (see for instance [130–137]), though the breaking of Lorentz invariance has not been the primary focus of their investigations. In these cases, the theory contains massless Goldstone bosons, which participate in long-ranged gravitational interactions and alter the Newtonian and post-Newtonian limits of the theory. Equivalently, we may also think of these additional fields as additional polarizations of the graviton. From this perspective, broken Lorentz invariance offers a new framework to study modifications of gravity, and may cast some light onto theories that have been already proposed.

This chapter is based on the paper [138] and it is structured as follows. In Section 3.2 we generalize the coset construction of Callan, Coleman, Wess and Zumino to theories in which the group of global Lorentz transformations is spontaneously broken. In Section 3.3 we briefly review the role of the Lorentz group as an internal local symmetry group in generally covariant theories, and study the broken Lorentz invariance in this framework. Section 3.4 is devoted to an illustration of our formalism in theories in which the rotation group remains unbroken. We summarize our results in Section 3.5.

3.2 Broken Lorentz Invariance

In this section we explore how to construct theories in which the global symmetry of the action under a given Lorentz subgroup H is manifest (linearly realized), but the global symmetry under the "broken part" of the Lorentz group $L_+^\uparrow$ is hidden (non-linearly realized). This is indeed what happens in a theory in which the Lorentz group $L_+^\uparrow$ is spontaneously broken down to a subgroup H. After a brief review of the Lorentz group, we first consider how to parametrize the broken part of the Lorentz group, that is, the coset $L_+^\uparrow/H$. The corresponding parameters are the Goldstone bosons of the theory. We define the action of the full Lorentz group on this set of Goldstone bosons in such a way that they transform linearly under H, but non-linearly under $L_+^\uparrow/H$. Initially, the transformation that we consider is internal, that is, does not affect the spacetime coordinates of the Goldstone bosons. This is the way the Lorentz group acts in generally covariant theories, which we discuss in Section 3.3, but it is not the way it acts in theories in Minkowski spacetime, in which the Lorentz group is a spacetime symmetry. Hence, we subsequently extend our realization of the Lorentz group to a set of spacetime transformations.

In order to write down Lorentz-invariant theories in which the symmetry under H is manifest, we need to come up with appropriate "covariant" derivatives that transform like the Goldstone bosons themselves. As we shall see, once these covariant derivatives have been identified, the construction of actions invariant under the full Lorentz group becomes straight-forward, and simply reduces to the construction of theories in which invariance under the linearly realized H is explicit.

3.2.1 The Lorentz Group

The Lorentz group L is the set of transformations $\Lambda^a{}_b$ that leave the Minkowski metric invariant, $\eta_{ab}\Lambda^a{}_c\Lambda^b{}_d = \eta_{cd}$. Its component connected to the identity, the proper orthochronous Lorentz group $L_+^\uparrow$, admits 6 generators which can be subdivided into the generators of rotations, J^i, and those of boosts, K^i. The orthochronous group is an invariant subgroup of the Lorentz group and its generators satisfy the following commutation relations:

$$[J_i, J_j] = i\epsilon_{ijk} J^k, \tag{3.1a}$$

$$[J_i, K_j] = i\epsilon_{ijk} K^k, \tag{3.1b}$$

$$[K_i, K_j] = -i\epsilon_{ijk} J^k. \tag{3.1c}$$

Another important subgroup of the Lorentz group is the discrete subgroup $V \equiv \{1, P, T, PT\}$ spanned by the parity transformation P and the time reversal T. It turns out that any Lorentz transformation $\Lambda^a{}_b$ can be expressed as a combination of an orthochronous transformation and, possibly, a parity transformation P and/or a time reversal T. For this reason, the orthochronous group $L_+^\uparrow$ may be understood as the coset

$$L_+^\uparrow = L/V. \tag{3.2}$$

The elements of V also define a map whose square is the identity and which preserves the commutation relations of the Lie algebra (3.1),

$$P \quad : J^i \mapsto PJ^iP^{-1} = J^i, \quad K^i \mapsto PK^iP^{-1} = -K^i \tag{3.3a}$$

$$T \quad : J^i \mapsto TJ^iT^{-1} = J^i, \quad K^i \mapsto TK^iT^{-1} = -K^i. \tag{3.3b}$$

In the rest of this chapter I will be mostly concerned with the breaking of the proper orthochronous Lorentz group $L_+^\uparrow$.

3.2.2 Coset Construction

Suppose now that the proper orthochronous Lorentz group $L_+^\uparrow$ ("Lorentz group" for short) is spontaneously broken down to a subgroup $H \subset L_+^\uparrow$. In the simplest models of this kind, the breaking occurs because the potential energy of a vector field has a minimum at a non-zero value of the field, in analogy with spontaneous symmetry breaking in scalar field theories with Mexican-hat potentials. Perhaps more interesting are cases in which Lorentz invariance is broken "dynamically," that is, when a strong interaction causes fermion bilinears to condense into spacetime vectors [139–141]. This is analogous to the way in which chiral invariance is broken in QCD. The formalism we develop here however does not depend on the actual mechanism that triggers the symmetry breaking, and only relies on the unbroken group H.

Let $\mathcal{H}$ be the Lie algebra of H, which we assume to be semisimple. Although the Lorentz group is not compact, it is simple, so the Killing form $(\cdot,\cdot)$ is non-degenerate and may be regarded as a scalar product on $\mathcal{H}$. We may then uniquely decompose the Lie algebra of $L_+^\uparrow$ into the algebra of H and its orthogonal complement, which we denote by $\mathcal{C}$,

$$L_+^\uparrow = \mathcal{H} \oplus \mathcal{C}. \tag{3.4}$$

Hence, by definition, for any $t \in \mathcal{H}$ and any $x \in \mathcal{C}$, $(t, x) = 0$. In the following we assume that the set of unbroken generators t^i is a basis of $\mathcal{H}$, and that the set of broken generators x^m forms a basis of $\mathcal{C}$. In any representation, l^k collectively denotes the generators of the Lorentz group, $k = 1, \ldots, 6$.

For any $t \in \mathcal{H}$, the map $f_t : x \in \mathcal{C} \mapsto [t, x]$ is linear. Moreover, for any $t' \in \mathcal{H}$ we have

$$(t', [t, x]) = ([t', t], x) = 0, \tag{3.5}$$

where we have used the properties of the Killing form and that $[t, t'] \in \mathcal{H}$. Therefore, f_t maps $\mathcal{C}$ into itself.[2] In fact, the commutator defines a homomorphism of $\mathcal{H}$ into the linear maps of $\mathcal{C}$. Hence, the matrices $C(t)$ with elements defined by

$$[t, x^m] = i C(t)_n{}^m \, x^n \tag{3.6}$$

provide a representation of $\mathcal{H}$. In particular, equation (3.6) implies that, for any element of the unbroken group $h \in H$ and for any $x \in \mathcal{C}$,

$$h \, x \, h^{-1} \in \mathcal{C}. \tag{3.7}$$

[2] It is at this point where the assumption of a semisimple group becomes necessary. As an illustration of this point, consider the case where the unbroken group is spanned by the single generator $t \equiv K^1 + J^2$. Then, the commutation relations (3.1) imply $[t, K^3] = it$, which is not in $\mathcal{C}$.

Following the standard coset construction of Callan, Coleman, Wess and Zumino [19, 105] (see [36, 142] for brief reviews), we can write down realizations of the Lorentz group, in which any given set of fields transform in a linear representation of the unbroken group H. For that purpose, let us first introduce a convenient parametrization of the coset space $L_+^\uparrow/H$. Any element $\gamma \in L_+^\uparrow/H$ can be expressed as

$$\gamma(\pi) = \exp(i\pi_m\, x^m), \tag{3.8}$$

where a sum over indices in opposite locations is always implied. The fields $\pi_m = \pi_m(x)$ correspond to the Goldstone bosons of the theory. If there are M broken generators of the Lorentz group, there are M Nambu-Goldstone bosons π_m.[3]

We may now introduce a realization of the group $L_+^\uparrow$ on this set of Goldstone bosons. By definition, any $g \in L_+^\uparrow$ can be uniquely decomposed into the product of an element of the unbroken group $h \in H$ and a representative γ of the coset space $L_+^\uparrow/H$, such that $g = \gamma\, h$. Therefore, the product $g\,\gamma(\pi) \in L_+^\uparrow$ also has a unique decomposition

$$g\,\gamma(\pi(x)) = \gamma(\pi'(x))\, h(\pi(x), g), \qquad \text{with} \qquad \gamma(\pi') \in L_+^\uparrow/H\ , \quad h(\pi, g) \in H. \tag{3.9}$$

Equation (3.9) defines a non-linear realization of the Lorentz group by mapping π into π' for any given $g \in L_+^\uparrow$. Notice however that this representation becomes linear when g belongs to H. In fact, because of equation (3.7) we must have that $\bar{h}\,\gamma(\pi)\,\bar{h}^{-1} = \gamma(\pi')$ for every $\bar{h} \in H$, and a comparison with equation (3.9) implies

$$h(\pi, \bar{h}) = \bar{h}. \tag{3.10}$$

In particular, use of equations (3.6), (3.8) and (3.10) shows that in this case the Goldstone bosons transform in a linear representation of the unbroken group H,

$$h \in H : \pi_m \mapsto \pi'_m = R(h)_m{}^n\, \pi_n, \quad \text{with} \quad R\,(\exp it) \equiv \exp\left[iC(t)\right]. \tag{3.11}$$

Therefore, the Goldstone bosons have the same “quantum numbers” as the broken generators x^m. For a *compact*, connected, semi-simple Lie group G broken down to H, the uniqueness of the transformation law (3.9), up to field-redefinitions, was proved in [105].

[3]See [35] for exceptions to this argument in the case of spontaneous breaking of translations.

3.2.3 Covariant Derivatives

Thus far, the realization of the Lorentz group that we have defined in equation (3.9) treats the Lorentz group as an internal symmetry; the spacetime arguments on both sides of the equation coincide. This is going to be useful in our discussion of the Lorentz group in generally covariant theories, but it is not the way the Lorentz group acts in conventional field theories in Minkowski spacetime, in which the Lorentz group is a group of spacetime symmetries. Following [108, 110], we define now a non-linear realization of the Lorentz group as a spacetime symmetry by

$$g : \gamma(\pi(x)) \mapsto \gamma(\pi'(x')), \quad \text{where} \quad g\, e^{iP_\mu x^\mu} \gamma(\pi(x)) = e^{iP_\mu x'^\mu} \gamma(\pi'(x')) h(\pi(x), g). \tag{3.12}$$

This implicitly defines a realization of the Lorentz group on the coordinates x^μ and the fields $\pi(x)$. In particular, under an arbitrary element $g \in L_+^\uparrow$, equation (3.12) implies

$$g : x^\mu \mapsto x'^\mu = \Lambda^\mu{}_\nu(g) x^\nu, \quad \gamma(\pi(x)) \mapsto \gamma(\pi'(x')) = \gamma(\pi'(x)), \tag{3.13}$$

with $g P_\mu g^{-1} = \Lambda^\nu{}_\mu(g) P_\nu$ and $\gamma(\pi'(x))$ defined in equation (3.9).

Because we are interested in theories in which the Lorentz group is a set of global symmetries, any action constructed from the Goldstone bosons π can only depend on their derivatives. In order to introduce appropriate covariant derivatives, in analogy with the conventional prescription [19], we expand an appropriately modified [108, 110] Maurer-Cartan form in the basis of the Lie algebra,

$$\boldsymbol{\Omega}_\mu \equiv \frac{1}{i} \gamma^{-1} e^{-iP\cdot x} \partial_\mu (e^{iP\cdot x} \gamma) \equiv e_\mu{}^a P_a + D_{\mu m}\, x^m + E_{\mu i}\, t^i \equiv \boldsymbol{e}_\mu + \boldsymbol{D}_\mu + \boldsymbol{E}_\mu, \tag{3.14}$$

which immediately implies that

$$e_\mu{}^a = \Lambda_\mu{}^a(\gamma). \tag{3.15}$$

The field $e_\mu{}^a$ is the analogue of the vierbein that we shall introduce in Section 3.3. Both transform similarly under the Lorentz group, and this leads to formally identical expressions in both cases. But the reader should nevertheless realize that the "vierbein" (3.15) and the vierbein of Section 3.3 are actually different objects.

The transformation properties of $\boldsymbol{e}, \boldsymbol{D}$ and $\boldsymbol{E}$ follow from the definition (3.12). Under an arbitrary $g \in L_+^\uparrow$, they transform according to

$$g: \quad \boldsymbol{e}_\mu(x) \mapsto \boldsymbol{e}'_\mu(x') = \Lambda_\mu{}^\nu(g)\, h(\pi, g)\boldsymbol{e}_\nu(x) h^{-1}(\pi, g), \tag{3.16a}$$

$$\boldsymbol{D}_\mu(x) \mapsto \boldsymbol{D}'_\mu(x') = \Lambda_\mu{}^\nu(g)\, h(\pi, g)\boldsymbol{D}_\nu(x) h^{-1}(\pi, g), \tag{3.16b}$$

$$\boldsymbol{E}_\mu(x) \mapsto \boldsymbol{E}'_\mu(x') = \Lambda_\mu{}^\nu(g) \left[h(\pi, g)\boldsymbol{E}_\nu(x) h^{-1}(\pi, g) - i h(\pi, g)\partial_\nu h^{-1}(\pi, g)\right], \tag{3.16c}$$

where $h(\pi, g)$ is defined in equation (3.9). Therefore, none of these quantities really transforms covariantly, since the spacetime index μ and the components of the different fields transform under different group elements. To define fully covariant quantities, let us introduce the inverse of the quantity defined in equation (3.15),

$$e^\mu{}_a = \Lambda_a{}^\mu(\gamma^{-1}). \tag{3.17}$$

This is indeed the (transposed) inverse of $e_\mu{}^a$ because it follows from equation (3.15) that $e^\mu{}_a e_\mu{}^b = \delta_a{}^b$. Then, the quantities

$$\boldsymbol{\mathcal{D}}_a \equiv e^\mu{}_a \boldsymbol{D}_\mu, \quad \boldsymbol{\mathcal{E}}_a \equiv e^\mu{}_a \boldsymbol{E}_\mu, \tag{3.18}$$

do transform covariantly under the Lorentz group,

$$\boldsymbol{\mathcal{D}}_a(x) \mapsto \boldsymbol{\mathcal{D}}'_a(x') = \Lambda(h(\pi, g))_a{}^b\, h(\pi, g)\boldsymbol{\mathcal{D}}_b(x) h^{-1}(\pi, g), \tag{3.19a}$$

$$\boldsymbol{\mathcal{E}}_a(x) \mapsto \boldsymbol{\mathcal{E}}'_a(x') = \Lambda(h(\pi, g))_a{}^b \left[h(\pi, g)\boldsymbol{\mathcal{E}}_b(x) h^{-1}(\pi, g) - i h(\pi, g)\partial_b h^{-1}(\pi, g)\right], \tag{3.19b}$$

where $\partial_a \equiv e^\mu{}_a \partial_\mu$. We identify $\boldsymbol{\mathcal{D}}_a$ with the covariant derivative of the Goldstone bosons π_m, and $\boldsymbol{\mathcal{E}}_a$ with a "gauge field" that will enter the couplings between the Goldstone bosons and other matter fields. The transformation rules (3.19) are again non-linear in general, but, because of equation (3.10), they reduce to a linear transformation if $g \in H$. Note that under $g \in L_+^\uparrow$, the components of the covariant derivative $\boldsymbol{\mathcal{D}}_a$ transform as

$$g: \mathcal{D}_{am}(x) \mapsto \mathcal{D}'_{am}(x') = \Lambda_a{}^b(h(\pi, g)) R_m{}^n \mathcal{D}_{bn}(x), \tag{3.20}$$

where the matrix R is the one we introduced in equation (3.11).

For specific calculations, it is often required to have concrete expressions for the covariant derivatives. It follows from the definitions (3.8) and (3.14) that

$$\mathcal{D}_{am} = \partial_a \pi_m - i\pi_n (x^n_{(4)})_a{}^b \partial_b \, \pi_m + \frac{1}{2}\pi_n \partial_a \pi_p \, C^{np}{}_m + \mathcal{O}(\pi^3), \tag{3.21}$$

where $x^n_{(4)}$ is the fundamental (form) representation of the Lorentz generator x^n, and the $C^{np}{}_m$ are the structure constants of the Lie algebra $\mathcal{H}$ in our basis of generators.

Parity and Time Reversal

In certain cases, we can also define the transformation properties of the Goldstone bosons under parity and time reversal, or, in general, under an appropriate subgroup of $V \equiv \{1, P, T, PT\}$. Let V_H denote the "stabilizer" of H, that is, the set of all elements $v \in V$ that leave H invariant, $v \, h \, v^{-1} \in H$ for all $h \in H$. This is a subgroup of V, which may contain just the identity, either P or T, or V itself. Because $\mathcal{H}$ is invariant under V_H, the latter defines an homomorphism on $\mathcal{C}$ by conjugation,

$$v \, x^m v^{-1} = V_n{}^m \, x^n. \tag{3.22}$$

The two sets $L^{\uparrow}_{+} V_H$ and HV_H are two subgroups of L, and, by definition, HV_H is a subgroup of $L^{\uparrow}_{+} V_H$. Thus, just as in Section 3.2.2 , we may define a realization of V_H (which is now contained in $L^{\uparrow}_{+} V_H$) on the coset

$$\frac{L^{\uparrow}_{+} V_H}{HV_H} = \frac{L^{\uparrow}_{+}}{H}. \tag{3.23}$$

In particular, for $g \in L^{\uparrow}_{+} V_H$ and $\gamma(\pi) \in L^{\uparrow}_{+}/H$ we set

$$g\gamma(\pi) = \gamma(\pi')h(\gamma, g)v(\gamma, g), \quad \text{with} \quad h(\gamma, g) \in H \text{ and } v(\gamma, g) \in V_H. \tag{3.24}$$

If $g \in L^{\uparrow}_{+}$, this definition reduces to that of equation (3.9). For $v \in V_H$ it leads to

$$v : \gamma(\pi) \mapsto \gamma(\pi') = v \, \gamma(\pi) \, v^{-1}, \tag{3.25}$$

which can be extended to include the arguments of the Goldstone boson fields as before,

$$v : \gamma(\pi(x)) \mapsto \gamma(\pi'(x')), \quad \text{where} \quad v \, e^{iP\cdot x} \gamma(\pi(x)) \, v^{-1} = e^{iP\cdot x'} \gamma(\pi'(x')). \tag{3.26}$$

Under these group elements the Goldstone bosons change according to

$$v : \pi_m \mapsto \pi'_m(x') = V_m{}^n \, \pi_n(x), \tag{3.27}$$

and, from equation (3.20), their covariant derivatives according to

$$v : D_{am}(x) \mapsto D'_{am}(x') = V_a{}^b V_m{}^n D_{bn}(x), \tag{3.28}$$

where $v P_a v^{-1} = V_a{}^b P_b$.

3.2.4 Invariant Action

If we are interested in the low-energy limit of theories in which Lorentz-invariance is broken, we can restrict our attention to their massless excitations. This is a restatement of the Appelquist-Carazzone theorem [143], though the latter has been actually proven only for renormalizable Lorentz-invariant theories in flat spacetime. Typically, massless fields are those protected by a symmetry, and always include the Goldstone bosons, since invariance under the broken symmetry prevents them from entering the action undifferentiated. Therefore, the low-energy effective action of any theory in which Lorentz invariance is broken must contain the covariant derivatives of the Goldstone bosons. To leading order in the low-energy expansion, we can restrict our attention to the minimum number of spacetime derivatives, namely, two.

The tensor product representation in equation (3.20) under which the covariant derivatives transform is in general reducible. Let $\Lambda \otimes R = \oplus_i R^{(i)}$ be its Clebsch-Gordan series, and let $\mathcal{D}^{(i)}$ be the linear combination of covariant derivatives that furnishes the i-th irreducible representation. Some of these representations may be singlets, and we shall label them by s. Because the unbroken group is not necessarily compact, the non-trivial irreducible representations are generally not unitary. In any case, if $G^{(i)}$ is invariant under the i-th representation of the unbroken group H, i.e. $R^{(i)T} G^{(i)} R^{(i)} = G^{(i)}$, then the Lagrangian

$$\mathcal{L} = \sum_s F_s \mathcal{D}^{(s)} + \sum_i F_i \, \mathcal{D}^{(i)T} G^{(i)} \mathcal{D}^{(i)} \tag{3.29}$$

transforms as a scalar under the Lorentz-group $L_+^\uparrow$. Here, F_s and F_i are free parameters in the effective action, which remain undetermined by the symmetries of the theory. In order to construct a Lorentz-invariant action, we just need a volume element that transforms appropriately under our realization of the Lorentz group. This is in general given by [110]

$$d^4V \equiv d^4x \det e_\mu{}^a, \tag{3.30}$$

which, because of equation (3.15), results in $d^4V = d^4x$. (Inside the determinant, the vierbein should be regarded as a 4×4 matrix with rows labeled by μ and columns labeled by a.) The functional

$$S = \int d^4V \boldsymbol{\mathcal{L}} \tag{3.31}$$

is then invariant under the action of the Lorentz group defined by equation (3.12).

3.2.5 Couplings to Matter

The formalism can be also extended to capture the effects of Lorentz breaking on the matter sector. As mentioned above, at low-energies we can restrict our attention to massless (or light) fields, which are typically those that are prevented from developing a mass by a symmetry like chiral or gauge invariance. We consider couplings to the graviton in Section 3.3.

Let ψ be any matter field that transforms under any (possibly reducible) representation $\mathscr{R}(h)$ of the unbroken group H, with generators t_i. Let us now *define* the transformation law under the full Lorentz group to be [105]

$$g : \psi(x) \mapsto \psi'(x') = \mathscr{R}(h(\pi, g))\, \psi(x), \tag{3.32}$$

where x' and $h(\pi, g)$ are given in equation (3.12). We can also construct covariant derivatives under the Lorentz group by setting,

$$\mathscr{D}_a\psi \equiv e^\mu{}_a \left[\partial_\mu \psi + i\boldsymbol{E}_\mu \psi\right] = \partial_a \psi + i\boldsymbol{\mathcal{E}}_a \psi, \tag{3.33}$$

where $\boldsymbol{E}_\mu$ is defined in equation (3.14). The covariant derivative transforms just as the field itself, under a representation of the same group element,

$$g : \mathscr{D}_a\psi(x) \mapsto \mathscr{D}'_a\psi'(x') = \Lambda(h(\pi, g))_a{}^b \mathscr{R}(h(\pi, g))\, \mathscr{D}_b\psi(x). \tag{3.34}$$

Therefore, any Lagrangian built out of d^4V, ψ, $\mathscr{D}_a\psi$ and $\mathcal{D}_{am}$ that is invariant under the unbroken group H is then invariant under the full Lorentz group.

With these ingredients we could develop a formulation of the standard model in which the Lorentz group is spontaneously broken to any subgroup. If the unbroken group is trivial, $H = 1$, this construction would be analogous to the standard model extension considered by Colladay and Kostelecky [101]. This chapter mainly focuses on the general formalism of broken Lorentz invariance, so we shall not carry out this program here. For the purpose of illustration however, and in order to establish the connection to previous work on the subject, let us consider a formulation of QED (quantum electro-dynamics) in which the Lorentz group is completely broken. For simplicity we consider a theory with a single "spinor" ψ_α of charge q coupled to a "photon" A_a. We use quotation marks because, according to (3.32), we assume that under the (completely) broken Lorentz group both fields are invariant. On the other hand, we require that the theory be invariant under gauge transformations, that is, we demand invariance under

$$\psi_\alpha \to e^{iq\chi}\psi_\alpha, \quad A_a \to A_a + \partial_a \chi, \tag{3.35}$$

where χ is an arbitrary spacetime scalar. If the Lorentz group is broken down to $H = 1$, there are six Goldstone bosons in the theory, and γ becomes $\gamma \equiv \exp(i\pi^k l_k)$, which, under the Lorentz group transforms as $g : \gamma \mapsto \gamma'(x') = g\,\gamma(x)$. Following (3.33) we introduce now the covariant derivatives

$$\mathscr{D}_a A_b \equiv \Lambda(\gamma^{-1})_a{}^\mu \partial_\mu A_b, \qquad \mathscr{D}_a \psi_\alpha \equiv \Lambda(\gamma^{-1})_a{}^\mu \partial_\mu \psi_\alpha, \tag{3.36}$$

which by construction are Lorentz-invariant (if the Lorentz group is completely broken, $\boldsymbol{E}_\mu \equiv 0$ by definition.) Gauge invariance then dictates that the derivatives of the fields must enter in the gauge invariant or covariant forms

$$F_{ab} \equiv \mathscr{D}_a A_b - \mathscr{D}_a A_b, \quad \nabla_a \psi_\alpha \equiv (\mathscr{D}_a - iqA_a)\psi_\alpha. \tag{3.37}$$

Any gauge invariant combination of these elements, such as

$$\mathcal{L}_{QED} = M^{abcd} F_{ab} F_{cd} + N^{\alpha\beta a} \psi^\dagger_\alpha \nabla_a \psi_\beta + P^{\alpha\beta} \psi^\dagger_\alpha \psi_\beta, \tag{3.38}$$

is also Lorentz invariant (for simplicity, we have not written down all the terms compatible with the two symmetries). In equation (3.38), the dimensionless matrices M, N and P are *constant* and *arbitrary*, up to the restrictions imposed by permutation symmetry and hermiticity. The Lagrangian (3.38) is thus the analogue of the extension of QED described in [101]. From a phenomenological perspective, its coefficients can be regarded as quantities to be determined or constrained by experiment, as in the standard model extension of [101]. But of course, as opposed to the latter, the Lagrangian (3.38) contains couplings to the Goldstone bosons, and should be supplemented with the Goldstone boson Lagrangian, which for a trivial H is

$$\mathcal{L}_\pi = G^{am} D_{am} + F^{mnab} \mathcal{D}_{am} \mathcal{D}_{bn}, \tag{3.39}$$

where $\mathcal{D}_{am}$ is given in equation (3.14), and $m, n = 1, \ldots, 6$. As we shall see in the next section, in a gravitational theory these covariant derivatives should be included in the Lagrangian too, but in that case they reduce to appropriate components of the spin connection. Note that in our conventions the Goldstone bosons are dimensionless. Thus the coefficients in G have mass dimension three, and those in F mass dimension two. In theories in which an internal symmetry is spontaneously broken, Lorentz invariance and invariance under the unbroken group often restrict the possible different mass scales appearing in the Lagrangian to a single scale. This single energy scale is the identified with the scale at which the internal symmetry group is broken. In our case however, the values of G and F are (up to symmetry under permutations) completely arbitrary, so the identification of a single energy scale at which Lorentz symmetry is broken is in general not possible.

The obvious problem with this approach is that the Lorentz group seems to be an unbroken symmetry in the matter sector. A generic Lagrangian like (3.38), constructed out of the standard model fields ψ, their covariant derivatives $\mathscr{D}_a\psi$ and the covariant derivatives of the Goldstones $\mathcal{D}_{am}$ would clearly violate Lorentz invariance, in flagrant conflict with experimental constraints [76]. Thus, we are forced to assume that these "Lorentz-violating" terms are sufficiently suppressed, which in our context requires specific relations between the coefficients in the effective Lagrangian.

To illustrate this point, let us briefly discuss how to construct scalars under linearly realized Lorentz transformations out of the ingredients at our disposal, namely, ψ, $\mathscr{D}_a\psi$ and $\mathcal{D}_{am}$. Imagine that the matter fields $\tilde{\psi}$ actually fit in a representation of the Lorentz group $\mathscr{R}(g)$. It is then more convenient to postulate that under the *full* Lorentz group, these fields transform as

$$g : \tilde{\psi}(x) \mapsto \tilde{\psi}'(x') = \mathscr{R}(g)\tilde{\psi}(x). \tag{3.40}$$

Then, any Lagrangian that is invariant (a scalar) under *global* Lorentz transformations,

$$\mathcal{L}_{\text{inv}}[\tilde{\psi}, \partial_\mu\tilde{\psi}] = \mathcal{L}_{\text{inv}}[\mathscr{R}(g)\tilde{\psi}, \mathscr{R}(g)\Lambda(g)_\mu{}^\nu\partial_\nu\tilde{\psi}], \quad g \in L_+^\uparrow, \tag{3.41}$$

is clearly invariant under the unbroken subgroup H of global transformations, and can thus be part of the effective Lagrangian in the broken phase. Note that these Lorentz invariant terms would not contain any couplings to the Goldstone bosons. But given the transformation law (3.40) we can also construct a new quantity that transforms under the non-linear realization of the Lorentz group (3.32),

$$\psi \equiv \mathscr{R}(\gamma^{-1})\tilde{\psi}, \tag{3.42}$$

and whose covariant derivative can again be defined by equation (3.33). In this case, however, the field ψ is to be understood simply as a shorthand for the right hand of equation (3.42), which contains the Goldstone bosons $\gamma(\pi)$. Given any Lagrangian $\mathcal{L}_{\text{break}}$ that is invariant under the linearly realized unbroken group H, but not invariant under linear representations of the full Lorentz group $L_+^\uparrow$,

$$\mathcal{L}_{\text{break}}[\psi, \partial_\mu\psi] = \mathcal{L}_{\text{break}}[\mathscr{R}(h)\psi, \mathscr{R}(h)\Lambda(h)_\mu{}^\nu\partial_\nu\psi], \quad h \in H, \tag{3.43}$$

we can then construct further invariants under Lorentz transformations,

$$\mathcal{L}_{\text{break}}[\mathscr{R}(\gamma^{-1})\tilde{\psi}, \mathscr{D}_\mu(\mathscr{R}(\gamma^{-1})\tilde{\psi})]. \tag{3.44}$$

Here, the appearance of the Goldstone bosons in those terms that violate the full Lorentz symmetry is manifest.

It seems now that the Lagrangians (3.41) and (3.44) do not fit into the general prescription to construct invariant Lagrangians that we described at the beginning of this subsection, but this is just an appearance. Suppose we perform a field redefinition $\mathscr{R}(\gamma^{-1})\tilde{\psi} \to \psi$, and assume that the new field ψ transforms as in equation (3.32). This field redefinition turns the Lagrangian in equation (3.44) into $\mathcal{L}_{\text{break}}[\psi, \mathscr{D}_\mu\psi]$, and takes the Lagrangian (3.41) into

$$\mathcal{L}_{\text{inv}}[\psi, \mathscr{D}_\mu\psi + iD_{\mu m}x^m\psi]. \tag{3.45}$$

Both Lagrangians are invariant under the linearly realized symmetry group H (and the non-linearly realized Lorentz group $L_+^\uparrow$), and both are solely constructed in terms of $\psi, \mathscr{D}_\mu\psi$ and $\mathcal{D}_{am}$.

Of course, a general Lagrangian invariant under H will have the form of equation (3.45) only for very particular choices of the coefficients that remain undetermined under the unbroken symmetry. From the point of view of the effective theory, this particular choice cannot be explained, though it is certainly compatible with the symmetries we are enforcing. To address it we would have to rely on specific models. Say, if Lorentz symmetry is broken in a hidden sector which is completely decoupled from the standard model, the breaking in the hidden sector should not have any impact on the visible sector. But of course, the two sectors must couple at least gravitationally. Then, if the scale of Lorentz-symmetry breaking is sufficiently small compared to the Planck mass, we expect a double suppression of Lorentz-violating terms in the matter sector: from the weakness of gravity, and from the smallness of the symmetry breaking scale. We defer the discussion of gravitation to the next section. Radiative corrections to Lorentz-violating couplings in the matter sector of Einstein-aether models [114] have been calculated in [144].

3.2.6 Broken Rotations

As an example of the formalism discussed so far, we shall briefly study a pattern of symmetry breaking in which the unbroken group H is non-compact. This is an interesting case since, for internal non-compact symmetry groups, the theory contains ghosts in the spectrum of Goldstone bosons [140, 142]. We show that, instead, it is certainly possible to have a well-behaved spectrum in a theory in which the Lorentz group is broken down to a non-compact subgroup. We consider the widely-studied case of unbroken rotations, $H = SO(3)$, in Section 3.4.

Suppose that the Lorentz group $L_+^\uparrow$ is broken down to the group of transformations that leave the vector field $A^\mu = (0, 0, 0, F)$ invariant. This breaking pattern was studied in references [139, 140], in which the "photon" of electromagnetism is identified with the Goldstone bosons associated with the breaking. The Lie algebra of the unbroken group H is then

$$\mathcal{H} = \text{Span}\{K^1, K^2, J^3\}, \tag{3.46}$$

which is simple, and isomorphic to the Lie algebra of the group of Lorentz transformations in three-dimensional spacetime $so(1, 2)$. Its orthogonal complement is spanned by

$$\mathcal{C} = \text{Span}\{J^1, J^2, K^3\}. \tag{3.47}$$

Because $\dim(\mathcal{C}) = 3$, there are three Goldstone bosons in the theory. It follows from the commutation relations (3.1) and equations (3.6) and (3.11) that $\pi_m \equiv (\pi_3, \pi_1, \pi_2)$ transforms like a Lorentz three-vector. It is thus convenient to let m run from 0 to 2 and identify $\pi_0 \equiv \pi_3$.

The covariant derivative $\mathcal{D}_{am}$ transforms in a reducible representation of the subgroup $H = SO(1, 2)$. In fact, the covariant derivative

$$\mathcal{D}_m \equiv \mathcal{D}_{3m} \tag{3.48}$$

is an $SO(1, 2)$ three-vector. The remaining irreducible spaces are spanned by the scalar φ, the vector a_{mn} and the tensor s_{mn} defined by

$$\varphi \equiv \mathcal{D}_m{}^m, \;\; a_{mn} \equiv \frac{1}{2}(\mathcal{D}_{mn} - \mathcal{D}_{nm}), \;\; s_{mn} \equiv \frac{1}{2}(\mathcal{D}_{mn} + \mathcal{D}_{nm}) - \frac{1}{3}\varphi\, \eta_{mn}, \tag{3.49}$$

where indices are raised with the (inverse) of the Minkowski metric in three dimensions, $\eta^{mn} = \text{diag}(-1, 1, 1)$ and $m = 0, 1, 2$. Scalar invariants are constructed then by appropriate contraction of indices,

$$\mathcal{L}_\pi = G_\varphi \varphi + F_\varphi\, \varphi^2 + F_D \mathcal{D}_m \mathcal{D}^m + F_a\, a_{mn} a^{mn} + F_\epsilon\, \epsilon_{mnp}\, a^{mn} \mathcal{D}^p + F_s\, s_{mn} s^{mn}. \tag{3.50}$$

For simplicity, let us now consider the case where $G_\varphi = 0$. Because to lowest order in the Goldstone bosons $\mathcal{D}_{mn} = \partial_m \pi_n + \cdots$, inspection of (3.50) reveals the lower-dimension analogue of a generalized vector field theory in which the vector field consists of the Goldstone bosons π_m. This analogy can be further strengthened by dimensionally reducing the four dimensional theory from four to three spacetime dimensions. Expanding the Goldstone bosons in Kaluza-Klein modes

$$\pi_m(t,x,y,z) = \sum_{k_z} \pi_m^{(k)}(t,x,y) e^{ikz}, \tag{3.51}$$

and inserting into the action we obtain, to quadratic order,

$$S = \sum_k S_k, \quad \text{where} \tag{3.52}$$
$$S_k[\pi_m^{(k)}] = \int dt\, d^2x \left[\frac{F_a + F_s}{4}\, a_{mn} a^{mn} + \left(F_\varphi + \frac{2F_s}{3}\right) (\partial_m \pi^m)^2 + F_D\, k^2\, \pi_m \pi^m \right].$$

Note that we have suppressed the index k of the Kaluza-Klein modes on the right hand side of equation (3.52). The Kaluza-Klein modes $\pi^{(k=0)}$ are massless, and transform like an $SO(1,2)$ vector. They can be thought of as the Goldstone bosons associated with the breaking $L_+^\uparrow \sim SO(1,3) \to SO(1,2)$ induced by the compactification.

The spectrum of excitations in the theory described by (3.52), and the conditions that stability imposes on the free parameters F_a, F_φ, F_s and F_D can be derived by relying on the similarity of the action S_k with the four-dimensional models analyzed in [137]. Since their stability analysis does not crucially depend on the dimensionality of spacetime, their results also apply in the case at hand.[4] Following the analysis in Section V of [137] we find:

i) If both $F_a + F_s$ and $F_\varphi + 2F_s/3$ are different from zero, the spectrum consists of an $SO(1,2)$ vector and an $SO(1,2)$ scalar. There is always a ghost at high spatial momenta ($k_x^2 + k_y^2 \gg k^2$).

[4] There is just one difference between the four-dimensional and the three-dimensional case: In four dimensions, the vector sector (under spatial rotations) contains two modes, while in three dimension the vector sector (under spatial rotations) only contains one mode.

ii) For $F_a + F_s = 0$ the theory is stable if $F_\varphi + 2F_s/3 > 0$ and $F_D < 0$. The spectrum consists of a scalar under $SO(1,2)$. If $F_D k^2 = 0$, there are no dynamical fields in the spectrum.

iii) For $F_\varphi + 2F_s/3 = 0$ the Lagrangian is the three-dimensional version of the Proca Lagrangian. The spectrum consists of a massive $SO(1,2)$ vector, with two polarizations. The theory is stable for $F_a + F_s > 0$ and $F_D < 0$. If $F_D k^2 = 0$ the vector is massless, with only one polarization. This last cast corresponds to electrodynamics in three spacetime dimensions.

Hence, as we anticipated there are theories in which the low-energy theory is free of ghosts. These are however non-generic, in the sense that they require the coefficients of certain terms otherwise allowed by Lorentz invariance to be zero.

3.3 Coupling to Gravity

In the previous section we have explored spontaneous symmetry breaking of Lorentz invariance in Minkowski spacetime, in which the Lorentz group is a global symmetry. Though this approach should appropriately capture the local physical implications of the breaking in non-gravitational phenomena, it certainly does not suffice to study arbitrary spacetime backgrounds, or the gravitational interactions themselves.

In order to extend these considerations to gravity, it is convenient to exploit the formal analogy between gravity and gauge theories. For that purpose, one introduces the Lorentz group $L_+^\uparrow$ as an "internal" group of symmetries, in addition to the symmetry under general coordinate transformations [145]. In theories with fermions (such as the standard model) this is actually mandatory, as the group of general coordinate transformations does not admit spinor representations. In the first part of this section we review the standard formulation of GR as a gauge theory of the Lorentz group [146]. In the second part we then extend this standard formulation to theories in which Lorentz invariance is "broken." As for global symmetries, we say that local Lorentz invariance is broken down to a subgroup H if the theory admits a generally covariant formulation in which invariance under local transformations in H is manifest (linearly-realized), and invariance under local transformations in the broken part of $L_+^\uparrow$ is hidden (non-linearly-realized.) Readers already familiar with the vierbein formalism can skip directly to Subsection 3.3.1.

3.3.1 Broken Lorentz Symmetry

The extension of this formalism to theories with broken Lorentz invariance is relatively straight-forward, and parallels the standard construction in flat spacetime. We begin by constructing the most general theory invariant under (linearly realized) local transformations in a Lorentz subgroup H and general coordinate transformations, and then we show that, by introducing Goldstone bosons, the theory can be made explicitly invariant under the full (non-linearly realized) Lorentz group.

Unitary Gauge

Let us first postulate the existence of a vierbein ${e_\mu}^a$ that transforms linearly under local transformations in a subgroup of the Lorentz group,

$$h(x) : {e_\mu}^a(x) \mapsto {e'_\mu}^a(x) = \Lambda^a{}_b(h)\, {e_\mu}^b, \quad h(x) \in H \subset L_+^\uparrow. \tag{3.53}$$

This particular transformation law shall later allow us to extend the local symmetry of the action from H to the full Lorentz group $L_+^\uparrow$. Given this vierbein, we *define* the spacetime metric to be

$$g_{\mu\nu} \equiv {e_\mu}^a {e_\nu}^b \eta_{ab}. \tag{3.54}$$

It follows then from the definition of the metric that the vierbein forms a set of orthonormal vectors, as in equation (2.61), and that the volume element (3.30) is invariant both under general coordinate and Lorentz transformations.

In order to construct derivatives that transform covariantly under local transformations in H, we need to postulate the existence of an appropriate connection $\boldsymbol{\omega}_\mu$. If we want to avoid introducing extraneous ingredients into the gravitational sector, we should construct such a gauge field solely in terms of the vierbein, as in the standard construction. Inspection of equations (2.66) and (2.70) reveals that if we define $\boldsymbol{\omega}_\mu$ by equation (2.70), under an element of H the connection transforms like

$$h(x) : \boldsymbol{\omega}_\mu \mapsto h\, \boldsymbol{\omega}_\mu\, h^{-1} + h\, \partial_\mu h^{-1}. \tag{3.55}$$

But as opposed to the original construction in which we demanded invariance under the full Lorentz group, the reduced symmetry in the broken case allows us to introduce additional covariant quantities. In particular, expanding the connection in the basis of broken and unbroken generators,

$$\boldsymbol{\omega}_\mu \equiv i \left(D_{\mu m}\, x^m + E_{\mu i}\, t^i \right) \equiv i \left(\boldsymbol{D}_\mu + \boldsymbol{E}_\mu \right), \tag{3.56}$$

it is then easy to verify that $\boldsymbol{D}_\mu$ transforms covariantly (under H), while $\boldsymbol{E}_\mu$ transforms like a gauge field,

$$h(x) \quad : \quad \boldsymbol{D}_\mu(x) \mapsto h \boldsymbol{D}_\mu(x) h^{-1}, \tag{3.57a}$$

$$\boldsymbol{E}_\mu(x) \mapsto h \boldsymbol{E}_\mu(x) h^{-1} - i\, h\, \partial_\mu h^{-1}. \tag{3.57b}$$

These transformation laws are analogous to those in equations (3.16). The only difference, setting $g = h$ and using equation (3.10), is that in the latter the Lorentz group acts a transformation in spacetime, which changes the spacetime coordinates of the fields, while here the Lorentz group acts internally, and thus leaves the spacetime dependence of the fields unchanged.

The transformation properties of $\boldsymbol{E}_\mu$ allow us to define another covariant derivative of the vierbein, $\bar{\nabla}_\rho {e_\mu}^a = \partial_\rho {e_\mu}^a - {\Gamma^\nu}_{\mu\rho} {e_\nu}^a - i E_{\rho i} (t^i_4)^a{}_b\, {e_\mu}^b$. But because $\nabla_\rho {e_\mu}^a = 0$, this derivative equals $-i D_{\nu m} (x^m_4)_a{}^b e^\mu{}_b$, and therefore does not yield any additional covariant quantity. Finally, from the connection $\boldsymbol{\omega}_\mu$ we define the curvature (2.72), which under (3.53) transforms like

$$h(x) : \boldsymbol{R}_{\mu\nu} \mapsto h\, R_{\mu\nu}\, h^{-1}. \tag{3.58}$$

In order to construct invariants under both diffeomorphisms and local Lorentz transformations, it is convenient to consider quantities that transform as scalars under diffeomorphisms, and tensors under the unbroken Lorentz subgroup H. We thus define, in full analogy with equations (3.18),

$$\boldsymbol{\mathcal{D}}_a \equiv e^\mu{}_a \boldsymbol{D}_\mu, \quad \boldsymbol{\mathcal{E}}_a \equiv e^\mu{}_a \boldsymbol{E}_\mu, \quad \boldsymbol{\mathcal{R}}_{ab} \equiv e^\mu{}_a e^\nu{}_b \boldsymbol{R}_{\mu\nu}. \tag{3.59}$$

The quantities $\boldsymbol{\mathcal{D}}_a$ and $\boldsymbol{\mathcal{E}}_a$ are the appropriate generalization of the covariant derivatives defined in equation (3.18), since they also transform like in equation (3.19), the only difference being again that here the Lorentz group acts as an internal transformation. As before, the covariant derivatives of any diffeomorphism scalar ψ that transforms in a representation of the unbroken group with generators t^i are defined by equation (3.33), where $E_{\mu i}$ is now given by equation (3.56).

By construction, any term solely built from the covariant quantities d^4V, $\mathcal{D}_{am}$, $\mathcal{R}_{ab}{}^{cd}$, ψ and $\mathscr{D}_a\psi$, which is invariant under global H transformations is also invariant under local transformations in H and diffeomorphisms. In particular, because the covariant derivatives D_{am} defined in (3.18) and the the covariant derivatives in equation (3.59) transform in the same way under H, the unbroken symmetries now allow us to write down linear and quadratic terms for the components of the connection $\boldsymbol{\omega}_\mu$ along the directions of the broken generators, as in equation (3.29). In an ordinary gauge theory, the quadratic terms give mass to some gauge bosons, but in our context, because the spin connection depends on derivatives of the vierbein, these quadratic terms cannot be properly considered as mass terms for the graviton. Since the spacetime metric is $g^{\mu\nu} = e^\mu{}_a e^{\nu a}$, a graviton mass term should be a quartic polynomial in the vierbein. But the only invariants one can construct from the vierbein $e^\mu{}_a$ are field-independent constants. In gravitational theories in which the spin connection is a fundamental field however, quadratic terms in the spin connection can be regarded as mass terms [147–149].

Manifestly Invariant Formulation

Let us assume now that we have constructed an H invariant action,

$$S[e,\psi] = S[\Lambda(h)e, \mathscr{R}(h)\psi], \quad h(x) \in H, \tag{3.60}$$

where the functional dependence emphasizes that only e and ψ are the "fundamental" fields of the theory, from which the remaining covariant quantities are constructed, as discussed above. We show next that by introducing the corresponding Goldstone bosons in the theory $\gamma \equiv \gamma(\pi_m)$, this symmetry can be extended to the full Lorentz group. To that end, let us assume that the vierbein $e^\mu{}_a$ transforms in a linear representation of the Lorentz group, as in equation (2.63), and let us define

$$\tilde{e}_\mu{}^a \equiv \Lambda^a{}_b(\gamma^{-1}) e_\mu{}^b, \tag{3.61}$$

where $\tilde{e}^\mu{}_a$ is to be regarded as a shorthand for the expression on the right hand side, and γ is a function of the Goldstone bosons defined in equation (3.8). Let us postulate that under local Lorentz transformations, $\gamma(\pi)$ transforms as in equation (3.9), while under $g(x) \in L_+^\uparrow$,

$$g(x) : \psi \mapsto \mathscr{R}(h(\pi, g))\psi. \tag{3.62}$$

In that case, it follows from the definition (3.61) that $\tilde{e}$ transforms analogously,

$$g(x) : \tilde{e}^\mu{}_a \mapsto \Lambda_a{}^b(h(\pi, g))\, \tilde{e}^\mu{}_b. \tag{3.63}$$

The transformation properties (3.62) and (3.63) and the invariance of the action (3.60) imply that a theory with

$$\tilde{S}[\gamma, e, \psi] \equiv S[\Lambda(\gamma^{-1})e, \psi] \tag{3.64}$$

is invariant under the full Lorentz group. In the Lorentz-invariant formulation of the theory in equation (3.64) the action appears to depend on the Goldstone bosons $\gamma(\pi)$. However, inspection of the right hand side of the equation reveals that such a dependence can be removed by the field redefinition (3.61). By a "field redefinition" we mean here a change of variables in the theory, which replaces the combination of two fields $\Lambda(\gamma^{-1})e$ by a single field, which we may call again e. Since the field variables we use do not have any impact on the physical predictions of a theory, we may thus replace $S[\Lambda(\gamma^{-1})e, \psi]$ by $S[e, \psi]$. In this "unitary gauge" we have effectively set $\gamma = 1$, and returned back to the original action in equation (3.60).

It is instructive to show how the introduction of the Goldstone bosons would make the theory manifestly invariant under local transformations. For simplicity, let us just focus on the gravitational sector. As mentioned above, the modified vierbein (3.61) transforms non-linearly under the action (2.63) of the Lorentz group, $g(x) \in L_+^\uparrow$. When we substitute this modified vierbein into the expression for the spin connection (2.70) we obtain

$$\tilde{\boldsymbol{\omega}}_\mu = \gamma^{-1} \left(\partial_\mu + \tilde{\boldsymbol{\omega}}_\mu\right) \gamma, \tag{3.65}$$

which is just the covariant generalization of the Maurer-Cartan form $\gamma^{-1}\partial_\mu\gamma$, and transforms non-linearly under (2.63),

$$g(x) : \tilde{\boldsymbol{\omega}}_\mu \mapsto h(\pi, g)\, \boldsymbol{\omega}_\mu\, h^{-1}(\pi, g) + h(\pi, g)\, \partial_\mu h^{-1}(\pi, g), \tag{3.66}$$

with $h(\pi, g)$ defined in equation (3.9). Therefore, if we expand this connection in the basis of the Lie algebra,

$$\tilde{\boldsymbol{\omega}}_\mu \equiv i \left[\tilde{D}_{\mu m}\, x^m + \tilde{E}_{\mu i}\, t^i\right], \tag{3.67}$$

we obtain covariant derivatives $\tilde{\boldsymbol{\mathcal{D}}}_a \equiv \tilde{e}^\mu{}_a \tilde{\boldsymbol{D}}_\mu$ and gauge fields $\tilde{\boldsymbol{\mathcal{E}}}_a \equiv \tilde{e}^\mu{}_a \tilde{\boldsymbol{E}}_\mu$ that transform like in equations (3.19), but with $x' = x$. The curvature tensor $\tilde{\boldsymbol{R}}_{\mu\nu}$ associated with $\tilde{\boldsymbol{\omega}}_\mu$ is in fact given by

$$\tilde{\boldsymbol{R}}_{\mu\nu} = \gamma^{-1} \boldsymbol{R}_{\mu\nu} \gamma, \tag{3.68}$$

where $\boldsymbol{R}_{\mu\nu}$ is the curvature tensor associated with the spin connection $\boldsymbol{\omega}_\mu$, derived itself from $e_\mu{}^a$. Under the action of elements $g(x) \in L_+^\uparrow$ on the vierbein (2.63), this curvature transforms non-linearly too,

$$g(x) : \tilde{\boldsymbol{R}}_{\mu\nu} \mapsto h(\pi, g)\, \tilde{\boldsymbol{R}}_{\mu\nu}\, h^{-1}(\pi, g). \tag{3.69}$$

It is thus clear from the transformation properties of these new quantities that if the original action S is invariant under H, the new action $\tilde{S}$ defined in equation (3.64) will be invariant under $L_+^\uparrow$. In fact, we could have reversed the whole construction. We could have started by defining a modified vierbein $\tilde{e}_\mu{}^a$, a modified covariant derivative $\tilde{\boldsymbol{D}}_\mu$ and a modified curvature tensor $\tilde{\boldsymbol{R}}_{\mu\nu}$ according to equations (3.61), (3.67) and (3.68). Then, any invariant action under H, solely constructed out of these ingredients would have been automatically and manifestly invariant under $L_+^\uparrow$.

3.4 Unbroken Rotations

We turn now our attention to cases in which the unbroken group is the rotation group, $H = SO(3)$, which is the maximal compact subgroup of $L_+^\uparrow$. This pattern of symmetry breaking is analogous to the spontaneous breaking of chiral invariance in the two quark model. In the latter, the chiral symmetry of QCD with two massless quarks, $SU(2)_L \times SU(2)_R$, is broken down to the isospin subgroup $SU(2)$, while in the former, the Lorentz group $SO(1,3) \sim SU(2) \times SU(2)$ is broken down to the diagonal subgroup of rotations $SO(3) \sim SU(2)$. Hence, the construction of rotationally invariant Lagrangians with broken Lorentz invariance is formally analogous to the construction of isospin invariant Lagrangians with broken chiral symmetry.

As in the two-quark model, the case for unbroken rotations can be motivated phenomenologically. If rotations were broken, we would expect the expansion of the universe to be anisotropic, in conflict with observations, which are consistent with a nearly isotropic cosmic expansion all the way from the initial stages of inflation. Our main goal here however is not to consider the phenomenology of theories with unbroken rotations, as this has been already extensively studied, but simply to illustrate how our formalism applies to theories with gravity. We shall see in particular how in this case our construction directly leads to the well-known Einstein-aether theories, which we show to be the most general class of theories in which rotations remain unbroken.

3.4.1 Coset Construction

In order to build the most general theory in which the rotation group remains unbroken, let us assume first that spacetime is flat, as in Section 3.2.2. In the case at hand, then, the generators of the unbroken group are the generators of rotations J_i, and the remaining "broken" generators are the boosts K_m. Therefore, the theory contains three Goldstone bosons π_m. Of particular relevance are the transformation properties of these Goldstone bosons under rotations. For an infinitesimal rotation

$t = \omega^i J_i$, equations (3.6) and (3.11) lead to

$$t : \pi_m \mapsto \pi'_m = \pi_m + (\omega \times \pi)_m \,. \tag{3.70}$$

In addition, since $P K^m P^{-1} = -K^m$ and $T K^m T^{-1} = -K^m$ we have, from (3.25) that $\pi_m \to -\pi_m$ under parity and time reversal. Therefore, the set of Goldstone bosons transform like a 3-vector under spatial rotations. These are analogous to the pions of spontaneously broken chiral invariance.

The restriction of the four-vector representation $\Lambda(g)$ to the subgroup of rotations H is reducible, $\mathbf{4} = \mathbf{1} \oplus \mathbf{3}$, so the tensor product representation of the rotation group in equation (3.20) is also reducible,

$$(\mathbf{1} \oplus \mathbf{3}) \otimes \mathbf{3} = \mathbf{3} \oplus \mathbf{1} \oplus \mathbf{3} \oplus \mathbf{5}. \tag{3.71}$$

(The different representations of the rotation group are labeled by their dimension. The dimension N of the representation is $N = 2S + 1$, where S is the spin of the representation.) More precisely, the covariant derivative

$$\mathcal{D}_m \equiv \mathcal{D}_{0m} \tag{3.72}$$

transforms like a spatial vector under rotations (spin one, $\mathbf{3}$), while D_{mn} transforms in the tensor product representation of rotations $\mathbf{3} \otimes \mathbf{3}$. Defining

$$\mathcal{D}_{mn} = \frac{1}{3}\, \varphi\, \delta_{mn} + a_{mn} + s_{mn}, \tag{3.73}$$

with a antisymmetric and s symmetric and traceless, leads to a scalar φ (spin zero, $\mathbf{1}$), a vector $a_{mn} \equiv \epsilon_{mnp} a^p$ (spin one, $\mathbf{3}$), and a traceless symmetric tensor s_{mn} (spin two, $\mathbf{5}$). Therefore, the most general Lagrangian density at most quadratic in the covariant derivatives, and invariant under the *full* Lorentz group is

$$\mathcal{L}_\pi = \frac{1}{2} \left(F_\varphi \, \varphi^2 + F_D \mathcal{D}_m \mathcal{D}^m + F_a \, a_{mn} a^{mn} + F_s \, s_{mn} s^{mn} \right), \tag{3.74}$$

where indices are raised with the (inverse) metric of Euclidean space, δ^{mn}. Note that we have omitted a linear term proportional to φ, and the parity-violating expression $\epsilon_{mnp} \, a^{mn} \mathcal{D}^p$ in the Lagrangian. As we show below, these terms are just total derivatives.

Let us now address the new ingredients that gravity introduces into the theory. As we discussed in Section 3.3.1, in a generally covariant theory we may choose to work in unitary gauge, in which the Goldstone bosons identically vanish. In this gauge, the covariant derivatives $\mathcal{D}_{am}$ defined above simply reduce to the spin connection along the appropriate generators, as in equations (3.56). Therefore, using the explicit form of the rotation generators in the fundamental representation, and $\mathrm{tr}(x^m_{(4)} \cdot x^n_{(4)}) = -2\delta^{mn}$, we find

$$\mathcal{D}_m = \omega_{0m0}, \quad \mathcal{D}_{mn} = \omega_{mn0}. \tag{3.75}$$

Recall that there are three broken generators which transform like vectors under rotations, which we label by m, n, and that the derivatives defined in equations (3.59) transform in the same way as the covariant derivatives defined in equation (3.18), with $x' = x$. Therefore, the Lagrangian (3.74) already contains all the rotationally invariant terms constructed from the undifferentiated spin connection.

To complete the most general gravitational action invariant under general coordinate and local Lorentz transformations, with at most two derivatives acting on the vierbein, we just need to add all invariant terms that can be constructed from the curvature alone. Without loss of generality, we may restrict ourselves to the components of the Riemann tensor in an orthonormal frame, $\mathcal{R}_{ab}{}^{cd}$. Then, indices along spatial direction transform like vectors, while indices along the time direction transform like scalars under rotations. Most of the invariants one can construct out of the Riemann tensor vanish because of antisymmetry. For instance, the term $\mathcal{R}_0{}^{mnp}\epsilon_{mnp}$ is identically zero because of the antisymmetry of the curvature tensor in the last three indices. In addition, the identity $[\nabla_\mu, \nabla_\nu]A^\rho = R_{\mu\nu}{}^\rho{}_\sigma A^\sigma$, in an orthonormal frame and up to boundary terms, implies the relation

$$\int d^4V \left[\mathcal{R}_{0m}{}^{0m} - \mathcal{D}_{mn}\mathcal{D}^{mn} + (\mathcal{D}_m{}^m)^2\right] = 0, \tag{3.76}$$

which can be used to eliminate a scalar term proportional to $\mathcal{R}_{0m}{}^{0m}$ from the action. As we mentioned earlier a term linear in the covariant derivative, $\varphi \equiv \mathcal{D}_m{}^m$, is a total derivative, since from equations (2.67) and (2.69)

$$\omega_{m0}{}^m = \partial_\mu e^\mu{}_0 + \Gamma^\mu{}_{\nu\mu} e^\nu{}_0 = \frac{1}{\det e}\partial_\mu(\det e\, e^\mu{}_0). \tag{3.77}$$

Similarly, one can show that $\epsilon_{mnp}\, a^{mn}\mathcal{D}^p$ is a total derivative too, since the latter equals $\epsilon^{mnpq}\nabla_m A_n \nabla_p A_q$, for $A_m = \delta_m{}^0$. We therefore conclude that the most general diffeomorphism invariant action invariant under local rotations is

$$S = \frac{M_P^2}{2}\int d^4V\,[\mathcal{R} + \mathcal{L}_\pi] + S_M, \tag{3.78}$$

where $\mathcal{R} \equiv \mathcal{R}_{ab}{}^{ab}$ is the Ricci scalar, the "Goldstone" Lagrangian $\mathcal{L}_\pi$ is given by equation (3.74), and S_M denotes the matter action. Tests of the equivalence principle [64] and constraints on Lorentz-violating couplings in the standard model [76] suggest that any Lorentz-violating term in the matter action S_M is very small. Hence, for phenomenological reasons, we assume that the breaking of Lorentz invariance is restricted to the gravitational sector. Therefore, S_M is taken to be invariant under Lorentz transformations, and the action (3.78) defines a metric theory of gravity.

3.4.2 The Einstein-aether

For unbroken rotations, the matrix γ that we introduced in Section 3.2.2 is a boost, $\gamma = \exp(i\pi_m K^m)$. Hence, instead of characterizing the Goldstone bosons by the set of three scalars π_m, we may simply describe them by the transformation matrix $\Lambda^a{}_0$ of the boost itself. The latter has four components,

$$u^a \equiv \Lambda^a{}_0, \tag{3.79}$$

but not all of them are independent, because Lorentz transformations preserve the Minkowski metric. In particular, the vector field u_a has unit norm

$$u_a u^a \equiv \eta_{ab}\Lambda^a{}_0\Lambda^b{}_0 = \eta_{00} = -1. \tag{3.80}$$

In the conventional approach to the formulation of the most general theory in which rotations remain unbroken, one would solve the constraint (3.80) by introducing an appropriate set of three parameters, and then identify their transformation properties under the Lorentz group [117]. One would then proceed to define covariant derivatives of these parameters, and use them to construct the most general theory compatible with the unbroken symmetry, just as we did.

In this case however, a simpler approach leads to the same general theory, but avoids introducing coset parametrizations and covariant derivatives altogether. Since the Lorentz transformation of a boost can be described by a the vector field (3.79), one may simply expect that the problem of constructing the most low-energy effective theory in which the rotation group remains unbroken just reduces to the problem of writing down the most general diffeomorphism invariant theory with the least numbers of derivatives acting on a unit norm vector field. This was precisely the problem that Jacobson and Mattingly studied in [114], which resulted in what they called the "Einstein-aether". The most general action in this class of theories is

$$S = \frac{M_G^2}{2} \int d^4V \Big[\mathcal{R} - c_1 \nabla_a u_b \nabla^a u^b - c_2 (\nabla_a u^a)^2 - c_3 \nabla_a u_b \nabla^b u^a + \\ + c_4 u^a u^b \nabla_a u_c \nabla_b u^c + \lambda (u_a u^a + 1) \Big], \quad (3.81)$$

where the parameters c_i are constant, and we have written down all the components of the "aether" vector field u^μ in an orthonormal frame, $u^a \equiv e_\mu{}^a u_\mu$, with covariant derivatives given by

$$\nabla_a u^b \equiv e^\mu{}_a \left(\partial_\mu u^b + \omega_\mu{}^b{}_c u^c \right). \quad (3.82)$$

The constraint $u^a u_a = -1$ on the norm of the field is enforced by the Lagrange multiplier λ. Hence, the action (3.81) is analogous to the linear σ-model in which chiral symmetry breaking was originally studied. In this formulation, the Lorentz group acts linearly on the vector field u^a, and, as we shall see, the fixed-norm constraint can be understood as limit in which the potential responsible for Lorentz symmetry breaking is infinitely steep around its minimum.

To establish the connection between the Einstein-aether (3.81) and the rotationally invariant action (3.78), we simply need to impose unitary gauge. We can solve the unit norm constraint in (3.81) by expressing the vector field u^a as a Lorentz transformation acting on an appropriately chosen vector $\tilde{u}^a$,

$$u^a = \Lambda^a{}_b(\pi) \tilde{u}^b, \quad \text{with} \quad \tilde{u}^a = \delta^a{}_0, \quad (3.83)$$

which is just a restatement of equation (3.79). Then, invariance under local Lorentz transformations implies that the aether action (3.81) can be equally thought of as a functional of $\tilde{u}^b$ and the transformed vierbein $\tilde{e}_\mu{}^a = (\Lambda^{-1}(\pi))^a{}_b\, e_\mu{}^b$. If we now redefine the vierbein field, $\tilde{e}_\mu{}^b \to e_\mu{}^a$, the Goldstone bosons π disappear from the action, and we are left with the theory in unitary gauge. In this gauge the vierbein is arbitrary, but (dropping the tildes) we can assume that $u^a = \delta^a{}_0$. In that case equation (3.82) gives in addition $\nabla_a u_b = \omega_{ab0}$, which, when substituted into the Einstein-aether action (3.81) precisely yields the action (3.78). The corresponding parameters M_P and F_i are expressed in terms of five linearly independent combinations of aether parameters,

$$M_P = M_G, \quad F_\varphi = -\frac{1}{3}(c_1+3c_2+c_3), \quad F_D = c_1+c_4, \quad F_a = c_3-c_1, \quad F_s = -(c_1+c_3), \tag{3.84}$$

and, therefore, the Einstein-aether is the most general low-energy theory in which the rotation group remains unbroken. The correspondence (3.84) also explains then why these particular combinations of the Einstein-aether parameters enter the predictions of the theory. In our language, they map into the different irreducible representations in which one can classify the covariant derivatives of the Goldstone bosons. The phenomenology of Einstein-aether theories is nicely reviewed in [75].

3.4.3 General Vector Field Models

In Einstein-aether theories, Lorentz invariance is broken because the vector field u^a develops a time-like vacuum expectation value. In this context, it is then natural to consider generic vector field theories in which a vector field develops a non-zero expectation value, and to study how the latter reduce to the Einstein-aether in the limit of low energies. This will also help us to illustrate our formalism in cases in which the spectrum of excitations contains a massive field, and how the latter disappears from the low-energy predictions of the theory.

The most general low energy effective action for a vector field non-minimally coupled to gravity which contains at most two derivatives and is invariant under local

Lorentz transformations and general coordinate transformations reads

$$S=\frac{1}{2}\int d^4V\Big[M_G^2\mathcal{R}+\frac{\alpha}{2}\,F_{ab}F^{ab}+\beta\,(\nabla_a A^a)^2+\,\beta_4\mathcal{R}A_aA^a+\beta_5\mathcal{R}_{ab}A^aA^b+ \qquad (3.85)$$

$$+\frac{A^aA^b}{\Lambda^2}\left(\alpha_1\,\nabla_aA_c\nabla_bA^c+\alpha_2\,\nabla_cA_a\nabla^cA_b+\alpha_3\,\nabla_aA_b\nabla_cA^c\right)+$$

$$+\gamma\,\frac{A^aA^bA^cA^d}{\Lambda^4}\nabla_aA_b\nabla_cA_d+\delta_1A_bA^b\nabla_aA^a-\Lambda^4\,V\Big].$$

Here, $F_{ab}\equiv\partial_aA_b-\partial_bA_a$, A^a are the components of the vector field in an arbitrary orthonormal frame, and the various coefficients α, α_i, β, β_i, γ, δ_1 and V should be regarded as arbitrary (dimensionless) functions of A_aA^a/Λ^2. Finally, M_G and Λ are the two characteristic energy scales of the effective theory, which is valid at energies $E\ll\min(\Lambda,M_G)$. In order to generate spontaneous breaking of Lorentz symmetry down to rotations we assume, without loss of generality, that the potential V is minimized by field configurations with $A_aA^a=-\Lambda^2$. Other low energy terms that do not appear in the expression (3.85) can be reduced to linear combinations of the terms above after integrations by parts. An action very similar to (3.85) has been already considered in [115], though the latter did not include the terms proportional to β_4 and δ_1, and all the other couplings were assumed to be constants rather than arbitrary functions of A^a. Models involving fewer terms have been studied for instance in [150–153] under the name of "bumblebee models," and in [137] under the name of "unleashed aether models."

In order to make contact with the formalism developed in the previous sections, we shall parametrize again the vector field as a Lorentz transformation acting on

$$A^a(x)=\delta^a_0\left(\Lambda+\sigma(x)\right), \qquad (3.86)$$

where the field σ is just a singlet under rotations. This is the same we did for the aether, the only difference being that there the fixed-norm constraint forced the field σ to vanish. As before, invariance under local Lorentz transformations then implies that the vector field can be taken to be given by (3.86). In this unitary gauge, the covariant derivative of A^a is

$$\nabla_aA^b=\delta^b_0(e^\mu{}_a\partial_\mu\sigma)+\eta^{bm}(\Lambda+\sigma)\,\mathcal{D}_{am}, \qquad (3.87)$$

where we have used equations (3.75). Thus, the action (3.85) can be expressed in terms of rotationally invariant operators that solely involve $\mathcal{R}_{ab}{}^{cd}$, $\mathcal{D}_{am}$, the scalar σ and its covariant derivative $\mathscr{D}_a\sigma = e^\mu{}_a\partial_\mu\sigma$.

It shall prove to be useful to expand the action (3.85) in powers of σ. To quadratic order, and to leading order in derivatives, the results is

$$S = \frac{1}{2}\int d^4V\Big[(M_G^2 - \bar{\beta}_4\Lambda^2)\mathcal{R} + \Lambda^2\left(\bar{\beta} + \frac{2\bar{\beta}_5}{3}\right)\varphi^2 + \Lambda^2(\bar{\alpha}_1 - \bar{\alpha})\mathcal{D}_m\mathcal{D}^m - \Lambda^2\bar{\beta}_5\, s_{mn}s^{mn}$$
$$+\Lambda^2(2\bar{\alpha} + \bar{\beta}_5)\, a_{mn}a^{mn} + \sigma(-2\bar{\delta}_1\Lambda^2\varphi + \cdots) + \sigma^2(-2\bar{V}''\Lambda^2 + \cdots) + \mathcal{O}(\sigma^3)\Big], \quad (3.88)$$

where the dots stands for the subleading terms in the derivative expansion and $\bar{V}''$ denotes the second derivative of the potential function with respect to its argument, evaluated at its minimum, where $A_aA^a = -\Lambda^2$. Similarly, $\bar{\alpha}, \bar{\beta}, \bar{\beta}_4, \bar{\beta}_5, \bar{\alpha}_1$ and $\bar{\delta}_1$ stand for the values of the couplings at the minimum of the potential. Apart from the additional rotationally invariant terms involving the field σ, the action (3.88) has manifestly the form (3.78) with $M_P^2 \equiv (1 - \bar{\beta}_4)M_G^2$.

We study the spectrum of this class of theories in Appendix 3.A. Their scalar sector consists of a massless excitation, one of the Goldstone bosons, and a massive field, whose mass is linear in $\bar{V}''$. We show in the appendix that in the low-momentum limit, the field σ has a vanishing matrix element between the massless scalar particle and the vacuum,

$$\lim_{p\to 0}\langle m = 0|\sigma(p)|0\rangle = 0. \quad (3.89)$$

Hence, if we are interested in low momenta and massless excitations, the field σ can be simply integrated out. At tree level, this can be easily done by solving the classical equations of motion to express σ in terms of the covariant derivatives $\mathcal{D}_{am}$. From (3.88), we see that to lowest order in derivatives the result is completely determined by the two terms proportional to σ^2 and $\sigma\varphi$. Thus, solving the corresponding linear equation,

$$\sigma = -\frac{\bar{\delta}_1^2}{2\bar{V}''}\varphi + \mathcal{O}(\partial^2/\Lambda), \quad (3.90)$$

and plugging back into the action (3.88) we get, to leading order in derivatives,

$$S = \frac{1}{2}\int d^4V\left[(M_G^2-\bar\beta_4\Lambda^2)\mathcal{R}+\Lambda^2\left(\bar\beta+\frac{\bar\delta_1^2}{2\bar V''}+\frac{2\bar\beta_5}{3}\right)\varphi^2+\Lambda^2(\bar\alpha_1-\bar\alpha)\mathcal{D}_m\mathcal{D}^m+\right.$$
$$\left.+\Lambda^2(2\bar\alpha+\bar\beta_5)\,a_{mn}a^{mn}-\Lambda^2\bar\beta_5\,s_{mn}s^{mn}\right]. \quad (3.91)$$

As expected the low energy action (3.91) has the form of (3.78). Integrating out the field sigma has simply renormalized the coefficients of the low energy theory, which are now given by

$$M_P^2 = M_G^2-\bar\beta_4\Lambda^2, \qquad F_\varphi=\left(\bar\beta+\frac{\bar\delta_1^2}{2\bar V''}+\frac{2\bar\beta_5}{3}\right)\frac{\Lambda^2}{M_P^2}, \qquad F_D=(\bar\alpha_1-\bar\alpha)\frac{\Lambda^2}{M_P^2},$$
$$F_a=(2\bar\alpha+\bar\beta_5)\frac{\Lambda^2}{M_P^2}, \qquad F_s=-\bar\beta_5\frac{\Lambda^2}{M_P^2}. \quad (3.92)$$

By combining these relations with equations (3.84), one can easily derive the dispersion relations and residues of the massless excitations in the model (3.85) from the known aether theory results [75]. Equations (3.92) show from the very beginning that the couplings γ, α_2 and α_3 will not enter the low-energy phenomenology. A "brute force" calculation based on the action (3.85) tends to obscure this fact, as shown explicitly in Appendix 3.A, although the final results are of course identical.

Alternatively, if we are interested only in the low energy phenomenology of the theory, we can choose to drop the field σ from the onset, as massive excitations will not give any observable contribution at low energies [143]. In the limit $\bar V''\to\infty$ where the massive mode becomes infinitely heavy, the potential may be replaced by a fixed-norm constraint, as in Einstein-aether theories. In fact, when $\bar V''\to\infty$, equation (3.90) implies that σ can be simply set to zero, and the general class of vector field models described by (3.85) directly reduces to the Einstein-aether. After introducing a rescaled vector $A^a\equiv\Lambda u^a$ and integrating some terms by parts, the coefficients c_i in (3.81) can be easily mapped onto the couplings in (3.85) as follows:

$$\alpha=-c_1\frac{M_G^2}{\Lambda^2}, \quad \beta=-(c_1+c_2+c_3)\frac{M_G^2}{\Lambda^2}, \quad \beta_5=(c_1+c_3)\frac{M_G^2}{\Lambda^2}, \quad \alpha_1=c_4\frac{M_G^2}{\Lambda^2},$$
$$\alpha_2=\alpha_3=\beta_4=\gamma=\delta_1=0. \quad (3.93)$$

Once again, equations (3.93) can be easily combined with the known Einstein-aether results [75] to immediately obtain the dispersion relations and the residues for the massless propagating modes in the specific model (3.85).

3.5 Summary

In this chapter we have generalized the effective Lagrangian construction of Callan, Coleman, Wess and Zumino to the Lorentz group. In flat spacetime, the Lorentz group is a global symmetry, and its breaking implies the existence of Goldstone bosons, one for each broken Lorentz generator. The broken global symmetry is not lost, and is realized non-linearly in the transformation properties of these Goldstone bosons and the matter fields of the theory. Because the Lorentz group is a spacetime symmetry, the Goldstone bosons transform non-trivially under the Lorentz group, and can be classified in linear representations of the unbroken subgroup. The same non-linearly realized global symmetry prevents the Goldstone bosons from entering the Lagrangian undifferentiated, which allows us to identify them as massless excitations. Because spacetime derivatives transform non-trivially under the Lorentz group, the covariant derivatives of Goldstone bosons typically furnish reducible representations of the unbroken Lorentz subgroup. The Lorentz group does not seem to be broken in the standard model sector, so any eventual breaking of this symmetry must be confined to a hidden sector of the theory. In that respect, phenomenologically realistic theories must resemble models of gravity-mediated supersymmetry breaking [154–156]. In both cases, a spacetime symmetry is broken in a hidden sector, the breaking is communicated to the standard model by the gravitational interactions, and, for phenomenological reasons, the symmetry breaking scale has to be sufficiently low.

Given an internal symmetry group, one always has a choice to make it global or local. But in the case of the Lorentz group this choice does not seem to exist. Any generally covariant theory that contains spinor fields, such as the standard model coupled to GR, requires that Lorentz transformations be an internal local symmetry, very much like a group of internal gauge symmetries. We have therefore extended the construction of actions in which global Lorentz invariance is broken to generally covariant formulations in which the group of local Lorentz transformations is non-linearly realized on the fields of the theory, which at the very least must contain the covariant derivatives of the Goldstone bosons and the vierbein, which describes the gravitational field. But in this case, since the Lorentz group is a local symmetry, it is possible and simpler to work in a formulation in which the Goldstone bosons are absent, and Lorentz symmetry is explicitly broken. In this "unitary gauge," the theory remains generally covariant, but Lorentz symmetry is lost. Even though the lost invariance under the Lorentz group can always be restored by introducing the appropriate Goldstone bosons, this restored symmetry is merely an artifact.

Generally covariant theories with broken Lorentz invariance differ significantly from their fully symmetric counterparts. In unitary gauge for instance, the covariant derivatives of the Goldstone bosons that the unbroken symmetry allows us to write down simply become the spin connection along the broken generators. This is just the Higgs mechanism. But in a generally covariant theory without extraneous additional fields, this connection is expressed in terms of the vierbein, so these terms actually represent kinetic terms for some of its components. Thus, instead of a massive theory of gravity, when Lorentz invariance is broken we obtain a theory with additional massless excitations (in Minkowski spacetime), which we can interpret as extra graviton polarizations in unitary gauge, or simply as the Goldstone bosons of the theory in general.

We have illustrated these issues for cases in which the rotation group remains unbroken. In particular, we have rigorously shown that the most general low-energy effective theory with unbroken spatial rotations is the Einstein-aether, and how generic vector field theories reduce to the latter at low energies.

The construction of low-energy effective theories that we have described here provides us with a tool to explore Lorentz symmetry breaking systematically and in a model-independent way. It identifies first how the Lorentz group acts on the field of the theory, it removes the clutter of particular models by focusing on the relevant fields at low energies, and it uniquely enumerates all the invariants under the unbroken symmetries.

Appendix 3.A Vector-Tensor EFTs

In this appendix we study the spectrum of excitations in the vector-tensor theories introduced in Section 3.4.3, in which Lorentz symmetry is broken down to rotations. Although such a study is usually carried out in the standard metric formulation (see for example [115]), in what follows we adopt instead the vierbein formalism which we already used in the main body of this chapter.

3.A.1 Perturbations

Our starting point is the action (3.85), which is a functional of the vierbein ${e_\mu}^a$ and the vector field A^a, and describes the behavior of both light and heavy modes. Perturbations of the vierbein around the Minkowski solution ${e_\mu}^a = {\delta_\mu}^a$ can be decomposed into scalars, vectors and tensors under spatial rotations as follows:

$$\delta {e_0}^0 = \phi, \tag{3.94a}$$

$$\delta {e_0}^i = \partial_i B + S_i, \tag{3.94b}$$

$$\delta {e_i}^0 = -\partial_i C - T_i, \tag{3.94c}$$

$$\delta {e_i}^j = -\delta_{ij}\psi + \partial_i\partial_j E + \epsilon_{ijk}\partial^k D - \partial_{(i}F_{j)} + \epsilon_{ijk}W^k + \frac{1}{2}h_{ij}. \tag{3.94d}$$

In this decomposition ϕ, B, C, ψ, E, D are scalars, S_i, T_i, F_i, W_i are transverse vectors, $\partial_i S^i = \cdots = \partial_i W^i = 0$, and h_{ij} is a transverse and traceless tensor, ${h_i}^i = \partial_i h^{ij} = 0$. Here, $i = 1, 2, 3$ labels spatial indices, which we raise and lower with the flat metric δ^{ij}.

Scalars, vectors and tensors transform in different irreducible representations of the rotation group and therefore do not couple from each other in the free theory. As we show in Section 3.4.3, no matter what the spacetime background is, we can always use invariance under local boosts to impose the "unitary gauge" condition (3.86), namely

$$A^a(x) = \delta^a_0 \left(\Lambda + \sigma(x)\right). \tag{3.95}$$

The field σ is a scalar under rotations.

Gauge fixing

At this point, not all the scalars and vectors in equations (3.94) and (3.95) describe independent degrees of freedom, because of the residual gauge invariance associated with general coordinate transformations and the unbroken group of local rotations. In fact, under infinitesimal coordinate transformations ($x^\mu \to x^\mu + \xi^\mu$) and local Lorentz rotations ($e_i{}^\mu \to e_i{}^\mu + \omega^k \epsilon^{ij}{}_k e_j{}^\mu$) the fluctuations of the vierbein around a Minkowski background (3.94) transform in the following way:

$$\delta e_0{}^0 \quad \to \quad \delta e_0{}^0 - \partial_t \xi^0, \tag{3.96a}$$

$$\delta e_0{}^i \quad \to \quad \delta e_0{}^i - \partial_t \partial^i \xi - \partial_t \xi_T^i, \tag{3.96b}$$

$$\delta e_i{}^0 \quad \to \quad \delta e_i{}^0 - \partial_i \xi^0, \tag{3.96c}$$

$$\delta e_i{}^j \quad \to \quad \delta e_i{}^j - \partial_i \partial^j \xi - \partial_i \xi_T^j + \epsilon_i{}^{jk} \partial_k \omega + \epsilon_i{}^j{}_k \omega_T^k, \tag{3.96d}$$

where we have decomposed ξ^μ and ω^i into the scalars ξ^0, ξ, ω and the transverse vectors ξ_T^i and ω_T^i ($\partial_i \xi_T^i = \partial_i \omega_T^i = 0$). Comparison of equations (3.94) and (3.96) then shows that, by performing an appropriately chosen rotation together with a general coordinate transformation, one can set for instance $F_i = W_i = 0$ and $C = D = E = 0 = 0$. Thus, we are eventually left with only four scalars (ϕ, B, ψ and σ), two vectors (S_i and T_i) and one tensor (h_{ij}). This is the same number of degrees of freedom one obtains in the metric formulation of the theory, after completely fixing the gauge.

3.A.2 Tensor Sector

As we mention above, in the free theory, scalars, vectors and tensors decouple from each other. Let us therefore start by considering the tensor sector, which is described by the quadratic Lagrangian

$$\mathcal{L}_t = \frac{1}{4} \left\{ \left[M_G^2 - \left(\bar{\beta}_4 + \bar{\beta}_5 \right) \Lambda^2 \right] \dot{h}_{ij} \dot{h}_{ij} - \left[M_G^2 - \bar{\beta}_4 \Lambda^2 \right] \partial_k h_{ij} \partial_k h_{ij} \right\}, \tag{3.97}$$

from which we can immediately read off the residue and the speed of sound of the tensor modes,

$$Z_t^{-1} = \frac{M_G^2 - \left(\bar{\beta}_4 + \bar{\beta}_5\right)\Lambda^2}{2}, \qquad c_t^2 = \frac{M_G^2 - \bar{\beta}_4\Lambda^2}{M_G^2 - (\bar{\beta}_4 + \bar{\beta}_5)\Lambda^2}. \tag{3.98}$$

Once again, $\bar{\beta}_4$ and $\bar{\beta}_5$ stand for the values of the couplings at the minimum of the potential, and a similar notation applies in what follows to the other couplings too. The tensor sector is ghost free provided $(\bar{\beta}_4 + \bar{\beta}_5) \ll (M_G/\Lambda)^2$. We should also impose $\bar{\beta}_4 \ll (M_G/\Lambda)^2$ in order to ensure classical stability. The results (3.98) agree with the ones of aether models with parameters given by equation (3.92), and they also reduce to the ones found by Gripaios [115] in the limit where $\Lambda \ll M_G$.

3.A.3 Vector Sector

The Lagrangian for the vector modes is only slightly more complicated, and reads

$$\begin{aligned}\mathcal{L}_v \;=\; & \frac{1}{2}\Big\{\left[M_G^2 - \bar{\beta}_4\Lambda^2\right]\partial_i(T_j + S_j)\partial_i(T_j + S_j) + 2(\bar{\alpha}_1 - \bar{\alpha})\Lambda^2\,\dot{T}_i\dot{T}_i + \\ & +(\bar{\beta}_5 + 2\bar{\alpha})\Lambda^2\,\partial_i T_j\partial_i T_j - \bar{\beta}_5\Lambda^2\,\partial_i S_j\partial_i S_j\Big\}.\end{aligned} \tag{3.99}$$

The field S_i only appears in the Lagrangian density through the combination $\partial_i S_j$ and does not propagate. Its equation of motion can be easily solved to get

$$\left[M_G^2 - (\bar{\beta}_4 + \bar{\beta}_5)\Lambda^2\right] S_i = -\left[M_G^2 - \bar{\beta}_4\Lambda^2\right] T_i, \tag{3.100}$$

which, when substituted back in (3.99) gives

$$\mathcal{L}_v = (\bar{\alpha}_1 - \bar{\alpha})\Lambda^2\,\dot{T}_i\dot{T}_i + \left(\bar{\alpha} - \frac{\bar{\beta}_5^2\Lambda^2}{2\left[M_G^2 - (\bar{\beta}_4 + \bar{\beta}_5)\Lambda^2\right]}\right)\partial_i T_j\partial_i T_j. \tag{3.101}$$

Therefore, only two massless vector modes propagate, with residue and a speed of sound given by

$$Z_v^{-1} = 2(\bar{\alpha}_1 - \bar{\alpha})\Lambda^2, \qquad c_v^2 = \frac{1}{\bar{\alpha} - \bar{\alpha}_1}\left(\bar{\alpha} - \frac{\bar{\beta}_5^2\Lambda^2}{2\left[M_G^2 - (\bar{\beta}_4 + \bar{\beta}_5)\Lambda^2\right]}\right). \tag{3.102}$$

In empty space, the vector sector of GR is non-dynamical. However, the breakdown of Lorentz invariance gives dynamics to this sector, even in the absence of matter fields.

Of course, these two vector modes correspond to two of the Goldstone bosons of the "spontaneously broken" phase. They are well behaved in the limit $\Lambda \ll M_G$ provided $(\bar{\alpha}_1 - \bar{\alpha}) > 0$ and $\bar{\alpha} < 0$. Notice that this result does not agree with [115], though it does agree with the result found in aether theories [75], upon the identification in equations (3.92).

3.A.4 Scalar Sector

Let us finally consider the scalar sector, which now contains both massive and massless fields. To quadratic order in the perturbations, its Lagrangian density is given by

$$\begin{aligned}
\mathcal{L}_s \;=\; & \frac{1}{2}\Big\{2(M_G^2 - \bar{\beta}_4\Lambda^2)(\partial_i\psi\partial_i\psi - 2\partial_i\phi\partial_i\psi) + \bar{\beta}\Lambda^2(\Delta B)^2 \\
& +(3\bar{\beta}\Lambda^2 + 2\bar{\beta}_5\Lambda^2 + 2\bar{\beta}_4\Lambda^2 - 2M_G^2)(3\dot{\psi}^2 + 2\dot{\psi}\Delta B) \\
& +(\bar{\alpha}_1 - \bar{\alpha})\Lambda^2\partial_i\phi\,\partial_i\phi + (\bar{\beta} - \bar{\alpha}_1 - \bar{\alpha}_2 - \bar{\alpha}_3 + \bar{\gamma})\dot{\sigma}^2 - (\bar{\alpha} - \bar{\alpha}_2)\partial_i\sigma\partial_i\sigma + \\
& -2\bar{V}''\Lambda^2\sigma^2 + [(-4\bar{\beta}_4 + 4\bar{\beta}_4' - 2\bar{\beta}_5 + 2\bar{\beta}_5' + \bar{\alpha}_3 - 2\bar{\beta})\dot{\sigma} + 2\bar{\delta}_1\Lambda\sigma](\Delta B + 3\dot{\psi}) + \\
& -\partial_i\sigma\partial_i[(4\bar{\beta}_4 - 4\bar{\beta}_4' + 2\bar{\beta}_5 - 2\bar{\beta}_5' + 2\bar{\alpha})\phi - 8(\bar{\beta}_4 - \bar{\beta}_4')\psi]\Big\}.
\end{aligned} \tag{3.103}$$

The scalars ϕ and B only appear in the Lagrangian trough the combinations $\partial_i\phi$ and ΔB, so they can be easily eliminated by solving their classical equations of motion. At this point, it is more convenient to switch to Fourier space, and write the action for the two remaining scalars in the form

$$S_s = -\frac{1}{2}\int d^4k\, X^\dagger D X, \quad \text{with} \qquad X \equiv \begin{pmatrix} \sigma(k) \\ \psi(k) \end{pmatrix} \tag{3.104}$$

and

$$D \equiv \begin{pmatrix} a_1\omega^2 + a_2k^2 + a_3\Lambda^2 & a_4\omega^2 + a_5k^2 + ia_6\Lambda\,\omega \\ a_4\omega^2 + a_5k^2 - ia_6\Lambda\,\omega & a_7\omega^2 + a_8k^2 \end{pmatrix}. \tag{3.105}$$

Here, the (dimension two) coefficients a_i are some complicated functions of the various coupling constants of the model. In particular, a_3 and a_6 are the only couplings that break the $\mathbb{Z}_2$ symmetry $A^a \to -A^a$.

The inverse of the matrix D is just the field propagator. In order to find the propagating modes we just have to find the values of ω^2 at which its eigenvalues have poles, or, equivalently, the values of ω^2 at which the eigenvalues of D have zeros. Requiring that $\det(D)$ vanish we thus arrive at the frequencies of the two propagating modes,

$$\omega_1^2 = m_1^2\Lambda^2 + c_1^2 k^2 + \mathcal{O}(k^4/\Lambda^2), \qquad \omega_2^2 = c_2^2 k^2 + \mathcal{O}(k^4/\Lambda^2), \tag{3.106}$$

with

$$m_1^2 = \frac{a_6^2 - a_3 a_7}{a_1 a_7 - a_4^2}, \tag{3.107a}$$

$$c_1^2 = \frac{a_8(a_3 a_4^2 - a_1 a_6^2) + (a_6^2 - a_3 a_7)(2a_4 a_5 - a_2 a_7)}{(a_6^2 - a_3 a_7)(a_1 a_7 - a_4^2)}, \tag{3.107b}$$

$$c_2^2 = \frac{a_3 a_8}{a_6^2 - a_3 a_7}. \tag{3.107c}$$

In the absence of fine-tuning, the first mode has a mass of order Λ and can be excluded from the low-energy theory. On the other hand, the speed of sound of the massless mode,

$$c_2^2 = \frac{(2\bar{V}''\bar{\beta} + \bar{\delta}_1^2)\left[2M_G^2 - (2\bar{\beta}_4 - \bar{\alpha} + \bar{\alpha}_1)\Lambda^2\right]\left[M_G^2 - \bar{\beta}_4\Lambda^2\right]}{(\bar{\alpha} - \bar{\alpha}_1)\left[M_G^2 - (\bar{\beta}_4 - \bar{\beta}_5)\Lambda^2\right]\left[2\bar{V}''\left(2M_G^2 - (2\bar{\beta}_4 + 2\bar{\beta}_5 + 3\bar{\beta})\Lambda^2\right) - 3\,\bar{\delta}_1^2\Lambda^2\right]}, \tag{3.108}$$

coincides with the speed of sound of the scalar mode in aether theories [75], after substitution of equations (3.92). Note that the terms $\mathcal{O}(k^4/\Lambda^2)$ in equation (3.106) cannot be trusted since our starting point was an effective action in which all the terms with more than two derivatives were excluded.

As in the vector sector, in the absence of matter fields the scalar sector of GR is non-dynamical. But again, the breakdown of Lorentz invariance gives dynamics to this sector. This captures of course the existence of a Goldstone boson in the scalar sector of the theory, which, together with the two massless modes we found in the vector sector, play the role of the three Goldstone bosons associated with the broken boost generators.

The residues of the scalar modes can be determined using the general result [137]

$$\frac{1}{Z_{1,2}} = -\frac{1}{\mathrm{tr}(D)}\,\frac{\partial}{\partial\omega^2}\det(D)\bigg|_{\omega^2=\omega_{1,2}^2}, \tag{3.109}$$

which, in our case, yields

$$Z_1^{-1} = \frac{a_6^2(a_1+a_7) - a_3(a_4^2+a_7^2)}{(a_1a_7 - a_4^2)(a_6^2 - a_3a_7)} + \mathcal{O}(k^4/\Lambda^2), \qquad Z_2^{-1} = \frac{a_3}{a_3a_7 - a_6^2} + \mathcal{O}(k^4/\Lambda^2). \tag{3.110}$$

Like for the speed of sound, the residue of the massless mode

$$Z_2^{-1} = \frac{2\left[M_G^2 - (\bar\beta_4+\bar\beta_5)\Lambda^2\right]\left[3\bar\delta_1^2\Lambda^2 - 2\bar V''(2M_G^2 - (2\bar\beta_4+2\bar\beta_5+3\bar\beta)\Lambda^2)\right]}{(\bar\delta_1^2 + 2\bar V''\bar\beta)\Lambda^2} + \mathcal{O}(k^4/\Lambda^2) \tag{3.111}$$

agrees with that obtained in aether theories [75], upon the identification (3.92). Once again, the terms $\mathcal{O}(k^4/\Lambda^2)$ in the residues are out of the reach of validity of the effective theory we wrote down.

To conclude, it is interesting to point out that none of the results concerning the massless modes depend on α_2, α_3, γ, nor on the derivatives of β_4 and β_5. A brute-force approach like the one we just followed makes this look like the result of accidental cancellations. Notice for instance that in fact the free scalar Lagrangian (3.103) does depend on α_2, α_3, γ, as well as on the derivatives of β_4 and β_5. The low-energy effective action (3.91) on the other hand makes this manifest from the very beginning.

3.A.5 The field σ

We obtained the low energy effective Lagrangian (3.91) by integrating out the field σ. In that context, we claimed that this procedure was justified because that the matrix element of σ between the vacuum and a state with one massless particle vanishes in the low-momentum limit (see equation (3.89)). We are now in a position to prove this result.

As we have seen above, the scalar spectrum consists of a massive field s_1 and a massless field s_2. We can thus express the field σ as a linear combination of the two canonically normalized fields,

$$\sigma = \kappa_1 s_1 + \kappa_2 s_2, \tag{3.112}$$

in which κ_1 and κ_2 are momentum-dependent coefficients. Therefore, using the reduction formula, the matrix element for emission of a massless excitation in equation

(3.89) can be written as

$$\langle m=0,p|\sigma(p')|0\rangle = \lim_{\omega\to\omega_2} i\,(\omega_2^2-\omega^2)\langle s_2(p)\sigma(p')\rangle_T =$$
$$= i\kappa_2 \lim_{\omega\to\omega_2}(\omega_2^2-\omega^2)\langle s_2(p)s_2(p')\rangle_T = \delta(p+p')\,\kappa_2, \tag{3.113}$$

where $p=(\omega,k)$, the energy ω_2 was defined in equation (3.106), and $\langle f(p)g(p')\rangle_T$ is the Fourier transform of the corresponding Green's function. The value of κ_2 can be readily calculated by noting that

$$-i\delta(p+p')D_{\sigma\sigma}^{-1}(p) = \langle\sigma(p)\sigma(p')\rangle_T = \kappa_1^2\langle s_1(p)s_1(p')\rangle_T + \kappa_2^2\langle s_2(p)s_2(p')\rangle_T \tag{3.114}$$
$$= \delta(p+p')\left(\frac{i\kappa_1^2}{\omega^2-\omega_1^2}+\frac{i\kappa_2^2}{\omega^2-\omega_2^2}\right). \tag{3.115}$$

Hence,

$$\kappa_2^2 = \lim_{\omega\to\omega_2}(\omega_2^2-\omega^2)\,D_{\sigma\sigma}^{-1} = \frac{a_6^2\,a_8}{(a_3a_7-a_6^2)^2}\frac{k^2}{\Lambda^2}+\mathcal{O}(k^4/\Lambda^4), \tag{3.116}$$

which clearly shows that κ_2 vanishes in the low-momentum limit.

Chapter 4

Scalar-Tensor Theories of Gravity and WEP violations

4.1 Introduction

Einstein based the development of GR on two pillars: general covariance and the equivalence principle. Since then, physicists have often wondered whether there are any alternatives to GR, which, while preserving its theoretical framework and phenomenological successes also avoid some of the shortcomings sometimes attributed to it. Among the phenomenological successes of GR, the equivalence principle—the proportionality of inertial and gravitational mass—is the most accurately tested and constrained one. Indeed, experiments at the University of Washington limit the relative difference in acceleration towards the earth of two test spheres of different atomic compositions to be less than one part in 10^{12} [157]. Therefore, any putative alternative theory of gravitation has to pass the significant hurdle of the equivalence principle.

Arguably, the simplest way to modify GR is to add a scalar field to the gravitational sector. Since gravitation is a long-ranged interaction, such a scalar would have to be sufficiently light to be considered part of the gravitational field. Whereas it is straightforward to include such a scalar field while preserving diffeomorphism invariance, the most general diffeomorphism-invariant theory with a light scalar would generically lead to strong violations of the equivalence principle [158]. There is nevertheless a subclass of scalar-tensor theories which respect the weak form of the equivalence principle, at least at tree level, and thus provides a natural class of phenomenologically viable alternatives to GR. (As shown by Nordtvedt [159], theories in this class do violate the strong equivalence principle, although these violations are negligible in laboratory-sized experiments.) The first such theory was proposed by Pascual Jordan [160], after criticism by Fierz [161] of an earlier proposal of the former [162]. Essentially the same theory was later revived by Brans and Dicke [63], whose names are usually associated with the class of scalar-tensor theories we study here. Further extensions and generalizations within this class were later considered by different authors [163].

What distinguishes these weak equivalence principle-preserving scalar-tensor theories is the existence of a formulation of the theory—a conformal frame—in which the scalar field only couples to gravity (at tree level). It follows then, by construction, that these theories preserve the weak equivalence principle classically, since their matter sector is the same as that of GR. Of course, the question is what happens to the equivalence principle when quantum fluctuations are turned on, and, more generally, whether quantum corrections preserve the structure of this subclass of scalar-tensor theories. This is not just a purely academic question, because even Planck-suppressed interactions eventually generated by loops would lead to departures from the weak equivalence principle that are experimentally ruled out. The question is most conveniently addressed in the Einstein conformal frame of these theories, in which the propagators of the graviton and the scalar are diagonal. Although in this frame the scalar couples directly to matter, it is easy to check that the equivalence principle is preserved at tree level. However, because the field couples directly to matter, it is hard to see why quantum corrections would not lead to violations of the equivalence principle.

The impact of quantum corrections on the equivalence principle has been the subject of a small but interesting debate in the literature. In the first article on the topic we were able to find, Fujii argued that quantum corrections should violate the equivalence principle [164]. He explicitly calculated one-loop quantum corrections to the vertex for scalar emission by a photon, with matter fields running inside the loop, and argued that the latter do seem to violate the equivalence principle. But somewhat later the same author realized that this purported violation disappears if one employs dimensional regularization instead of a cut-off [165]. Unaware of these results, Cho also argued that the equivalence principle should be violated in scalar-tensor theories [166], though some of his arguments seemed to be in conflict with the explicit calculation performed by Fujii in [165]. Up to that point, whether or why quantum corrections preserve the equivalence principle remained unclear, to say the least. Recently, Hui and Nicolis have shed more light on the issue by providing explicit examples for *massive* fields showing that matter loops do not lead to violations of the weak equivalence principle in scalar-tensor theories [167]. They argue that this is due to the linear coupling of the scalar to the trace of the energy momentum tensor: Because the energy-momentum tensor is conserved, the scalar couples to a charge density given by the time-time component of the energy-momentum tensor, which they identify with the mass density.

In this chapter, which is based on the paper [168], we extend these arguments further. As we shall see, the equivalence principle in scalar-tensor theories has a two-fold origin: A broken Weyl symmetry that relates the couplings of the scalar to those of the graviton, and diffeomorphism invariance, which significantly constrains the couplings of the graviton (and demands in particular that the latter couple to a conserved quantity, the energy-momentum tensor.) Diffeomorphism invariance implies that in the limit of zero momentum transfer the vertex for graviton emission by matter—the gravitational mass—has to be proportional to the inertial mass [169–171], and it is the broken Weyl symmetry what makes the couplings to the scalar inherit that property. Moreover, because this Weyl symmetry is broken, there is a corresponding Ward identity for the broken symmetry that exactly predicts the size of those quantum corrections that violate the equivalence principle: They have to be proportional to three inverse powers of the gravitational couplings.

4.2 Formalism

4.2.1 Action Principle

The scalar-tensor theories we are about to study are characterized by the existence of a conformal frame, the Jordan frame, in which bosonic matter is minimally coupled to the spacetime metric. Out of all possible scalar-tensor theories, this restriction singles out a very specific class of theories in which the weak equivalence principle holds, at least classically.

By definition, the gravitational sector of any scalar-tensor theory consists of a scalar ϕ and a rank two symmetric tensor $g_{\mu\nu}$, the metric. Our universe contains fermionic fields however, and this conventional formulation has to be replaced by one in terms of the scalar ϕ and the vierbein $e_\mu{}^a$ (see [146] for a review.) In this language, the scalar-tensor theories we consider here have an action functional

$$S_J = \int d^d x \det e \left[F(\phi)R - G(\phi)\partial_\mu\phi\partial^\mu\phi - W(\phi)\right] + S_M^J[e_\mu{}^a, \psi_\alpha], \tag{4.1}$$

where the index J denotes Jordan frame quantities, ψ_α is a set of matter fields and R is the Ricci scalar associated with the metric

$$g_{\mu\nu} = \eta_{ab} e_\mu{}^a e_\nu{}^b. \tag{4.2}$$

Note that we work in an arbitrary number of dimensions d, and that matter is now minimally coupled to the vierbein field, which is what singles out the class of theories we consider in this chapter. To some extent the dynamics of the gravitational sector are unimportant; our considerations can be easily generalized to even more general forms of the gravitational sector of the action.

The action (4.1) is invariant under two symmetry groups: diffeomorphisms and local Lorentz transformations. Because any spacetime tensor can be converted into a diffeomorphism scalar by contraction with the vierbein, we can assume that all matter fields are diffeomorphism scalars. In that case, under infinitesimal diffeomorphisms

$x^\mu \to x'^\mu = x^\mu + \xi^\mu(x)$ the fields of the theory transform according to

$$e_\mu{}^a \to e'_\mu{}^a = e_\mu{}^a + \Delta e_\mu{}^a, \qquad \Delta e_\mu{}^a = -\xi^\nu \partial_\nu e_\mu{}^a - e_\nu{}^a \partial_\mu \xi^\nu, \tag{4.3a}$$

$$\phi \to \phi' = \phi + \Delta\phi, \qquad \Delta\phi = -\xi^\mu \partial_\mu \phi, \tag{4.3b}$$

$$\psi_\alpha \to \psi'_\alpha = \psi_\alpha + \Delta\psi_\alpha, \qquad \Delta\psi_\alpha = -\xi^\mu \partial_\mu \psi_\alpha. \tag{4.3c}$$

Under local Lorentz transformations $\Lambda(x) \in SO(1,3)$ the different fields transform in the corresponding representation of the Lorentz group,

$$e_\mu{}^a \to e'_\mu{}^a = \Lambda^a{}_b \, e_\mu{}^b, \tag{4.4a}$$

$$\phi \to \phi' = \phi, \tag{4.4b}$$

$$\psi_\alpha \to \psi'_\alpha = D(\Lambda)_\alpha{}^\beta \psi_\beta, \tag{4.4c}$$

where D is the linear representation of the Lorentz group under which the matter fields transform.

Our goal is to investigate the gravitational interactions experienced by the different matter fields. In the quantum theory these interactions are mediated by the interchange of gravitons and scalar particles. However, in the action (4.1) the graviton and scalar propagators are typically not diagonal. Hence, it is convenient and customary to introduce a new set of variables in terms of which the propagators become diagonal. This set of new variables define what is usually known as the Einstein frame, in which the action reads

$$S_E = S_{EH}[e_\mu{}^a] + S_\phi[e_\mu{}^a, \phi] + S_M^E[f(\phi/M) e_\mu{}^a, \psi_\alpha], \tag{4.5a}$$

where

$$S_{EH} = \int d^d x \det e \, \frac{M_P^2}{2} R, \quad S_\phi = \int d^d x \det e \left[-\frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi) \right]. \tag{4.5b}$$

Of course, the choice of conformal frame is a matter of convenience, and both (4.1) and (4.5) are physically equivalent, as recognized early on by Dicke [172] (in the quantum theory, the equivalence follows from the invariance of S-matrix elements under field redefinitions.) For convenience and simplicity we take however (4.5) as the starting point of our considerations. We also assume that the equations of motion admit a solution with constant value $\bar{\phi}$ of the scalar field, which for simplicity we take to be $\bar{\phi} = 0$, and at this minimum we define

$$m_\phi^2 \equiv \left. \frac{d^2 V}{d\phi^2} \right|_{\bar{\phi}=0} . \tag{4.6}$$

If both $f(0)$ and $f'(0)$ differ from zero, we may assume without loss of generality the normalization conditions

$$f(0) = 1, \quad f'(0) = 1. \tag{4.7}$$

Note that in d spacetime dimensions, M_P and M do not have mass dimension one. Instead they have the same mass dimension as the scalar and the graviton.

4.2.2 The Weak Equivalence Principle

Recall that the weak equivalence principle states that in a gravitational field all neutral test bodies fall with the same acceleration, or, more simply, that gravitational and inertial mass are proportional to each other. To see how the equivalence principle emerges in the classical theory defined by the action (4.5a), consider the tree-level diagram in figure 4.1.1, in which two different matter particles scatter through scalar exchange on a Minkowski background. The amplitude of the diagram in figure 4.1.1 is[1]

$$\mathcal{M}_\phi = -\frac{1}{(2\pi)^{3d-1}} [u_\beta^\dagger(p'_A) \gamma_\phi^{\beta\alpha} u_\alpha(p_A)] \, \frac{1}{q^2 + m_\phi^2} \, [u_\beta^\dagger(p'_B) \gamma_\phi^{\beta\alpha} u_\alpha(p_B)], \tag{4.8}$$

where $\gamma_\phi^{\beta\alpha}$ is the tree-level amplitude for scalar emission by matter, the $u_\alpha(p)$ are the appropriate mode functions for the external particles, $q \equiv p'_A - p_A$ is the momentum transfer, and $(q^2 + m_\phi^2)^{-1}$ is the scalar propagator. We are interested here in the potential energy between two static bodies, that is, on a scalar whose four-momentum q^μ approaches zero: $p'_A \to p_A$, $p'_B \to p_B$.

[1] We mostly follow the conventions of [54]. In these conventions, the propagator carries a factor of $(2\pi)^{-d}$, each external line contributes a factor of $(2\pi)^{-(d-1)/2}$, and the relation between S-matrix elements and the amplitudes $\mathcal{M}$ for an initial state i and a final state f is $S_{fi} = \delta_{fi} - 2\pi i \delta(p_f - p_i) \mathcal{M}_{fi}$. See the next subsections for additional information on our conventions for vertices and propagators.

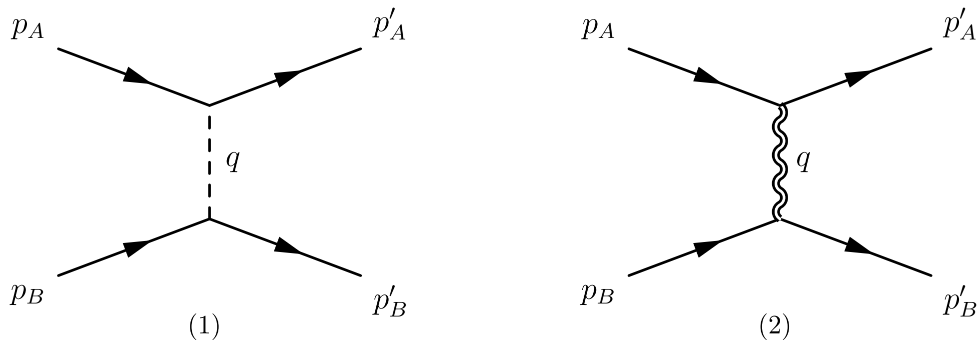


Figure 4.1: Scalar and graviton exchange between two different matter species. Continuous lines denote matter fields (bosonic or fermionic), dashed lines label the scalar ϕ, and wiggly lines label the graviton.

Inspection of the way ϕ enters the action (4.5a) reveals that in flat spacetime the scalar vertex γ_ϕ is related to the vertex for graviton emission $(\gamma_h)^{\mu\nu}$ by

$$M\gamma_\phi^{\beta\alpha} = 2M_P(\gamma_h^{\beta\alpha})^\mu{}_\mu, \tag{4.9}$$

where we have used equation (4.7). The graviton vertex γ_h is proportional to the quadratic component of the energy momentum tensor in flat space, so equation (4.9) is just roughly the statement that the scalar couples to the trace of the energy-momentum tensor (the factor of two stems from the identification of the vierbein as "half" a graviton). As we shall see, it follows from diffeomorphism invariance alone that in momentum space, and in the limit of zero momentum transfer, this tree-level graviton vertex has to be of the form

$$2M_P(\gamma_h^{\beta\alpha})^{\mu\nu} = \pi^{\beta\alpha}(p)\,\eta^{\mu\nu} - p^{(\mu}\frac{\partial\pi^{\beta\alpha}}{\partial p_{\nu)}}, \tag{4.10}$$

where a parenthesis next to an index denotes symmetrization, p^μ is the momentum of matter, and $\pi^{\alpha\beta}$ is the tree-level self-energy, that is, minus the inverse of the tree-level propagator. The reader can easily verify this relation in the cases of a scalar, a spin half fermion and spin one vector. Hence, because of equation (4.9), an analogous relation applies for the amplitude for scalar emission,

$$M\gamma_\phi^{\beta\alpha} = d\,\pi^{\beta\alpha} - p^\mu\frac{\partial\pi^{\beta\alpha}}{\partial p^\mu}, \tag{4.11}$$

which, again, can be checked independently for scalars, spinors and vectors. On shell, the self-energy π vanishes by definition. Contracting then equation (4.11) with the appropriate mode functions we find for all three types of matter fields that, on shell,

$$u^\dagger_\beta \gamma_\phi^{\beta\alpha} u_\alpha = \frac{(2\pi)^d}{M}\frac{p^2}{p^0} = -\frac{(2\pi)^d}{M}\frac{m_I^2}{p^0}, \tag{4.12}$$

where m_I is the inertial mass of the particle, defined to be the value of $-p^2$ at the zero of the self-energy, and we have also used that for free fields of arbitrary spin

$$u^\dagger_\beta \frac{\partial \pi^{\beta\alpha}}{\partial p_\mu} u_\alpha = -(2\pi)^d \frac{p^\mu}{p^0}. \tag{4.13}$$

In particular, note that equation (4.12) implies that massless particles do not couple to the scalar at tree level, even if the field Lagrangian is not conformally invariant, as happens for instance for a massless scalar. Hence, the scalar interaction does not contribute to the bending of light, and the experimental constraints on the Eddington parameter γ thus demand that the scalar interaction be much weaker than gravity, $M_P \ll M$ [64]. Finally, substituting equation (4.12) into (4.8) and taking the limit of non-relativistic massive particles, $p^0 \approx m_I$, we arrive at

$$\mathcal{M}_\phi = -\frac{1}{(2\pi)^{d-1}}\frac{m_A m_B}{M^2}\frac{1}{q^2 + m_\phi^2}, \tag{4.14}$$

where m_A and m_B are, respectively, the inertial masses of particles A and B.

As we mentioned above, we want to calculate the potential energy for two static bodies, at fixed spatial distance $\vec{r}$ in $d = 4$ spacetime dimensions. To this end, we simply need to Fourier transform the non-relativistic limit of the scattering amplitude (4.14) back to real space. Since in the non-relativistic limit $q^2 = \vec{q}^{\,2}$, we obtain

$$V(\vec{r}) \equiv \int d^3q \, e^{i\vec{q}\cdot\vec{r}} \mathcal{M}_\phi(\vec{q}) = -\frac{m_A m_B}{M^2}\frac{e^{-m_\phi r}}{4\pi r}. \tag{4.15}$$

Hence, the force mediated by the scalar ϕ is proportional to the inertial mass, with a proportionality factor $1/M^2$ that is universal: The scalar interaction respects the weak equivalence principle.

Of course, if we calculate the potential energy due to graviton exchange, we also find that the latter respects the weak equivalence principle. As before, on shell, using equations (4.10) and (4.13) we find

$$u_\beta^\dagger (\gamma_h^{\beta\alpha})^{\mu\nu} u_\alpha = \frac{(2\pi)^d}{2M_P} \frac{p^\mu p^\nu}{p^0}. \tag{4.16}$$

Since with our conventions the graviton propagator in d spacetime dimensions (say, in de Donder gauge) is

$$i[\pi_h^{-1}(q)]_{\mu\nu,\rho\sigma} = \frac{-4i}{(2\pi)^d q^2} \left\{ \frac{1}{2} \left(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho} \right) - \frac{1}{d-2} \eta_{\mu\nu}\eta_{\rho\sigma} \right\} \tag{4.17}$$

we obtain in the non-relativistic limit that the amplitude associated with the diagram in figure 4.1.2 in $d = 4$ is

$$\mathcal{M}_h = -\frac{m_A m_B}{2M_P^2} \frac{1}{q^2}. \tag{4.18}$$

Again, the amplitude is proportional to the inertial masses of both particles, with a proportionality constant $1/M_P^2$ that is universal. The origin of this result is the tree-level Ward-Takahashi identity (4.10). The latter relates emission of a graviton—the gravitational mass—to the self-energy of matter—the inertial mass. It just so happens that, due to the structure of the matter action in (4.5a), the scalar couplings "inherit" this Ward identity, ultimately leading to the preservation of the weak equivalence principle in the scalar sector (at tree level). We explore whether these features survive in the quantum theory next.

4.2.3 Quantization

For the purpose of quantization, it shall prove to be useful to work with the quantum effective action Γ, the sum of all one-particle-irreducible (1PI) diagrams with a given number of external lines. In order to calculate the effective action, we expand the fields in quantum fluctuations around a given (but arbitrary) background. We thus write

$$e_\mu{}^a = \bar{e}_\mu{}^a + M_P^{-1} \delta e_\mu{}^a, \tag{4.19a}$$

$$\phi = \bar{\phi} + \delta\phi, \tag{4.19b}$$

$$\psi_\alpha = \bar{\psi}_\alpha + \delta\psi_\alpha, \tag{4.19c}$$

where overbars denote background values, and deltas quantum fluctuations. Plugging equation (4.19a) into (4.2) we find

$$g_{\mu\nu} \equiv \bar{g}_{\mu\nu} + M_P^{-1}\, h_{\mu\nu} + \mathcal{O}(M_P^{-2}), \quad \text{with} \quad \bar{g}_{\mu\nu} = \eta_{ab}\bar{e}_\mu{}^a \bar{e}_\nu{}^b, \quad h_{\mu\nu} = \delta e_{\mu\nu} + \delta e_{\nu\mu}, \tag{4.20}$$

and $\delta e_{\mu\nu} \equiv \bar{e}_{\nu a}\delta e_\mu{}^a$ (note that the location of the vierbein indices is important.) Hence, the symmetric part of the vierbein fluctuations, $h_{\mu\nu}$, is the graviton field; its antisymmetric part $a_{\mu\nu} \equiv \delta e_{\mu\nu} - \delta e_{\nu\mu}$ is non-dynamical [173]. It follows then by definition that[2]

$$\delta e_{\mu\nu} = \frac{h_{\mu\nu}}{2} + \frac{a_{\mu\nu}}{2}. \tag{4.21}$$

As in any non-abelian gauge theory, we quantize the theory defined by (4.5) using the functional integral formalism. Because the action (4.5) is invariant under two groups of local symmetries (diffeomorphisms and Lorentz transformations), we need to fix both gauges and introduce the corresponding ghost fields. Hence, our total action becomes

$$S_{\text{tot}} = S_E + S_{GF} + S_G, \tag{4.22}$$

where S_E is given in equation (4.5b), S_{GF} is the gauge-fixing term, and S_G the action for the ghosts. In the background field method, the gauge fixing term is such that the total action S_{tot} in equation (4.22) is invariant under a set of symmetries in which the background fields transform like the fields themselves, that is, under equations (4.3). For concreteness, and following [173], we impose the de Donder (harmonic) gauge condition to fix the diffeomorphism gauge, and an algebraic term to fix the Lorentz frame,

$$\begin{aligned} S_{GF} = -\frac{1}{4}\int d^dx \det\bar{e}\, \Bigg[\bar{g}^{\mu\nu}\left(\bar{\nabla}_\rho h^\rho{}_\mu - \frac{1}{2}\bar{\nabla}_\mu h^\rho{}_\rho\right)\left(\bar{\nabla}_\rho h^\rho{}_\nu - \frac{1}{2}\bar{\nabla}_\nu h^\rho{}_\rho\right) \\ + \bar{g}^{\mu\rho}\bar{g}^{\nu\sigma}\frac{a_{\mu\nu}a_{\rho\sigma}}{2M_P^2}\Bigg]. \end{aligned} \tag{4.23}$$

With this choice of gauge fixing, the action for the diffeomorphism ghosts ζ^μ and the Lorentz ghosts $\theta^{\mu\nu}$ becomes

$$S_G = -\frac{1}{\sqrt{2}}\int \det\bar{e}\left[\zeta^{\dagger\mu}\left(\bar{g}_{\mu\nu}\bar{\Box} - \bar{R}_{\mu\nu}\right)\zeta^\nu + \frac{M_P^2}{2}\bar{g}_{\mu\rho}\bar{g}_{\nu\sigma}\theta^{\dagger\mu\nu}\theta^{\rho\sigma}\right]. \tag{4.24}$$

[2] Roughly speaking, just as we think of the vierbein as the square root of the metric, we can think of a vierbein fluctuation as half a graviton.

We employ dimensional regularization, which preserves the gauge symmetries of the theory while rendering the theory finite. The effective action is the path integral over these fluctuations, with the prescribed values of the background fields kept fixed,

$$\exp(i\Gamma[\bar{e}_\mu{}^a, \bar{\phi}, \bar{\psi}_\alpha]) \equiv \int_{1PI} \mathcal{D}\delta e\, \mathcal{D}\delta\phi\, \mathcal{D}\delta\psi\, \mathcal{D}\zeta\, \mathcal{D}\theta\, \exp[iS_{\text{tot}}]. \tag{4.25}$$

The integral is restricted to run only over all one-particle-irreducible vacuum diagrams. The end result of this construction is that the effective action remains invariant under diffeomorphisms and Lorentz transformations, even though these symmetries had to be broken to define the path integral.

4.2.4 Gravitational Interactions

Consider now the scattering of two distinguishable particles described by the matter fields ψ_α (and their adjoints $\psi^\dagger_\alpha$ when appropriate). Restricting ourselves to interactions mediated by the vierbein and the scalar, these are determined by the two diagrams in figure 4.2, the counterparts of the two tree-level diagrams of figure 4.1. In real space, the 1PI vertices are given by functional derivatives of the quantum effective action evaluated at vanishing field fluctuations. In particular, in view of (4.20) and (4.21), the irreducible vertices for emission of a graviton and a scalar by matter are

$$\begin{aligned} (\Gamma_h^{\beta\alpha})^{\mu\nu}(z;y,x) &\equiv \frac{1}{2M_P}\frac{\delta^3\Gamma}{\delta\bar{\psi}_\alpha(x)\delta\bar{\psi}^\dagger_\beta(y)\delta\bar{e}_{(\mu}{}^a(z)}\bar{e}^{\nu)a}(z), \\ \Gamma_\phi^{\alpha\beta}(z;y,x) &\equiv \frac{\delta^3\Gamma}{\delta\bar{\psi}_\alpha(x)\delta\bar{\psi}^\dagger_\beta(y)\delta\bar{\phi}(z)}, \end{aligned} \tag{4.26}$$

while the self-energies of the graviton and the scalar (minus the inverse of their propagator) are given by[3]

$$(\Pi_h)^{\mu\nu,\rho\sigma}(y,x) \equiv \bar{e}^{(\rho a}(x)\frac{\delta^2\Gamma}{\delta\bar{e}_{\sigma)}{}^a(x)\delta\bar{e}_{(\mu}{}^b(y)}\bar{e}^{\nu b}(y), \quad \Pi_\phi(y,x) \equiv \frac{\delta^2\Gamma}{\delta\bar{\phi}(x)\delta\bar{\phi}(y)}. \tag{4.27}$$

[3]Because the effective action is diffeomorphism invariant by construction, we need to add to it an additional gauge fixing term to define the graviton propagator [174].

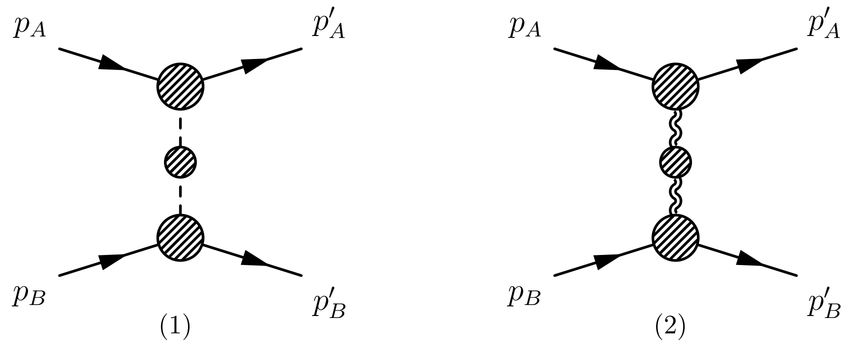


Figure 4.2: A light scalar (dashed) and a massless graviton mediate long-ranged interactions through the interchange of a single quantum. Each blob in a vertex represents the sum of all one-particle-irreducible diagrams (1PI) with the corresponding number of external lines, and external propagators stripped off. Each blob with two external lines represents the full propagator, the sum of all (connected) diagrams with the corresponding type of external lines.

These functional derivatives are evaluated in a Minkowski spacetime background with vanishing scalar and matter fields,

$$\bar{\phi} = 0, \quad \bar{\psi}_\alpha = 0, \quad \bar{e}_\mu{}^a = \delta_\mu{}^a, \tag{4.28}$$

though we do not make this explicit (it should be clear from the context.) If, aside from the vierbein, the background does not contain any Lorentz vectors, the variational derivative $\delta^2\Gamma/(\delta\bar{e}_\mu{}^a\delta\bar{\phi})$ vanishes as a consequence of Lorentz invariance. Therefore, there is no need to consider diagrams with one incoming scalar and one outgoing graviton. Note that the cubic vertices above describe the couplings of unrenormalized fields. To calculate physical scattering amplitudes we have to multiply these amplitudes with the appropriate wave function renormalization constants.

Scattering amplitudes are typically calculated in momentum space, so it is convenient to work with the momentum-space vertices and self-energies defined above. In our conventions, one of the vertex momenta is incoming (p_1), the other two (p_2 and p_3) are outgoing, and a momentum-conserving delta function has been split off,

$$\Gamma(p_2, p_1)\, \delta(p_1 - p_2 - p_3) \equiv \int d^dx\, d^dy\, d^dz\, \Gamma(z; y, x)\, e^{-ip_3 z} e^{-ip_2 y} e^{ip_1 x}. \tag{4.29}$$

In this way, the scattering amplitude is given by

$$\mathcal{M} = \frac{1}{(2\pi)^{2d-1}} \Big[\Gamma_\phi(p'_A, p_A)\Pi_\phi^{-1}(q)\Gamma_\phi(p'_B, p_B) + \Gamma_h(p'_A, p_A)\Pi_h^{-1}(q)\Gamma_h(p'_B, p_B) \Big], \tag{4.30}$$

where $q \equiv p'_A - p_A = p'_B - p_B$ is the momentum transfer and the $\Gamma's$ have been contracted with the appropriate mode functions for the matter fields, $\Gamma_f(p', p) \equiv u^\dagger_\beta(p')\Gamma^{\beta\alpha}_f(p', p)u_\alpha(p)$. The potential energy is determined by the values of the irreducible vertices and propagators at zero momentum transfer, $q = 0$. Using the definition of potential energy and the Fourier transform in (4.15), the gravitational potential in $d = 4$ becomes

$$V(r) \approx -\frac{1}{2(2\pi)^9}\left[\Gamma_\phi(p_A, p_A)\Gamma_\phi(p_B, p_B)\frac{Z_\phi e^{-m_\phi r}}{r} + \Gamma_h(p_A, p_A)\Gamma_h(p_B, p_B)\frac{Z_h}{2r}\right], \tag{4.31}$$

where we have used the spectral representation for the scalar propagator, and Z_ϕ and Z_h respectively are the residues of the scalar and graviton propagators. We assume that m_ϕ^{-1} is much larger than the scales r under consideration, so that we can think of the force mediated by ϕ effectively as a long-ranged interaction (we do not consider the Chameleon mechanism here [66].) What matters for our purposes is that the potential energy is determined by the vertices for scalar and graviton emission, and, therefore, the latter dictate the fate of the equivalence principle in the quantum theory.

4.3 Ward Identities

Because the quantum effective action is invariant under diffeomorphisms, it satisfies a set of Ward-Takahashi identities that relate the full vertex for graviton emission Γ_h to the full matter self-energy Π, as we shall derive next. These Ward identities are ultimately responsible for the validity of the equivalence principle in the quantum theory, as far as the couplings of matter to the graviton are concerned.

The origin of the Ward identity for graviton emission is that the vierbein transforms non-trivially under diffeomorphisms, even for a trivial vierbein background (flat spacetime.) This is why diffeomorphism invariance strongly restricts the couplings of matter to the graviton. In particular, it is possible to derive the weak equivalence principle in S-matrix theory solely from the requirement that S-matrix elements be invariant under diffeomorphisms acting on the polarization vectors of the graviton

[169].

The case of scalar emission however is quite different. The existence of a scalar field ϕ coupled to matter does not require nor entail any particular symmetry. In particular, because the change in the scalar field ϕ under diffeomorphisms vanishes at zero background field, diffeomorphisms have nothing to say about the couplings of the scalar to matter. This is why there is no a priori reason to expect that the couplings of the scalar field to matter respect the equivalence principle in the quantum theory. In fact they do not, as we also show further below. Nevertheless, because the scalar field only couples to matter in the combination $f(\phi/M)\, e_\mu{}^a$, its couplings inherit the Ward identity satisfied by the graviton to all orders in the matter coupling constants.

4.3.1 Graviton Emission

Our first goal is to derive the Ward identity for graviton emission. Such an identity was proven for arbitrary bosonic matter fields by DeWitt in [171], following the derivation in [170] for scalar matter. We basically extend here DeWitt's derivation to the vierbein formulation of the theory.

Let us consider the self-energy of the matter fields ψ_α in the presence of a background vierbein and a background scalar, and a vertex with an additional vierbein line,

$$\Pi^{\beta\alpha}(y,x) \equiv \frac{\delta^2\Gamma}{\delta\bar{\psi}_\alpha(x)\delta\bar{\psi}^\dagger_\beta(y)}, \qquad (\Gamma_e^{\beta\alpha})^\mu{}_a(z;y,x) \equiv \frac{1}{2M_P}\frac{\delta\Pi^{\beta\alpha}(y,x)}{\delta e_\mu{}^a(z)}. \tag{4.32}$$

Because the effective action is invariant under diffeomorphisms it does not change under the infinitesimal transformation (4.3),

$$\int d^d z \left[\frac{\delta\Gamma}{\delta\bar{e}_\mu{}^a(z)}\Delta\bar{e}_\mu{}^a(z) + \frac{\delta\Gamma}{\delta\bar{\phi}(z)}\Delta\bar{\phi}(z) + \Delta\bar{\psi}_\alpha(z)\frac{\delta\Gamma}{\delta\bar{\psi}_\alpha(z)}\right] = 0. \tag{4.33}$$

Therefore, acting on this equation with two functional derivatives with respect to the matter fields we obtain

$$\int d^d z \left[\frac{\delta\Pi^{\beta\alpha}(y,x)}{\delta\bar{e}_\mu{}^a(z)}\Delta\bar{e}_\mu{}^a(z) + \frac{\delta\Pi^{\beta\alpha}(y,x)}{\delta\bar{\phi}(z)}\Delta\bar{\phi}(z) \right. \\ \left. + \frac{\delta\Delta\bar{\psi}_\gamma(z)}{\delta\bar{\psi}_\alpha(x)}\Pi^{\beta\gamma}(y,z) + \frac{\delta\Delta\bar{\psi}^\dagger_\gamma(z)}{\delta\bar{\psi}^\dagger_\beta(y)}\Pi^{\gamma\alpha}(z,x)\right] = 0. \tag{4.34}$$

Using the transformation (4.3) and the definitions (4.32), and evaluating the last equation in our background (4.28) we then get

$$2M_P \int d^d z\, \delta_\nu{}^a \xi^\nu(z) \frac{\partial}{\partial z^\mu} (\Gamma_e^{\beta\alpha})^\mu{}_a(z; y, x) + \frac{\partial}{\partial y^\mu} \left[\xi^\mu(y) \Pi^{\beta\alpha}(y, x)\right] + \frac{\partial}{\partial x^\mu} \left[\xi^\mu(x) \Pi^{\beta\alpha}(y, x)\right] = 0. \tag{4.35}$$

In momentum space, with our momentum conventions (4.29), this becomes the identity

$$2M_P(p'_\mu - p_\mu)(\Gamma_e^{\beta\alpha})^\mu{}_\nu(p', p) = p'_\nu\, \Pi^{\beta\alpha}(p) - p_\nu\, \Pi^{\beta\alpha}(p'), \tag{4.36}$$

which in the limit of zero momentum transfer $p' \to p$ and after symmetrization reduces to the Ward-Takahashi identity for graviton emission,

$$2M_P(\Gamma_h^{\beta\alpha})^{\mu\nu}(p, p) = \Pi^{\beta\alpha}(p)\, \eta^{\mu\nu} - p^{(\mu} \frac{\partial \Pi^{\beta\alpha}}{\partial p_{\nu)}}. \tag{4.37}$$

An analogous identity holds in electromagnetism.

The self-energy is the sum of the tree-level contribution $\pi^{\beta\alpha}$ and the sum of all one-particle-irreducible self-energy diagrams[4] $\Delta\pi^{\beta\alpha}$,

$$\Pi^{\beta\alpha} = \pi^{\beta\alpha} + \Delta\pi^{\beta\alpha}. \tag{4.38}$$

It is convenient to work in renormalized perturbation theory, with fields whose self-energy corrections vanish on shell, and whose propagators have unit residue at the corresponding pole,

$$\Delta\pi^{\beta\alpha}\bigg|_{OS} = 0, \quad \frac{\partial \Delta\pi^{\beta\alpha}}{\partial p_\mu}\bigg|_{OS} = 0. \tag{4.39}$$

The irreducible vertex $(\Gamma_h^{\beta\alpha})^{\mu\nu}$ is also the sum of the tree contribution $(\gamma_h^{\beta\alpha})^{\mu\nu}$ and the contribution from loop diagrams $(\Delta\gamma_h^{\beta\alpha})^{\mu\nu}$,

$$(\Gamma_h^{\beta\alpha})^{\mu\nu} = (\gamma_h^{\beta\alpha})^{\mu\nu} + (\Delta\gamma_h^{\beta\alpha})^{\mu\nu}. \tag{4.40}$$

Because the Ward identity (4.37) is merely an expression of diffeomorphism invariance, it also holds in the limit in which all coupling constants of the theory go to

[4]For a scalar, $\pi \equiv -(2\pi)^d(p^2 + m^2)$, and $\Delta\pi = (2\pi)^d \pi^*$, where π^* is what is usually called the self-energy insertion [54].

zero, in which we can approximate all quantum amplitudes by tree-level expressions. Hence, the tree vertex and the tree-level self-energy obey the identity (4.10), the tree-level counterpart of equation (4.37), as the reader can explicitly check.

We are ready now to derive the main result of this subsection. Substituting equations (4.38) and (4.40) into the Ward-Takahashi identity (4.37), using the tree-level relation (4.10), and going on shell, equations (4.39), we conclude that

$$(\Delta\gamma_h^{\beta\alpha})^{\mu\nu}\Big|_{OS} = 0. \tag{4.41}$$

On shell, and in the limit of zero-momentum transfer, quantum corrections to the gravitational vertex vanish. Since, as we have seen in Subsection 4.2.2, tree-level (classical) amplitudes do respect the equivalence principle, so do the quantum corrected ones. As before, this result has an analogous counterpart in electromagnetism, which guarantees the non-renormalization of the electric charge (up to an overall wave function renormalization constant) at zero momentum transfer.

4.3.2 Scalar Emission

Let us turn our attention now to the emission of a scalar by matter. Although there is no analogous Ward identity for scalar emission, because of the structure of the couplings of ϕ to matter the vertex for scalar emission is closely related to that for graviton emission, whose properties it partially inherits. To see this, note that the matter action S_M^E in (4.5a) is invariant under the set of infinitesimal transformations

$$\psi_\alpha \to \psi'_\alpha = \psi_\alpha, \tag{4.42a}$$

$$\phi \to \phi' = \phi + \epsilon\, M, \tag{4.42b}$$

$$e_\mu{}^a \to e'_\mu{}^a = e_\mu{}^a - \epsilon\, \frac{f'}{f} e_\mu{}^a, \tag{4.42c}$$

where ϵ is an arbitrary function on spacetime. For certain functions $f(\phi/M)$, namely, exponentials, this transformation can be promoted to a group of $U(1)$ transformations that act on ϕ by a shift, and on the vierbein by a Weyl transformation. In that particular case, the transformations (4.42) are linear in the fields, though, in general, the transformation (4.42c) is non-linearly realized. Whatever the case, if (4.42) were an exact, linearly-realized, global symmetry of the full action we would get, plugging the transformation rules (4.42) into the general identity (4.34), and evaluating at our background (4.28)

$$\int d^d z \left[M \frac{\delta \Pi^{\beta\alpha}(y,x)}{\delta \bar{\phi}(z)} - \frac{f'(0)}{f(0)} \frac{\delta \Pi^{\beta\alpha}(y,x)}{\delta \bar{e}_\mu{}^a(z)} \delta_\mu{}^a \right] = 0, \tag{4.43}$$

where we have used that linear symmetries of the action are symmetries of the effective action. Using equations (4.26) and (4.27) this would lead immediately to the zero momentum identity

$$M\, \Gamma_\phi^{\beta\alpha}(p,p) = 2 M_P (\Gamma_h^{\beta\alpha})^\mu{}_\mu(p,p), \tag{4.44}$$

which relates the vertex for scalar emission to that for graviton emission. Since the latter satisfies equation (4.37) it would then follow in the limit of zero momentum transfer that

$$M\, \Gamma_\phi^{\beta\alpha}(p,p) = d\, \Pi^{\beta\alpha}(p) - p^\mu \frac{\partial \Pi^{\beta\alpha}}{\partial p^\mu}, \tag{4.45}$$

and, as in the graviton case, using the tree-level relation (4.11) this would finally yield

$$\Delta\gamma_\phi \bigg|_{OS} = 0, \tag{4.46}$$

which states that quantum corrections to the scalar vertex in the limit of zero-momentum transfer vanish. Since the tree-level scalar vertex does respect the equivalence principle, so would quantum corrections. Note that if, in addition, the transformation (4.42) were a local symmetry, we would be able to eliminate ϕ altogether from the theory by choosing the appropriate gauge.

But, of course, the full action is not invariant under the global transformation (4.42), and moreover, in general, the transformation (4.42) is non-linear. From this point of view, equation (4.43) is just an approximation to zeroth order in symmetry-breaking terms of a general Ward-Takahashi identity that we derive in Appendix 4.A. To apply the general Ward identity (4.129) to our case, consider a linear version of the Weyl transformation (4.42) acting on the field fluctuations,

$$\phi \to \phi' = \phi + \epsilon\, M, \tag{4.47a}$$

$$\delta e_\mu{}^a \to \delta e'_\mu{}^a = \delta e_\mu{}^a - \epsilon\, \frac{f'(0)}{f(0)}(\bar{e}_\mu{}^a + \delta e_\mu{}^a). \tag{4.47b}$$

Using the normalization conditions (4.7), substituting equations (4.47) into (4.129), and taking two functional derivatives with respect to the matter fields yields the analogue of equation (4.44), modulo corrections due to the fact that the transformation (4.42) is not an exact (linear) symmetry of the action,

$$M\,\Gamma_\phi^{\beta\alpha}(p,p) = 2M_P(\Gamma_h^{\beta\alpha})^\mu{}_\mu(p,p) + \Gamma_\Delta^{\beta\alpha}. \tag{4.48}$$

Here, as we detail in Appendix 4.A, $\Gamma_\Delta^{\alpha\beta}$ is the sum of all one-particle-irreducible diagrams with two external fields ψ_α and ψ_β (with amputated propagators), and a vertex insertion of Δ, the change of the Lagrangian density under the linear transformation (4.47), carrying zero momentum into the diagram.

In order to determine the explicit form of Δ we note that, from the action (4.5),

$$\frac{\delta S_\phi}{\delta \phi} = \det e \left[\Box\phi - \frac{dV}{d\phi}\right], \tag{4.49a}$$

$$e_\mu{}^a \frac{\delta S_\phi}{\delta e_\mu{}^a} = -\det e \left[\frac{d-2}{2}\partial_\mu\phi\partial^\mu\phi + d\,V(\phi)\right], \tag{4.49b}$$

$$e_\mu{}^a \frac{\delta S_{EH}}{\delta\, \delta e_\mu{}^a(z)} = \det e\, \frac{(d-2)\,M_P^2}{2} R, \tag{4.49c}$$

$$e_\mu{}^a \frac{\delta S_M}{\delta\, \delta e_\mu{}^a} \equiv \det e\, f^d\, T_M{}^\mu{}_\mu, \tag{4.49d}$$

$$\frac{\delta S_M}{\delta \phi} = \det e\, f^d \frac{f'}{M f}\, T_M{}^\mu{}_\mu, \tag{4.49e}$$

where R the scalar curvature and

$$(T_M)_\mu{}^\nu \equiv \frac{f\, e_\mu{}^a}{\det(f\, e)} \frac{\delta S_M}{\delta(f e_\nu{}^a)} \tag{4.50}$$

is the energy-momentum tensor of matter, which depends on ϕ because we assume that the matter action is of the form (4.5a). Hence, using equations (4.128), (4.47) and (4.7) we arrive at

$$\Delta = M \frac{\delta S_\phi}{\delta\phi(x)} - e_\mu{}^a \frac{\delta(S_\phi + \delta S_{EH} + \delta S_{GF})}{\delta\, \delta e_\mu{}^a(x)} + \det e\, f^d \left(\frac{f'}{f} - 1\right) T_M{}^\mu{}_\mu, \tag{4.51}$$

which we should expand around our background (4.28) in order to calculate the corresponding diagrams. The key of this result is that the correction term proportional to the energy-momentum tensor of matter is at least proportional to ϕ. This reflects that the transformation (4.47) leaves the part of the matter action proportional to $f(0)$ invariant.

A graphical representation of equation (4.48) to leading order in the gravitational couplings is given in figure 4.3, and helps to understand the different corrections due the broken Weyl symmetry. In our background, $\delta S_\phi/\delta\phi$ contains a linear term in ϕ, whose insertion in a vertex does not lead to any 1PI diagrams. The next contribution from $\delta S_\phi/\delta\phi$ stems from a term proportional to $\phi\, h/M_P$, and thus, the sum of all diagrams with an insertion of $M\,\delta S_\phi/\delta\phi$ and two external matter lines contributes a term of order M_P^{-2} to the equation in figure 4.3. (The diagram only has two external matter lines, so the scalar and graviton lines must end at a vertex in the diagram. Since the latter respectively couple with strength M^{-1} and M_P^{-1}, the suppression must be at least of order $M/M_P \times M^{-1} \times M_P^{-1} = M_P^{-2}$.) Similarly, because $e_\mu{}^a \delta S_\phi/\delta(\delta e_\mu{}^a)$ is at least quadratic in the scalar ϕ, insertion of this vertex yields a contribution of order M^{-2}, from the vertices at which the two scalar lines must end. In our background the variational derivative $e_\mu{}^a \delta S_{EH}/\delta(\delta e_\mu{}^a)$ is at least quadratic in the graviton field, and, therefore, the vertex containing (4.49c) yields a contribution of order M_P^{-2}, the same as that from the gauge fixing term, which is also quadratic in the graviton. Since the ghost action does not contain $h_{\mu\nu}$ nor ϕ, it is invariant under the transformation (4.47). Finally the vertex insertion proportional to $T^\mu_{M\mu}$ in equation (4.51) is linear in ϕ/M, and thus contributes a correction of order M^{-2} to the proper vertex, unless $f''(\bar\phi = 0) = 1$, for which this term would be proportional to $(\phi/M)^2$, and hence would contribute a factor of order M^{-3}. An extreme example of the latter is an exponential, for which the insertion proportional to $T^\mu_{M\mu}$ would be absent altogether. Overall, because of (4.37), this translates into the approximate scalar Ward-Takahashi identity

$$M\,\Gamma_\phi^{\beta\alpha}(p,p) = d\,\Pi^{\beta\alpha}(p) - p^\mu \frac{\partial \Pi^{\beta\alpha}}{\partial p^\mu} + \mathcal{O}(M_P{}^{-2}) + \mathcal{O}(M^{-2}). \tag{4.52}$$

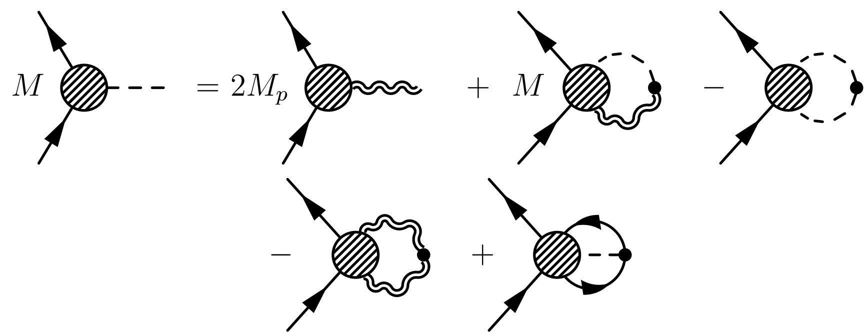


Figure 4.3: Diagrammatic expression of equation (4.48) to leading order in the gravitational couplings M_P^{-1} and M^{-1}. The irreducible vertex for scalar emission equals the trace of that for graviton emission plus or minus corrections terms. In each correction term, the blob represents the sum of all 1PI diagrams with the corresponding number of external lines and the vertex insertion marked by a dot.

Expanding the scalar vertex on the left hand side of the last equation into a tree-level contribution γ_ϕ and loop corrections $\Delta\gamma_\phi$, using equations (4.38) and (4.39) for the right hand side, and employing that the tree-level couplings of the scalar do respect the equivalence principle, equation (4.11), we thus finally get

$$\Delta\gamma_\phi\Big|_{OS} = \mathcal{O}(M^{-1}{M_P}^{-2}) + \mathcal{O}(M^{-3}). \tag{4.53}$$

Quantum corrections to scalar couplings do violate the equivalence principle, but by terms suppressed by three powers of the gravitational couplings. Since experimental constraints require $M_P \ll M$ [64], the dominant violations are of order $M^{-1}M_P^{-2}$. Although we have assumed for concreteness that the dynamics of the graviton and scalar fields is described by equation (4.5b), it is straightforward to extend our analysis to more general forms. As long as the latter do not preserve the Weyl symmetry (4.47), there should be violations of the weak equivalence principle in those theories too.

4.3.3 Extension of the Weyl Symmetry to the Full Action

We have previously noted that exponentials $f = \exp(\phi/M)$ play a special role in the action (4.5a), since for such functions the Weyl transformation (4.42) is a linearly realized, exact symmetry of the matter action, $\Delta S_M^E = 0$. In this case, the last term in equation (4.51) is absent, and the corresponding equivalence principle violating corrections to the scalar vertex proportional to $T^{\mu}_{M\mu}$ vanish. It is then natural to ask whether this Weyl symmetry can be extended to the rest of the action.

Consider first the scalar field action S_ϕ. To render it invariant under the transformation (4.42) we just need to interpret ϕ as the Goldstone boson of a spontaneously broken Weyl symmetry. In that case, a mass term is forbidden by the global shift symmetry $\phi \to \phi + \epsilon M$, and field derivatives need to enter with appropriate factors of $\exp(\phi/M)$,

$$\tilde{S}_\phi = -\frac{1}{2}\int d^dx \; \det e \; \exp\left[\frac{(d-2)\phi}{M}\right] g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi. \tag{4.54}$$

It is then easy to check then, that this new action is invariant under global Weyl transformations, $\Delta\tilde{S}_\phi = 0$. In such a theory, the correction terms in equation (4.51) coming from the change of S_ϕ under the Weyl transformation would vanish.

Along the same lines, we can also extend the Einstein-Hilbert action to a globally Weyl invariant expression,

$$\tilde{S}_{EH} = \int d^dx \det e \exp\left[\frac{(d-2)\phi}{M}\right] \frac{M_P^2}{2} R, \tag{4.55}$$

which, again remains invariant under (4.47), $\Delta\tilde{S}_{EH} = 0$. For such an action, the correction terms in (4.51) stemming from the change of $\tilde{S}_{EH}$ would again vanish.

However, we cannot make the full action Weyl invariant while keeping intact its scalar-tensor nature. In fact, if the total action reads

$$\tilde{S}_{\text{tot}} = \tilde{S}_{EH} + \tilde{S}_\phi + S_M\left[\exp(\phi/M){e_\mu}^a, \psi\right], \tag{4.56}$$

the field redefinition $\tilde{e}_\mu{}^a \equiv e^{\phi/M} e_\mu{}^a$ leads to

$$\tilde{S}_{\text{tot}} = \int d^dx \det\tilde{e}\left[\frac{M_P^2}{2}\tilde{R} - \frac{1}{2}\left(1 + (d-2)(d-1)\frac{M_P^2}{M^2}\right)\tilde{g}^{\mu\nu}\partial_\mu\phi\partial_\nu\phi\right] + S_M[\tilde{e}_\mu{}^a, \psi]. \tag{4.57}$$

This is just the action of GR minimally coupled to matter with an extended matter sector consisting of a minimally coupled massless scalar. Because there are no vertices with an odd power of ϕ in this theory, the amplitude for emission of a single scalar by matter vanishes (in any case, the scalar field couples derivatively, so it cannot mediate a long-ranged interaction.)

It is also instructive to consider the action (4.5) in flat space, with gravitation turned off ($M_P \to \infty$). Though the broken Weyl symmetry (4.47) acts non-trivially on the metric, this approximate symmetry does not get lost. Indeed, with all matter fields taken to be diffeomorphism scalars, in flat spacetime and for an exponential f the Weyl transformation (4.42c) has the same effect on the vierbein as the infinitesimal coordinate dilatation

$$x^\mu \to \left(1 - \epsilon \, \frac{f'(0)}{f(0)}\right) x^\mu. \tag{4.58}$$

Therefore, in that case, as a consequence of diffeomorphism invariance, the matter action in the Einstein frame possesses an exact dilatation symmetry under which the fields transform according to

$$\psi_\alpha \to \psi_\alpha + \epsilon \, x^\mu \partial_\mu \psi_\alpha, \tag{4.59a}$$

$$\phi \to \phi + \epsilon (M + x^\mu \partial_\mu \phi), \tag{4.59b}$$

where we have used the normalization conditions (4.7).

The dilatation (4.59a) does not act conventionally on the matter fields. To bring it to its usual form it is convenient to redefine the matter fields. Suppose that the kinetic term of the matter field ψ_α contains n derivatives. Then, diffeomorphism invariance implies that each derivative is accompanied by the inverse of the vierbein, and that the integration measure $d^d x$ is multiplied by $\det e$. Therefore, in the Einstein frame the kinetic term of the field ψ_α is proportional to f^{d-n}. Let us hence redefine

$$\tilde{\psi}_\alpha = f^{\frac{d-n}{2}} \psi_\alpha. \tag{4.60}$$

Then, by construction, the kinetic term of $\tilde{\psi}_\alpha$ does not contain factors of f (though there may be additional derivative interactions), and the matter action is invariant under

$$\tilde{\psi}_\alpha \to \tilde{\psi}_\alpha + \epsilon \left(\frac{d-n}{2} + x^\mu \partial_\mu\right) \tilde{\psi}_\alpha, \tag{4.61a}$$

$$\phi \to \phi + \epsilon (M + x^\mu \partial_\mu \phi), \tag{4.61b}$$

where we have used again equation (4.7). Acting on the matter fields, this is now a conventional dilatation, since $(d-n)/2$ is the scaling dimension of the field ψ_α. The inhomogeneous term in the transformation of ϕ underscores its interpretation as a pseudo Nambu-Goldstone boson of an approximate, spontaneously broken conformal symmetry, although, even for a massless ϕ, the scalar field action is not invariant under (4.61).

Because the field ϕ transforms inhomogeneously under (4.61), the vertex for scalar emission satisfies a Ward-Takahashi identity (4.129) related to this broken symmetry [175],

$$M\,\Gamma_\phi^{\beta\alpha}(p,p) + \left(p^\mu \frac{\partial}{\partial p^\mu} - n\right)\Pi^{\beta\alpha}(p) = \Gamma_\Delta^{\beta\alpha}, \tag{4.62}$$

where, again, $\Gamma_\Delta^{\alpha\beta}$ is the sum of all 1PI diagrams with two external ψ lines and a vertex insertion of Δ, the change in the Lagrangian density under the infinitesimal transformation (4.61). This equation is the flat space counterpart of equation (4.52), and also guarantees that, for fields renormalized on shell, quantum corrections to the vertex for scalar emission are determined by the change of the action under the broken symmetry (4.61).

The dilatation (4.58) is part of the conformal group, the set of all coordinate transformations that preserve the Minkowski metric up to an overall conformal factor. Along the same lines as for dilatations, as a consequence of diffeomorphism and Lorentz invariance, it is easy to show that, for an exponential f, the matter action in flat space is symmetric under the full conformal group, acting again on the scalar ϕ linearly, but inhomogeneously. There exist then additional Ward identities related to the full conformal symmetry of the theory, though we shall not write them down. Although conformal symmetries are typically anomalous, the structure of the couplings to ϕ in the matter action for an exponential f guarantees that the symmetry remains intact in the dimensionally regularized theory. If f is not an exponential, or if scalar kinetic term is not conformally invariant, conformal symmetry is broken, and the corresponding Ward identities contain the appropriate vertex insertions, as in the dilatation case.

4.4 Specific Examples

Our next goal is to illustrate our main results with a set of concrete examples that show the nature and size of the equivalence principle violations in scalar-tensor theories. We only address this issue for scalars and spin half fermions; the vertex for scalar emission by a gauge boson vanishes on shell as a consequence of Lorentz and gauge invariance, so there is no need to consider this case in the context of the weak equivalence principle.

We first check explicitly in a one-loop calculation that couplings to matter do not lead to any violations of the weak equivalence principle. For scalars, these results partially overlap and complement previous work in the literature [167]. In addition, we verify that one-loop corrections involving the scalar ϕ do result in violations of the equivalence principle, in agreement with the Ward identity (4.53). To avoid the complications of index algebra, we focus on loops of ϕ for simplicity, but we expect analogous violations from diagrams in which the loop contains at least one graviton.

4.4.1 Scalar Matter

Let us assume for the time being that matter consists of scalar particles χ, which for simplicity interact through a cubic coupling with another species of scalar particles σ. Then, in the Jordan frame, the matter action is

$$\mathcal{L}_M^J = -\frac{1}{2}g^{\mu\nu}\partial_\mu\chi\partial_\nu\chi - \frac{1}{2}m^2\chi^2 - \frac{1}{2}g^{\mu\nu}\partial_\mu\sigma\partial_\nu\sigma - \frac{1}{2}m_\sigma^2\sigma^2 - \frac{\lambda}{2}\sigma\chi^2. \tag{4.63}$$

We are going to calculate quantum corrections to the vertex for emission of a scalar ϕ by matter χ. In order to obtain the action in the Einstein frame, we apply the conformal transformation implicit in (4.5a). As we have seen, exponentials play a somewhat special role in scalar-tensor theories, so, for the purposes of illustration we choose

$$f\left(\frac{\phi}{M}\right) = \exp\left(\frac{\phi}{M}\right). \tag{4.64}$$

Since we are interested in corrections to the vertex at most of order $1/M^3$ we then expand the Einstein-frame action in Minkowski space to third order in ϕ, and drop some of the terms that do not enter our calculation,

$$
\begin{aligned}
\mathcal{L}^E_M = -\frac{1}{2}\partial_\mu\chi\partial^\mu\chi - \frac{m^2}{2}\chi^2 - \frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma - \frac{m_\sigma^2}{2}\sigma^2 - \frac{\lambda}{2}\sigma\chi^2 - \frac{d\,\lambda}{2M}\phi\sigma\chi^2 \\
-\frac{1}{2}\frac{\phi}{M}\left[(d-2)\partial_\mu\chi\partial^\mu\chi + d\,m^2\chi^2\right] - \frac{1}{2}\frac{\phi}{M}\left[(d-2)\partial_\mu\sigma\partial^\mu\sigma + d\,m_\sigma^2\sigma^2\right] \\
-\frac{1}{4}\frac{\phi^2}{M^2}\left[(d-2)^2\partial_\mu\chi\partial^\mu\chi + d^2m^2\chi^2\right] - \frac{1}{12}\frac{\phi^3}{M^3}\left[(d-2)^3\partial_\mu\chi\partial^\mu\chi + d^3m^2\chi^2\right] + \cdots .
\end{aligned} \tag{4.65}
$$

Note that some of the couplings above are redundant, and can be removed away by a field redefinition. Although the field redefinition simplifies the Feynman rules, it somewhat obscures the symmetry between the couplings of the scalar and the graviton, so we shall mostly proceed with the Lagrangian (4.65). Of course either formulations yield the same S-matrix elements.

Matter Loops

Our first goal is to explicitly show that one-loop corrections in which matter fields run inside the loop do respect the equivalence principle. In order to do so, it is simpler (and more revealing) to verify first the Ward-Takahashi identity (4.52). Consider for that purpose the order λ^2 correction to the amplitude for emission of a scalar ϕ by a matter field χ. At this order, the correction is given by the four diagrams in figure 4.4, where χ lines are labeled with an arrow, σ lines are plain and ϕ lines are dashed. Because we are interested in the limit of zero momentum transfer, we consider equal incoming and outgoing momenta. Using the vertices implied by the Lagrangian (4.65), and combining denominators using Feynman parameters in the standard way [54], we find

$$
i\Delta\gamma_1 = -\frac{2\lambda^2}{M}\int d^dk \int_0^1 dx\, \frac{x[(d-2)(p^2(1-x)^2+k^2)+dm^2]}{[k^2+p^2x(1-x)+m^2x+m_\sigma^2(1-x)]^3}, \tag{4.66a}
$$

$$
i\Delta\gamma_2 = -\frac{2\lambda^2}{M}\int d^dk \int_0^1 dx\, \frac{(1-x)[(d-2)(p^2x^2+k^2)+d\,m_\sigma^2]}{[k^2+p^2x(1-x)+m^2x+m_\sigma^2(1-x)]^3}, \tag{4.66b}
$$

$$
i\Delta\gamma_3 + i\Delta\gamma_4 = \frac{2d\,\lambda^2}{M}\int d^dk \int_0^1 dx\, \frac{1}{[k^2+p^2x(1-x)+m^2x+m_\sigma^2(1-x)]^2}, \tag{4.66c}
$$

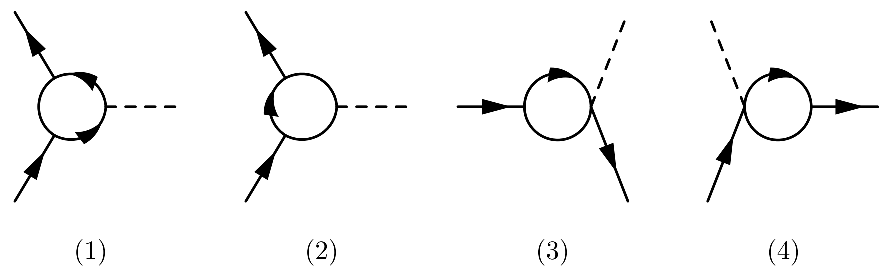


Figure 4.4: One-loop corrections to the vertex for scalar emission by matter

where we have dropped the $i\epsilon$ factors in the propagators. Combining all the contributions in (4.66) we thus conclude that the total vertex correction is

$$i\Delta\gamma \equiv i\sum_{i=1}^{4}\Delta\gamma_i = \frac{\lambda^2}{M}\int d^dk\int_0^1 dx\,\frac{4k^2+4p^2x(1-x)}{[k^2+p^2x(1-x)+m^2x+m_\sigma^2(1-x)]^3}. \quad (4.67)$$

The interactions of matter χ with the field σ also modify the self-energy of matter. At order λ^2, the self-energy corrections are described by the diagram in figure 4.5, which leads to

$$i\Delta\pi = \lambda^2\int d^dk\int_0^1 dx\,\frac{1}{[k^2+p^2x(1-x)+m^2x+m_\sigma^2(1-x)]^2}, \quad (4.68a)$$

and directly yields

$$i\left(d\Delta\pi - p^\mu\frac{\partial\Delta\pi}{\partial p^\mu}\right) = \lambda^2\int d^dk\int_0^1 dx\left(\frac{d}{[k^2+p^2x(1-x)+m^2x+m_\sigma^2(1-x)]^2}+\right.$$
$$\left.+\frac{4p^2x(1-x)}{[k^2+p^2x(1-x)+m^2x+m_\sigma^2(1-x)]^3}\right). \quad (4.68b)$$

The integrals over loop momenta in equations (4.67) and (4.68) can be explicitly carried out by rotating the integration contour counterclockwise into Euclidean momenta and making use of the well known relation

$$\int d^dk_E\frac{(k^2)^n}{[k^2+\Delta^2]^m} = \pi^{d/2}\frac{\Gamma(\frac{d+2n}{2})}{\Gamma(\frac{d}{2})}\frac{\Gamma(m-\frac{d+2n}{2})}{\Gamma(m)}\Delta^{d+2n-2m}, \quad (4.69)$$

which immediately confirms the Ward identity (4.52).

When we calculate S-matrix elements (as opposed to Green's functions) it is convenient to work in the OS scheme of renormalized perturbation theory. We then need to introduce appropriate field renormalization and mass counterterms to enforce our renormalization conditions (4.39). With $\chi \to Z^{1/2}\chi$ and $m^2 \to m^2 - \delta m^2$ the counterterm Lagrangian becomes

$$\begin{aligned}\mathcal{L}_C^E = & - \frac{1}{2}(Z-1)(\partial_\mu\chi\partial^\mu\chi + m^2\chi^2) + \frac{1}{2}Z\delta m^2\chi^2 - \\ & - \frac{1}{2}\frac{\phi}{M}\left\{(Z-1)[(d-2)\partial_\mu\chi\partial^\mu\chi + d\,m^2\chi^2] - d\,Z\delta m^2\chi^2\right\} + \cdots ,\end{aligned} \tag{4.70}$$

with Z and δm^2 chosen to satisfy the conditions (4.39),

$$Z - 1 = \frac{1}{(2\pi)^d}\frac{d\Delta\pi}{dp^2}\bigg|_{p^2=-m^2}, \quad Z\delta m^2 = -\frac{\Delta\pi(-m^2)}{(2\pi)^d}. \tag{4.71}$$

These counterterms yield the additional contributions to the vertex amplitude

$$i\Delta\gamma_5 = -i\frac{(2\pi)^d}{M}\left\{(Z-1)\left[(d-2)p^2 + d\,m^2\right] - d\,Z\delta m^2\right\}. \tag{4.72}$$

Using the Ward identity (4.52), evaluated at $p^2 = -m^2$, it is now straightforward to see that the total vertex correction vanishes. Alternatively, bringing all the factors in $\Delta\gamma_i$ to a common denominator, and simplifying the resulting numerator we find that the total vertex correction is

$$i(\Delta\gamma_\phi)_{OS} \equiv i\sum_{i=1}^{5}\Delta\gamma_i = \frac{\lambda^2}{M}\int d^dk\int_0^1 dx\,\frac{(4-d)k^2 - d[m^2x^2 + m_\sigma^2(1-x)]}{[k^2 + m^2x^2 + m_\sigma^2(1-x)]^3}. \tag{4.73}$$

Using equation (4.69) in (4.73) yields again $(\Delta\gamma_\phi)_{OS} = 0$, in agreement with our general result (4.53). The corresponding cancellation among the five different diagrams is an expression of diffeomorphism and Weyl invariance. In the Lagrangian (4.65), the vertex to which a single scalar ϕ is attached could be replaced by one to which a single graviton is attached. Since the Ward identity (4.41) guarantees that the sum of all diagrams that contribute to the vertex correction for graviton emission vanishes in the appropriate kinematic limit, this result transfers to the vertex for emission of a scalar particle.

This also explains why the total vertex correction does not vanish if we simply use a cut-off to regularize the theory. If we cut off the Euclidean momentum integrals at $k_E = \Lambda$ in $d = 4$ we get, from equation (4.73),

$$(\Delta\gamma_\phi)_{OS} = -\frac{2\pi^2\lambda^2}{M}\int_0^1 dx\left(1 + \frac{m^2x^2 + m_\sigma^2(1-x)}{\Lambda^2}\right)^{-2}. \tag{4.74}$$

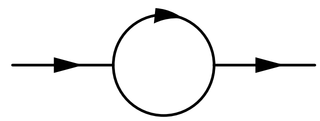


Figure 4.5: One loop correction to the self-energy of matter

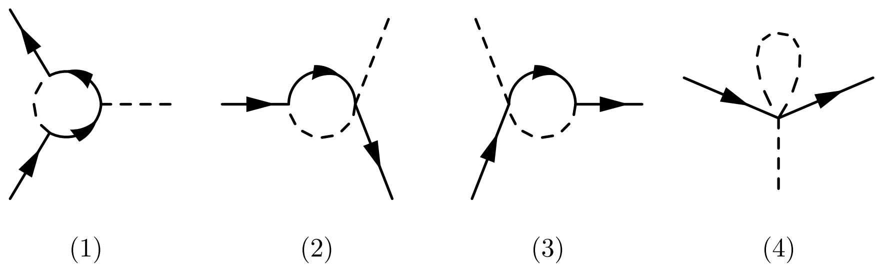


Figure 4.6: One-loop corrections to the vertex for scalar emission at order $1/M^3$. Continuous lines denote matter fields, while dashed lines label the scalar ϕ.

This remains finite in the limit $\Lambda \to \infty$, but does not vanish. The origin of the non-zero correction is of course the breaking of diffeomorphism invariance by the momentum cut-off, which leads to a breakdown of the Ward-Takahashi identity for graviton emission (4.10), but does not affect the relation (4.9) between the vertex and the graviton vertex. Although the quantum theory of massless spin two particles with non-derivative couplings to matter requires diffeomorphism invariance [51], the coupling of a spin zero scalar ϕ to matter does not demand any symmetry. In other words, by regulating the momentum integrals with a cut-off, we are not breaking any symmetry in the scalar sector that is not already broken, so a momentum cut-off appears to be a perfectly valid regularization method. In this light, even our claim that matter loops do respect the equivalence is somewhat misleading.

Scalar Loops

We proceed now to calculate corrections to the vertex that include the scalar ϕ running inside a loop. These are described by the four diagrams in figure 4.6, which respectively lead to the four vertex corrections

$$i\Delta\gamma_1 = -\frac{2}{M^3}\int d^dk\,dx\,x\,\times \tag{4.75a}$$
$$\times\,\frac{[(d-2)p\cdot(p(1-x)-k)+d\,m^2]^2[(d-2)(p(1-x)-k)^2+d\,m^2]}{[k^2+p^2x(1-x)+m^2x+m_\phi^2(1-x)]^3},$$

$$i\Delta\gamma_2 = \frac{1}{M^3}\int d^dk\,dx\,\times \tag{4.75b}$$
$$\times\,\frac{[(d-2)p(p(1-x)-k)+d\,m^2]\,[(d-2)^2\,p(p(1-x)-k)+d^2\,m^2]}{[k^2+p^2x(1-x)+m^2x+m_\phi^2(1-x)]^2},$$

$$i\Delta\gamma_3 = i\Delta\gamma_2, \tag{4.75c}$$

$$i\Delta\gamma_4 = -\frac{1}{2M^3}\int d^dk\,\frac{(d-2)^3p^2+d^3\,m^2}{k^2+m_\phi^2}, \tag{4.75d}$$

where, from now on and as before, the integral over x covers the range from zero to one.

Because we want to show that ϕ loops do lead to violations of the equivalence principle, it is more convenient to work in an on-shell renormalization scheme (OS). The self-energy insertion $\Delta\pi$ is determined by the two diagrams in figure 4.7, and the corresponding corrections read

$$i\Delta\pi_1 = \frac{1}{M^2}\int d^dk\int_0^1 dx\,\frac{[(d-2)p\cdot(p(1-x)-k)+d\,m^2]^2}{[k^2+p^2x(1-x)+m^2x+m_\phi^2(1-x)]^2}, \tag{4.76}$$

$$i\Delta\pi_2 = -\frac{1}{2M^2}\int d^dk\,\frac{(d-2)^2p^2+d^2m^2}{k^2+m_\phi^2}. \tag{4.77}$$

In order to enforce the renormalization conditions (4.39), we introduce a renormalized field $\chi \to Z^{1/2}\chi$ and a renormalized mass $m^2 \to m^2 - \delta m^2$, which give the counterterms in the Lagrangian (4.70). But because we are dealing now with non-renormalizable interactions (operators of mass dimension higher than d), the self-energy also contains a divergent term proportional to p^4, which we cannot absorb simply by renormalization of fields and parameters present in the action (4.65). We are thus forced to introduce a new bare term with four derivatives and two fields, which we treat as a perturbation. In the Jordan frame Lagrangian, this can be taken to be proportional to $\det e\,(\Box\chi)^2$, which in the Einstein frame becomes

$$\mathcal{L}_C^E \supset \frac{Z\,\delta c}{2}\,f^{d-4}\cdot(\Box\chi)^2, \tag{4.78}$$

with $Z\delta c$ chosen to enforce for instance the additional renormalization condition

$$\left.\frac{d\Delta\pi}{d(p^4)}\right|_{p^2=-m^2} = 0. \tag{4.79}$$

(For simplicity we assume that the renormalized c vanishes.) The counterterms then yield additional vertex corrections, as in equation (4.72), but with the additional contribution from (4.78)

$$i\Delta\gamma_5 = -i\frac{(2\pi)^d}{M}\left\{(Z-1)\left[(d-2)p^2 + d\,m^2\right] - d\,Z\delta m^2 - (d-4)Z\,\delta c\,p^4\right\}. \tag{4.80}$$

From the structure of the self-energy corrections, it is clear that the counterterms are of order M^{-2}.

We are ready to compute now the total correction to the vertex $(\Delta\gamma)_{OS} = \sum_i \Delta\gamma_i$. To make our point, let us concentrate of the phenomenologically relevant case of $d = 4$ dimensions. In this limit, some of the momentum integrals diverge. It is relatively easy to isolate the residue of the pole as $d \to 4$, which, in the limit $m_\phi = 0$ and after performing a trivial integral over x reads

$$(\Delta\gamma_\phi)_{OS} = -\frac{4\pi^2}{M^3}\frac{16m^4 + 7m^2p^2 + p^4}{d-4} + \mathcal{O}[(d-4)^0]. \tag{4.81}$$

The form of this pole immediately reveals that the theory defined by the action (4.5a) is non-renormalizable, in the broad sense that we cannot absorb its divergences by appropriate renormalization of the coupling constants and parameters appearing in *any* matter action of the form (4.5a). Say, suppose that we introduce a renormalized coupling constant by replacing $M^{-1} \to M^{-1} - \delta M^{-1}$. This introduces additional counterterms in our theory, which to leading order in $1/M$ yield an additional vertex correction

$$i\Delta\gamma_6 = -i(2\pi)^d\delta M^{-1}\left[(d-2)p^2 + d\,m^2\right]. \tag{4.82}$$

But comparison of equation (4.81) with (4.82) quickly reveals that no single choice of δM^{-1} cancels all the residues at $d = 4$, and that, in fact, we would have to choose three independent counterterms to cancel the terms proportional to m^4, m^2p^2 and p^4. This means that our theory contains three independent coupling constants, instead of one, as we initially thought.

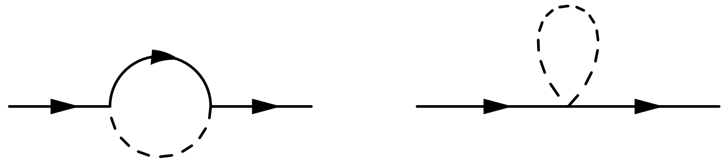


Figure 4.7: Self-energy of matter to order $1/M^2$.

What we are seeing here is that there is no symmetry that enforces the structure (4.5a) in scalar-tensor theories. In order to carry out the renormalization program we have to introduce all the terms compatible with the symmetries of the theory, which in this case *only* consists of diffeomorphism invariance. In particular, just in the scalar sector alone, we have to introduce a set of coupling constant $1/M_i^{(j)}$ for each linear coupling of ϕ to an operator quadratic in the scalar matter species χ_i,

$$\mathcal{L}_M^E \to \sum_i \left[-\frac{1}{2}\partial_\mu \chi_i \partial^\mu \chi_i - \frac{1}{2} m_i^2 \chi_i^2 - \frac{1}{2}\frac{\phi}{M_i^{(0)}} m_i^2 \chi_i^2 \right. \\ \left. -\frac{1}{2}\frac{\phi}{M_i^{(2)}}\partial_\mu \chi_i \partial^\mu \chi_i - \frac{1}{2}\frac{\phi}{M_i^{(4)}}(\Box \chi_i)^2 + \cdots \right]. \tag{4.83}$$

Because no common choice for all counterterms $\delta M_i^{(k)}$ can eliminate all the contributions to the pole at $d = 4$ in equation (4.81) for all matter species, and because the beta functions of the different coupling constants are determined by the coefficients of this pole [176], these different couplings run differently with scale under the renormalization group flow. Thus, once we include quantum corrections, the structure of (4.5) becomes untenable. The unnatural structure of the subclass of scalar-tensor theories we consider here has been repeatedly emphasized by Damour (see e.g. [177]).

Let us proceed anyway with the vertex correction and study its finite piece in the limit $d \to 4$. To simplify the algebra, we consider now on-shell momenta, $p^2 = -m^2$ and focus on the limit $m_\phi = 0$. In this case, the finite terms reduce to

$$(\Delta\gamma_\phi)_{OS} = \mathcal{O}\left(\frac{1}{d-4}\right) - \frac{4\pi^2}{M^3} m^4 \left[2 + 5\gamma + 5\log(\pi m^2)\right], \tag{4.84}$$

which again differs from zero. Of course, we should expect similar terms from the renormalization prescription that eliminates the pole at $d = 4$. Although we have explicitly calculated the corrections of order $(m/M)^3$, due to a scalar loop, we also expect non-vanishing corrections of order $m^3/(MM_P^2)$ due to a graviton loop, as we argued in Section 4.3.2.

Equations (4.81) and (4.84) explicitly show that quantum corrections in scalar-tensor theories generically lead to violations of the equivalence principle. Of course, to make a precise and definite prediction about the size of these violations, we need to specify a renormalization prescription to eliminate the poles at $d = 4$. In the absence of such a prescription, and on dimensional grounds, we generically expect the contribution of the scalar vertex to these violations to be of order $m^4/(MM_P^2)$ (to obtain the scattering amplitude one has to multiply this number by two powers of the appropriate mode function $u \propto 1/\sqrt{2p^0} \approx 1/\sqrt{2m}$). In that case, particles with different masses fall with different accelerations. In order to quantify the corresponding violations of the equivalence principle, it is conventional to quote the Eötvös parameter η, defined to be the relative difference in acceleration of two different test bodies A and B,

$$\eta_{AB} = 2\frac{a_A - a_B}{a_A + a_B}. \tag{4.85}$$

To leading order in gravitational couplings, $a_A + a_B$ is of order $1/M_P$, while our results indicate that $a_A - a_B$ is of order $(m_A^2 - m_B^2)/(MM_P^2)$. Hence, generically we expect the Eötvös parameter to be of order

$$\eta_{AB} \sim \frac{m_A^2 - m_B^2}{MM_P}, \tag{4.86}$$

which is negligible for practical purposes for elementary particle masses. But this does not necessarily rule out the phenomenological relevance of these corrections. If instead of using an on-shell renormalization scheme we had worked for instance with minimal subtraction (MS), we would have found an Eötvös parameter of order

$$\eta \sim \frac{\mu^2}{M_\mu M_P^\mu}\left[\frac{(m_\mu^A)^4}{(m_I^A)^2} - \frac{(m_\mu^B)^4}{(m_I^B)^2}\right], \tag{4.87}$$

where m_μ is the mass parameter in the MS scheme, and m_I is the inertial mass. The key is that for light scalars (in the presence of fine tuning) the inertial mass m_I may differ from the renormalized parameter $m_\mu\mu$ at a high scale $\mu \sim M_P$ by several orders of magnitude. In that case, the Eötvös parameter may be of order one, and thus these quantum violations are phenomenologically relevant. In any case, tests of the weak equivalence principle are not performed with elementary particles, but with macroscopic bodies instead. In order to predict the corresponding violations of the equivalence principle, we would have to proceed as in [158].

Ward-Takahashi Identity for Broken Symmetry

Our explicit calculation of the one-loop correction for scalar emission mediated by the scalar itself also allows us to check the Ward-identity (4.48) and illustrate its meaning. For that purpose let us rewrite equation (4.48) in the form

$$\Gamma_\phi - \frac{1}{M}\Gamma_\Delta = \frac{2M_P}{M}(\Gamma_h)^\mu{}_\mu. \tag{4.88}$$

On the left hand side of (4.88), the corrections to Γ_ϕ to order $1/M^3$ are determined by the four diagrams in figure 4.6, and are given by equations (4.75). As we mention in Appendix 4.A, Γ_Δ is given by all 1PI diagrams with two external matter lines, and an insertion of the vertex Δ, the change in the Lagrangian density under the transformation (4.47). To calculate the sum of these diagrams to order $1/M^3$ we just need to expand the change in the total action under the transformation (4.47) to quadratic order in ϕ. Since we are considering an exponential, equation (4.64), only S_ϕ changes under the transformation,

$$\Delta S_{\text{tot}} = \frac{1}{2}\int d^d x\ \left[(d-2)\partial_\mu\phi\partial^\mu\phi + d\, m_\phi^2\phi^2\right] \equiv \int d^d x\, \Delta. \tag{4.89}$$

To leading order, insertion of this vertex in a diagram with two external lines then leads to the two diagrams in figure 4.8, which, respectively, contribute

$$\Gamma^1_\Delta = -\frac{2i}{M^2}\int d^d k\, dx\,(1-x) \times \tag{4.90a}$$
$$\times\frac{[(d-2)(k+px)^2 + d\, m_\phi^2][(d-2)p\cdot(p(1-x)-k) + d\, m^2]^2}{[k^2+p^2x(1-x)+m^2x+m_\phi^2(1-x)]^3},$$

$$\Gamma^2_\Delta = \frac{i}{2M^2}\int d^d k\,\frac{[(d-2)k^2 + d\, m_\phi^2][(d-2)^2p^2 + d^2\, m^2]}{[k^2+m_\phi^2]^2}. \tag{4.90b}$$

To calculate the right hand side of equation (4.88) we need to expand the total

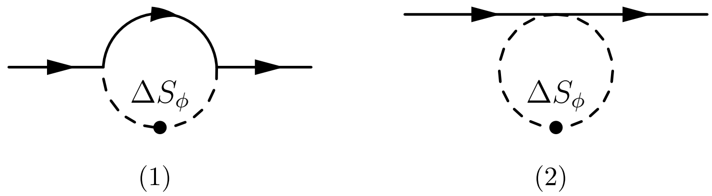


Figure 4.8: Diagrams with two external lines and the insertion of the two vertices in Eq.(4.89).

action to first order in the graviton, and second order in (ϕ/M),

$$\begin{aligned}
&\mathcal{L}_\phi^E + \mathcal{L}_M^E \supset \\
&-\frac{h_{\mu\nu}}{2M_P}\left[\eta^{\mu\nu}\left(\frac{1}{2}\partial_\rho\chi\partial^\rho\chi + \frac{1}{2}m^2\chi^2 + \frac{1}{2}\partial_\rho\phi\partial^\rho\phi + \frac{1}{2}m^2\phi^2\right) - \partial^\mu\chi\partial^\nu\chi - \partial^\mu\phi\partial^\nu\phi\right] - \\
&\qquad -\frac{h_{\mu\nu}}{2M_P}\frac{\phi}{M}\left[\eta^{\mu\nu}\left(\frac{d-2}{2}\partial_\rho\chi\partial^\rho\chi + \frac{d}{2}m^2\chi^2\right) - (d-2)\partial^\mu\chi\partial^\nu\chi\right] - \\
&\qquad -\frac{h_{\mu\nu}}{2M_P}\frac{\phi^2}{2M^2}\left[\eta^{\mu\nu}\left\{\frac{(d-2)^2}{2}\partial_\rho\chi\partial^\rho\chi + \frac{d^2}{2}m^2\chi^2\right\} - (d-2)^2\partial^\mu\chi\partial^\nu\chi\right]. \qquad (4.91)
\end{aligned}$$

Then, to order $1/M^2$, Γ_h on the right hand side of equation (4.88) is given by the six diagrams in figure 4.9, with vertices determined by the action (4.91). Let us label the contribution of the i-th diagram $(\Delta\gamma_i)^{\mu\nu}$. Then, we can write

$$(\Gamma_h)^{\mu\nu} = (\gamma_h)^{\mu\nu} + \sum_{i=1}^{6}(\Delta\gamma_i)^{\mu\nu}, \qquad (4.92)$$

where $(\gamma_h)^{\mu\nu}$ is the tree-level contribution.

Comparing the action (4.65) with (4.91) immediately reveals that the trace of the tree-level vertex for scalar emission by matter equals the trace of the tree-level vertex for graviton emission,

$$\gamma_\phi = \frac{2M_P}{M}(\gamma_h)^\mu{}_\mu. \qquad (4.93)$$

This is just a reflection of the invariance of the matter action under (4.47), as we discussed earlier. Therefore, it follows in addition that the contributions of diagram 4.6.1, equation (4.75b), and the trace of that of 4.9.1 are proportional to each other,

$$\Delta\gamma_1 = \frac{2M_P}{M}(\Delta\gamma_1)^\mu{}_\mu. \qquad (4.94)$$

Diagrams 4.6.2 and 4.6.3 are the same as those in 4.9.2 and 4.9.3. In fact, since the quartic vertex in (4.65) is proportional to the trace of the quartic vertex in equation (4.91), both pairs of diagrams basically yield identical contributions

$$\Delta\gamma_2 + \Delta\gamma_3 = \frac{2M_P}{M}\left[(\Delta\gamma_2)^\mu{}_\mu + (\Delta\gamma_3)^\mu{}_\mu\right]. \tag{4.95}$$

Similarly, diagrams 4.6.4 and 4.9.4 are also identical, and because the quintic vertex in (4.65) is proportional to the trace of the quintic vertex in (4.91), both diagrams are again proportional to each other,

$$\Delta\gamma_4 = \frac{2M_P}{M}(\Delta\gamma_4)^\mu{}_\mu. \tag{4.96}$$

Furthermore, it is clear from the structure of the couplings in (4.65) and (4.91) that these relations only apply for an exponential f.

On the other hand, comparison of figures 4.6 and 4.9 reveals that diagrams 4.9.5 and 4.9.6 do not have a scalar emission counterpart, simply because there is no analogous cubic vertex for ϕ in the action. This is corrected for by the two diagrams with an insertion of Δ in figure 4.8, whose contribution equals the trace of their graviton counterpart. To order $1/M^2$ this implies

$$-\Gamma_\Delta = \frac{2M_P}{M}\left[(\Delta\gamma_5)^\mu{}_\mu + (\Delta\gamma_6)^\mu{}_\mu\right]. \tag{4.97}$$

Together, equations (4.93), (4.94), (4.95), (4.96) and (4.97) immediately provide an explicit verification of equation (4.88).

We can further test the validity of equation (4.88) by noting that, because of equations (4.41) and (4.93), for fields renormalized on shell we should have

$$(\Delta\gamma_\phi)_{OS} = \frac{\Gamma_\Delta}{M}. \tag{4.98}$$

Indeed, we have explicitly checked that in the limit $d \to 4$ both the pole and the finite parts on both sides of the last equation agree.

4.4.2 Fermion matter

We turn our attention now to the vertex for scalar emission by fermionic matter. As we mentioned above, fermions are different from bosons because coupling them to

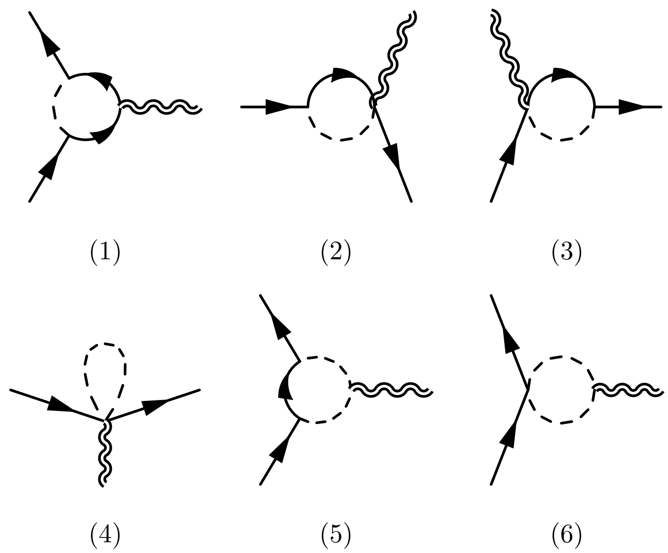


Figure 4.9: One-loop corrections to the vertex for graviton emission to order $1/M^2$.

gravity necessarily requires the introduction of the vierbein. This section illustrates that, as far as the equivalence principle is concerned, this property does not introduce any new ingredients, and that the properties of the vertex with fermion matter closely resemble those of the vertex with scalar matter.

Consider the Jordan-frame matter action

$$S_M^J = \int d^d x \det e \left[-\bar{\psi} e^\mu{}_a \gamma^a D_\mu \psi - m\bar{\psi}\psi - \frac{1}{2}\partial_\mu \chi \partial^\mu \chi - \frac{1}{2} m_\chi^2 \chi^2 - \lambda \chi \bar{\psi}\psi \right], \quad (4.99)$$

which simply describes the Yukawa interactions of a spin 1/2 fermion ψ with a massive Higgs-like scalar χ. Here, γ^a are the conventional Dirac matrices and D_μ is the covariant derivative of the spinor, which depends on the vierbein through the spin connection. To obtain the Einstein frame action, we replace $e_\mu{}^a$ by $f(\phi/M)e_\mu{}^a$, and expand the resulting expression to the desired order in ϕ around flat space. But in order to calculate S-matrix elements, it is simpler to work with a Lagrangian in which some of the interactions have been removed by a field redefinition. It is well-known [83] that the action of a massless spinor is invariant under the Weyl transformation

$$e_\mu{}^a \rightarrow f\, e_\mu{}^a, \quad (4.100a)$$

$$\psi \rightarrow f^{(1-d)/2}\, \psi. \quad (4.100b)$$

Thus, making these substitutions in the Jordan-Frame action (4.99) and expanding again to the required order we get in flat space

$$\mathcal{L}_M^E = -\bar{\psi}\gamma^\mu\partial_\mu\psi - m\left(1+\frac{\phi}{M}+\frac{\phi^2}{2M^2}+\frac{\phi^3}{6M^3}\right)\bar{\psi}\psi - \lambda\chi\left(1+\frac{\phi}{M}\right)\bar{\psi}\psi - \\ -\frac{1}{2}\left(1+(d-2)\frac{\phi}{M}\right)\partial_\mu\chi\partial^\mu\chi - \left(1+d\frac{\phi}{M}\right)\frac{1}{2}m_\chi^2\chi^2+\cdots, \quad (4.101)$$

where we have assumed again that f is an exponential, equation (4.64). Because the couplings in this Lagrangian are not of the form (4.5a), the vertex amplitudes do not obey the Ward identity (4.52), as can be easily verified at tree level. Instead, because the field redefinition (4.100) is of the form (4.60), the vertex obeys the dilatation Ward identity (4.62), as can be also easily verified at tree level. Note that in order to appropriately take into account the spinor field redefinition, we have to multiply the path integral measure by an appropriate Jacobian [178]. For an electrically neutral spinor, this has no effects to linear order in ϕ.

Matter Loops

Our first goal is to calculate the order λ^2/M corrections to the scalar-matter vertex induced by one-loop diagrams in which matter fields run inside the loop. The corresponding diagrams, figure 4.4, are the same as for scalar matter. We do not include external line corrections because we work in the OS scheme. In order to do so however, we need to introduce the appropriate counterterms to enforce our renormalization conditions (4.39). Introducing renormalized fields and mass parameters, $\psi \to Z_2^{1/2}\psi$ and $m \to m - \delta m$, we thus arrive at the counterterms

$$\mathcal{L}_C^E = -(Z_2-1)\left[\bar{\psi}\gamma^\mu\partial_\mu\psi + m\bar{\psi}\psi\right] + Z_2\delta m\bar{\psi}\psi - \left[(Z_2-1)m - Z_2\delta m\right]\frac{\phi}{M}\bar{\psi}\psi+\cdots, \quad (4.102)$$

where we have kept only those terms that are relevant for our calculation.

The determination of the amplitudes associated with the diagrams in figure 4.4 is straight-forward. To simplify the analysis, we concentrate in the limit of zero momentum transfer and on-shell momenta, which is the appropriate limit for our considerations. Following the standard Feynman rules (see e.g. [54]) we find that the contribution of the diagrams in figure 4.4 is

$$i\Delta\gamma_1 = -\frac{2\lambda^2 m}{M}\int d^d k\int_0^1 dx\,\frac{x\,[m^2(2-x)^2-k^2]}{[k^2+m^2x^2+m_\chi^2(1-x)]^3}, \tag{4.103a}$$

$$i\Delta\gamma_2 = -\frac{2\lambda^2 m}{M}\int d^d k\int_0^1 dx\,(1-x)\times \\ \times\frac{(2-x)\left[(d-2)(k^2-m^2x^2)+dm_\chi^2\right]-2(1-2/d)k^2x}{[k^2+m^2x^2+m_\chi^2(1-x)]^3}, \tag{4.103b}$$

$$i\Delta\gamma_3 = \frac{\lambda^2 m}{M}\int d^d k\int_0^1 dx\,\frac{2-x}{[k^2+m^2x^2+m_\chi^2(1-x)]^2}, \tag{4.103c}$$

$$i\Delta\gamma_4 = i\Delta\gamma_3, \tag{4.103d}$$

where we have used that on shell we may substitute $\not{p}$ by im.

In addition, we need to consider the contributions of the counterterms, which in this case reduce to

$$i\Delta\gamma_5 = -i(2\pi)^d(Z_2-1)\frac{m}{M}+i(2\pi)^d Z_2\frac{\delta m}{M}. \tag{4.104}$$

We choose these counterterms to enforce the on-shell renormalization conditions (4.39), which requires

$$Z_2-1 = -\frac{i}{(2\pi)^d}\frac{\partial\Delta\pi}{\partial\not{p}}\bigg|_{\not{p}=im}, \qquad Z_2\delta m = -\frac{\Delta\pi(im)}{(2\pi)^d}. \tag{4.105}$$

In order to calculate the values of the counterterms, we thus need to evaluate the self-energy correction. This is given by the diagram in figure 4.5, which finally leads to

$$i\Delta\pi = \lambda^2 m\int d^d k\int_0^1 dx\,\frac{2-x}{[k^2+m^2x^2+m_\chi^2(1-x)]^2}, \tag{4.106a}$$

$$\frac{\partial\Delta\pi}{\partial\not{p}} = -\lambda^2\int d^d k\,dx\left[\frac{1-x}{[k^2+m^2x^2+m_\chi^2(1-x)]^2}+\frac{4m^2(2-x)(1-x)x}{[k^2+m^2x^2+m_\chi^2(1-x)]^3}\right]. \tag{4.106b}$$

Then, the total loop correction to the vertex for scalar emission by matter is

$$(\Delta\gamma_\phi)_{OS} = \sum_{i=1}^{5}\Delta\gamma_i. \tag{4.107}$$

According to the Ward identity (4.52), the right hand side of equation (4.107) has to vanish, as the vertex correction only involves matter couplings in the loop. But as opposed to what happens in the scalar case, in which the Ward identity at one-loop can be readily verified, one has to complete a surprising amount of work here to show that $(\Delta\gamma_\phi)_{OS}$ equals zero. We leave this task for Appendix 4.B, in which we explicitly prove that, indeed,

$$(\Delta\gamma_\phi)_{OS} = 0, \tag{4.108}$$

in agreement with our general result (4.53). As before, the corresponding cancellation among the five different diagrams is an expression of diffeomorphism and Weyl invariance.

If we regularize the theory by introducing a momentum cut-off Λ, diffeomorphism invariance is broken again, and the cancellation (4.108) does not hold. Instead, say, in the limit $m_\phi \to 0$ we find that $(\Delta\gamma_\phi)_{OS}$ is logarithmically divergent,

$$(\Delta\gamma_\phi)_{OS} \to \frac{7\lambda^2\pi^2}{6}\frac{m}{M} - \lambda^2\pi^2\frac{m}{M}\int_0^1 dx\,(5-14x+6x^2)\log\frac{\Lambda^2}{m^2x^2}. \tag{4.109}$$

As in the scalar case, in order to renormalize this divergence we would have to introduce a coupling constant counterterm δM^{-1} to the Lagrangian, which would contribute

$$i\Delta\gamma_6 = -i(2\pi)^4\delta M^{-1}m \tag{4.110}$$

to the vertex amplitude. In that case, we could *impose* the condition $(\Delta\gamma_\phi)_{OS}+\Delta\gamma_6 = 0$, which would guarantee the preservation of the weak equivalence principle at one loop. But of course, since neither the Yukawa coupling λ nor the mass m are universal, this would lead to a collection of widely different set of bare coupling constants M_i, one for each fermion species, and it would remain a mystery why the renormalized vertex correction for all of them vanishes at zero momentum transfer. Otherwise, equation (4.109) implies generic values of the Eötvös parameter η_{AB} of order λ^2.

Scalar Loops

Having seen how matter loop corrections do respect the equivalence principle (in the dimensionally regularized theory), let us turn our attention to those corrections that do lead to violations. This time, instead of looking at diagrams with matter loops, we shall calculate the corrections caused by a scalar field loop, at order $1/M^3$.

The one-loop scalar field corrections to the scalar vertex are the same as for scalar matter; they are shown in figure 4.6. The self-energy corrections are also given by the diagrams in figure 4.7. Comparison of the corrections to the vertex caused by a fermion loop to those caused by the scalar shows that vertices and most of the diagrams basically agree if one replaces fermion lines with scalar lines. Therefore, we can borrow the results of the previous subsection, now keeping the momenta off-shell, simply by replacing λ by m/M, and m_χ^2 by m_ϕ^2. We do not need to consider the contribution of equation (4.103b), which does not have a counterpart in the scalar loop diagrams at order $1/M^3$. Therefore, the vertex loop correction is the sum of the four terms

$$i\Delta\gamma_1 = \frac{2m^3}{M^3}\int d^dk \int_0^1 dx\, x\, \frac{p^2(1-x)^2 + 2i\not{p}\, m(1-x) + k^2 - m^2}{[k^2 + p^2x(1-x) + m^2x + m_\phi^2(1-x)]^3}, \tag{4.111a}$$

$$i\Delta\gamma_2 = \frac{m^2}{M^3}\int d^dk \int_0^1 dx\, \frac{-i\not{p}(1-x) + m}{[k^2 + p^2x(1-x) + m^2x + m_\phi^2(1-x)]^2}, \tag{4.111b}$$

$$i\Delta\gamma_3 = i\Delta\gamma_2, \tag{4.111c}$$

$$i\Delta\gamma_4 = -\frac{m}{2M^3}\int d^dk\, \frac{1}{k^2 + m_\phi^2}. \tag{4.111d}$$

The contribution from the counterterms is still given by (4.104), with the latter determined by equations (4.105). But this time, there is a new contribution to the self-energy, captured by the second diagram in figure 4.7,

$$i\Delta\pi_1 = \frac{m^2}{M^2}\int d^dk \int_0^1 dx\, \frac{-i\not{p}(1-x) + m}{[k^2 + p^2x(1-x) + m^2x + m_\phi^2(1-x)]^2}, \tag{4.112a}$$

$$i\Delta\pi_2 = -\frac{m}{2M^2}\int d^dk\, \frac{1}{k^2 + m_\phi^2}. \tag{4.112b}$$

In this case, when we add the contributions from of order $(m/M)^3$, we find that the cancellations that occurred at order λ^2/M before do not operate. To actually see that the overall vertex correction $(\Delta\gamma_\phi)_{OS}$ indeed is different from zero, let us consider again the limit $d \to 4$. In this limit, the correction approaches

$$(\Delta\gamma_\phi)_{OS} = \frac{2\pi^2}{M^3}\frac{im^2\not{p} - 2m^3}{d-4} + \mathcal{O}[(d-4)^0], \tag{4.113}$$

which again shows that the theory defined by (4.1) is not renormalizable, in the sense that we cannot absorb this pole by renormalization of the coupling constant

M^{-1} in the Lagrangian (4.101). As in the scalar case, once this pole is removed by including the appropriate missing counterterms in the action, we expect then finite vertex corrections of order m^3/M^3 in the limit $d \to 4$, which lead to relative violations of the weak equivalence principle of order m^2/M^2.

Ward-Takahashi Identity for Broken Symmetry

We mentioned in Section 4.3.3 that in flat spacetime, scalar-tensor theories posses a broken dilatation symmetry (4.61), and a corresponding Ward identity for this broken symmetry, equation (4.62). Again, we can use the results of our explicit calculation of the vertex correction in the previous section to check the validity of the Ward identity (4.62), and vice-versa.

Since a fermion has scaling dimension $(d-1)/2$, the vertex for scalar emission Γ_ϕ by fermion matter obeys the identity (4.62) with $n = 1$. The Lagrangian (4.101) is not invariant under the dilatation (4.61), but instead changes by equation (4.89). Therefore, we should have

$$M\,\Gamma_\phi + \left(p^\mu \frac{\partial}{\partial p^\mu} - 1\right)\Pi = \Gamma_\Delta, \tag{4.114}$$

where Γ_Δ is the sum of all 1PI diagrams with two external lines and an insertion of the local operator Δ defined in equation (4.89). Recall that the Ward identity (4.48) does not hold in this case, as can be readily verified at tree level, because the field redefinition (4.100) has led to a matter action that is not of the form (4.5a).

At tree level, it is easy to check the validity of (4.114), since there is no tree-level diagram with an insertion of Δ and two fermion lines. At order $1/M^2$, the corresponding Feynman diagrams are those in figure 4.8, which yield the two correction terms

$$\Gamma^1_\Delta = -\frac{2im^2}{M^2}\int d^dk\,dx\,(1-x)\,\frac{\left[-i\not{p}(1-x)+m+i\not{k}\right]\left[(d-2)(k+x\,p)^2+d\,m_\phi^2\right]}{[k^2+p^2x(1-x)+m^2x+m_\phi^2(1-x)]^3}, \tag{4.115a}$$

$$\Gamma^2_\Delta = \frac{im}{2M^2}\int d^dk\,\frac{(d-2)k^2+d\,m_\phi^2}{[k^2+m_\phi^2]^2}. \tag{4.115b}$$

It is then easy to check for instance that the residue of the pole at $d = 4$ in $\Gamma_\Delta \equiv \Gamma^1_\Delta + \Gamma^2_\Delta$ actually agrees with equation (4.113), thus confirming the validity of the Ward identity (4.114).

4.5 Summary

We have studied the impact of quantum corrections on the weak equivalence principle in scalar-tensor theories that admit a Jordan-frame formulation, equation (4.1). To do so, it is convenient to work in the Einstein frame, in which the scalar and the graviton are decoupled in the free action, equation (4.5). In this frame the amplitude for scalar emission is universally proportional to the inertial mass at tree level, and the same result holds when we include quantum corrections that only involve matter loops. Once we include a scalar ϕ or a graviton in these loop corrections however, the equivalence principle is violated.

The origin of these results lies in the broken Weyl symmetry (4.47). The corresponding Ward identity for the broken symmetry (4.48) relates the 1PI vertex Γ_ϕ for scalar emission to that of the graviton Γ_h, and to the sum of all the diagrams with an insertion of a vertex proportional to the change of the Lagrangian density under the broken symmetry (4.48), Γ_Δ. Violations of the equivalence principle caused by the scalar interaction arise from those terms in the action that violate the shift symmetry (4.47). For an exponential, $f = \exp(\phi/M)$, the matter action is exactly symmetric under (4.47) and only S_ϕ and S_{EH} violate the Weyl symmetry. For other choices of f, such as a linear coupling in ϕ, even the matter Lagrangian is not exactly symmetric under this transformation. In both cases, because the only terms that violate the inhomogeneous Weyl symmetry involve terms quadratic in the scalar ϕ or the graviton, these violations of the equivalence principle are proportional to three powers of the gravitational couplings M^{-1} and M_P^{-1}. If we regularize the theory with a momentum cut-off, diffeomorphism invariance is broken, and even matter loops lead to violations of the weak equivalence principle caused by the scalar interaction. Although diffeomorphism invariance is required to couple a massless graviton to matter, there is no analogous constraint to couple a massive or massless scalar to matter. In particular, a momentum cut-off does break the Weyl symmetry (4.47), but the latter is broken anyway in the action (4.5).

The form of the quantum corrections to the scalar vertex Γ_ϕ implies that scalar-tensor theories with an Einstein frame formulation of the form (4.5a) are not renormalizable: Any matter action of the form (4.5a) does not contain enough counterterms to eliminate all the poles at $d = 4$ in the dimensionally regularized theory. To do so one has to include all the terms compatible with the symmetries of the action, which only consist of diffeomorphism invariance. Therefore, the structure of (4.5a) is not preserved by quantum corrections. From that point of view, assuming that the coupling of the scalar is universally characterized by a single coupling constant $1/M$ appears artificial.

The actual magnitude of the equivalence principle violations depends on the way the theory is regularized, and on the renormalization prescription that eliminates the remaining non-renormalizable divergences in the amplitudes. Generically, in the presence of a high momentum cut-off, we expect the Eötvös parameter of these theories to be of order one, which is strongly ruled out by experiment [157]. In the dimensionally regularized theory, we expect the Eötvös parameter to be of order $\Delta m^2/M_P^2$; this ratio is extremely small for typically inertial masses of elementary particles, but could be large if one of the mass parameters is defined away from the mass shell. In any case, we have not worked out the magnitude of the equivalence principle violations for macroscopic bodies, as appropriate for phenomenological considerations.

Finally, our results can be easily extended to similar classes of theories in which the matter action can be cast as in equation (4.5a), such as $f(R)$ gravity [18] or the Galileon [179]. Because both of them violate the Weyl symmetry (4.47), we expect them to behave like the scalar-tensor theories we have considered here.

Appendix 4.A Ward Identities for Broken Symmetries

It is well-known that linear symmetries of the action are also symmetries of the effective action. In this appendix we are concerned with transformations that, though linear, do not preserve the form of the action. As we shall show, in this case, the quantum effective action satisfies a Ward-Takahashi identity that relates the change of the effective action under the linear transformation to the change in total action functional under the broken symmetry. This general identity has been widely discussed in the literature, see e.g. [180], though its proof is difficult to find. Our derivation here closely follows the formalism of [181] (particularly its Section 12.6).

Consider the generating functional of an arbitrary theory that contains a set of fields χ^n in the presence of a corresponding set of currents J_n,

$$Z[J] = \int D\chi \exp\left(iS_{\text{tot}}[\chi] + i\int d^dx\, J_n(x)\chi^n(x)\right). \tag{4.116}$$

Suppose now that we change integration variables

$$\chi^n(x) \to \chi^n(x) + \epsilon\Delta\chi^n(x), \tag{4.117}$$

where $\Delta\chi^m$ is *linear* in the fields, and ϵ is an arbitrary infinitesimal constant that we use as an expansion parameter (the actual transformation (4.117) may be global or local). Then, invariance of the path integral under change of variables gives, to first order in ϵ,

$$\int D\chi \left(\Delta S_{\text{tot}}[\chi] + \int d^dx\, J_n(x)\Delta\chi^n(x)\right) \exp\left(iS_{\text{tot}}[\chi] + i\int d^dy\, J_n(y)\chi^n(y)\right) = 0, \tag{4.118}$$

where ΔS_{tot} is the total change in the action under the transformation (4.117), and we have also absorbed an eventual change of the functional measure into ΔS_{tot}.

In order to take into account the change of the action under the transformation, it turns out to be convenient to introduce a new generating functional $Z[J,B]$ with an additional (constant) source B for ΔS_{tot},

$$Z[J,B] \equiv \int D\chi \exp\left(iS_{\text{tot}}[\chi] + i\int d^dx\, J_n(x)\chi^n(x) + i\,B\,\Delta S_{\text{tot}}[\chi]\right). \tag{4.119}$$

Then, in terms of this new functional, equation (4.118) takes the form

$$\left(\frac{1}{i\, Z[J,B]} \frac{\partial Z[J,B]}{\partial B} + \int d^d x \, J_n(x) \, \langle \Delta \chi^n(x) \rangle_{J,B} \right) \Bigg|_{B=0} = 0, \tag{4.120}$$

where, for any functional $F[\chi]$ of the fields, we have defined

$$\langle F[\chi] \rangle_{J,B} \equiv Z^{-1}[J,B] \int D\chi \, F[\chi] \exp \left(iS_{\text{tot}}[\chi] + i \int d^d x \, J_n(x) \chi^n(x) + i\, B\, \Delta S_{\text{tot}}[\chi] \right). \tag{4.121}$$

We proceed now to turn Equation (4.120) into an equation for the effective action. We first define the generating function for connected diagrams in the presence of a source for ΔS_{tot} in the standard way,

$$i\, W[J,B] \equiv \log Z[J,B], \tag{4.122}$$

and then introduce the effective action by a Legendre transformation that only involves the curents J_n,

$$\Gamma[\bar{\chi}, B] \equiv W[J[\bar{\chi}, B], B] - \int d^d x \, J_n[\bar{\chi}, B] \bar{\chi}^n. \tag{4.123}$$

The currents $J[\bar{\chi}, B]$ in the last equation are such that the fields χ^n have prescribed expectation values[5] $\bar{\chi}^n(x)$,

$$\langle \chi^n(x) \rangle_{J,B} = \frac{\delta W[J,B]}{\delta J_n(x)} = \bar{\chi}^n(x). \tag{4.124}$$

Therefore, differentiation of equation (4.123) with respect to $\bar{\chi}^n$ and B respectively leads to the identities

$$J_n[\bar{\chi}, B] = -\frac{\delta \Gamma[\bar{\chi}, B]}{\delta \bar{\chi}^n}, \qquad \frac{\partial \Gamma[\bar{\chi}, B]}{\partial B} = \frac{\partial W[J[\bar{\chi}, B], B]}{\partial B} \Bigg|_J . \tag{4.125}$$

We are ready to put all these results together into equation (4.120). First, note that the first term on the left-hand side is simply the derivative of $W[J,B]$ with

[5]If the generating functional depends on the fields χ, and some background values $\bar{\chi}$ through gauge-fixing and ghost terms, the effective action is a functional of both the background fields $\bar{\chi}$ and the expectation values of the fields in the presence of the current, $\Gamma = \Gamma[\bar{\chi}, \langle \chi \rangle_J]$ (we set here $B = 0$ for simplicity.) The effective action in the background field method is defined by setting $\langle \chi \rangle = \bar{\chi}$, so, strictly speaking, the proper vertices are given by functional derivative of $\Gamma = \Gamma[\bar{\chi}, \langle \chi \rangle]$ with respect to $\langle \chi^n \rangle$. As shown in [174] however, the difference is irrelevant when computing S-matrix elements.

respect to B, which, because of (4.125) equals the derivative of the effective action $\Gamma[J, B]$ with respect to the same variable. Moreover, because we assume that $\Delta\chi^n$ is linear in the fields,

$$\langle \Delta\chi^n \rangle_{J,B} = \Delta\bar{\chi}^n, \tag{4.126}$$

so the second term on the left-hand side of equation (4.120) is the change in the effective action $\Delta\Gamma[\bar{\chi}] \equiv \Gamma[\bar{\chi}, 0]$ under the transformation (4.117). Therefore, equation (4.120) reads

$$\Delta\Gamma[\bar{\chi}] = \left.\frac{\partial\Gamma[\bar{\chi}, B]}{\partial B}\right|_{B=0}, \tag{4.127}$$

which states that at $B = 0$ the effective action $\Gamma[\bar{\chi}, B]$ is invariant under the transformation (4.117), supplemented with the additional transformation $B \to B - \epsilon$.

The right-hand side of equation (4.127) has a simple interpretation. Typically, ΔS_{tot} is the spacetime integral of a local operator,[6]

$$\Delta S_{\text{tot}} = \int d^d x \, \Delta. \tag{4.128}$$

In that case, $\Gamma[\bar{\chi}, B]$ is the generator of 1PI diagrams in a theory with an additional interaction $\int d^d x \, B \, \Delta$. Therefore, its derivative with respect to the "coupling constant" B at zero simply picks up those 1PI diagrams with a single insertion of the vertex Δ. Since the new interaction involves a spacetime integral, and B is a constant, such an insertion carries zero momentum into the diagram. Thus, denoting by $\Gamma_\Delta[\bar{\chi}] \equiv (\partial\Gamma/\partial B)\big|_{B=0}$ the generator of all 1PI diagrams with a vertex insertion of Δ we arrive at the main result of the appendix,

$$\Delta\Gamma[\bar{\chi}] = \Gamma_\Delta[\bar{\chi}], \tag{4.129}$$

the Ward identity for a broken symmetry expressed in terms of the effective action (variations of the same identity are also known as Slavnov-Taylor or Schwinger-Dyson equations.) By taking functional derivatives of equation (4.129) with respect to the matter fields one can then derive relations between the 1PI vertices of the theory. For

[6] Care should be exercised here because we are discarding a surface terms that may arise upon integration by parts when isolating the change in the action to first order in ϵ.

instance,

$$\Gamma_{\Delta}^{\beta\alpha} \equiv \left. \frac{\delta^2 \Gamma_\Delta}{\delta\bar{\psi}_\alpha(x)\delta\bar{\psi}^\dagger_\beta(y)} \right|_{\bar{\psi}=0} \tag{4.130}$$

is the sum of all 1PI diagrams with two external matter fields (with propagators stripped off) and an insertion of Δ. Note that if the theory is invariant under the transformation (4.117), $\Delta = 0$, equation (4.129) reduces to the well-known Slavnov-Taylor identity for a linear symmetry of the action.

Appendix 4.B Scalar Ward Identity for Fermions

In this appendix we verify that matter loop corrections do not renormalize the vertex for scalar emission by a fermion, equation (4.108).

With a scalar χ running inside the loop, the one-loop correction to the vertex for scalar emission by a fermion is determined by equations (4.103) and (4.104), whereas the one-loop correction to the self-energy of the fermion is given by equations (4.106). To verify the relation (4.108) we need to explicitly carry out the integrals over momenta and x.

The integrals over momenta can be easily performed using the identity (4.69). On shell, the remaining integrals over x turn out to be a sum of expressions of the general form

$$I_n = \int_0^1 dx\, x^n \left[m^2 x^2 + m_\chi^2 (1-x)\right]^{d/2-3}, \tag{4.131}$$

with integer n. After completing a square inside the square bracket, the integral can be re-expressed as

$$I_n = \int_0^1 dx\, x^n (m_\chi)^{d-6} \left(1 - \frac{1}{4r}\right)^{d/2-3} \left[1 + \frac{r}{1-\frac{1}{4r}} \left(x - \frac{1}{2r}\right)^2\right]^{d/2-3}, \tag{4.132}$$

where we have defined the dimensionless ratio

$$r \equiv \frac{m^2}{m_\chi^2}. \tag{4.133}$$

The scalar χ is stable upon decay onto two fermions if $m_\chi < 2m$. In that case

$(1 - 1/4r) > 0$, and we can introduce the new (real) integration variable

$$t = \sqrt{\frac{4r^2}{4r-1}} \left(x - \frac{1}{2r} \right) . \tag{4.134}$$

In terms of this variable, the integral (4.131) thus becomes

$$I_n = m_\chi^{d-6} \left(1 - \frac{1}{4r} \right)^{d/2-3} \sqrt{\frac{4r-1}{4r^2}} \int_{t_0}^{t_1} dt \left(\sqrt{\frac{4r-1}{4r^2}}\, t + \frac{1}{2r} \right)^n \left[1 + t^2 \right]^{d/2-3} , \tag{4.135}$$

where the lower and upper integration limits t_0 and t_1 are determined by respectively setting $x = 0$ and $x = 1$ in equation (4.134). Expanding the n-th power in equation (4.135) we further obtain a linear combination of integrals of the general form

$$J_m \equiv \int_{t_0}^{t_1} dt\, t^m \left[1 + t^2 \right]^{d/2-3} , \tag{4.136}$$

which can finally be expressed in terms of hypergeometric functions [182],

$$\begin{aligned} J_m \quad = \quad & \frac{t_1^{1+m}}{1+m}\, {}_2F_1 \left(3 - \frac{d}{2}, \frac{1+m}{2}, \frac{3+m}{2}; -t_1^2 \right) \\ & - \frac{t_0^{1+m}}{1+m}\, {}_2F_1 \left(3 - \frac{d}{2}, \frac{1+m}{2}, \frac{3+m}{2}; -t_0^2 \right) . \end{aligned} \tag{4.137}$$

In this way, after quite a bit of tedious but straight-forward algebra, collecting all the contributions from the integrals in equation (4.107) we find that they all add to zero, equation (4.108).

Chapter 5

Effective Theory of Cosmological Perturbations

5.1 Introduction

One of the main successes of inflation [77–79, 183] is the explanation of the origin of structure [42–46]. During slow-roll, the Hubble radius remains nearly constant, while cosmological modes are constantly pushed out of the horizon. Thus, local processes determine the amplitude and properties of perturbations at sub-horizon scales, which are transferred to cosmologically large distances by the accelerated expansion. In that sense, the sky is the screen upon which inflation has projected the physics of the microscopic universe.

The primordial perturbations seeded during inflation arise from quantum-mechanical fluctuations of the inflaton around its homogeneous value. Hence, their properties directly depend on the quantum state of the inflaton perturbations. Conventionally, this is taken to be a state devoid of quanta in the asymptotic past, raising the crucial question of whether we can trust cosmological perturbation theory—and its quantum nature—at such early times [96].

According to our present understanding, quantum eld theories and GR are merely low energy descriptions of a more fundamental theory of quantum gravity. In the case of inflation, the leading terms in the corresponding effective Lagrangian are the Einstein-Hilbert term plus the inflaton kinetic term and potential. In an EFT treatment, these terms are accompanied by all other possible operators compatible with the symmetries of the theory, namely, general covariance and any other symmetry of the inflaton sector. Higher dimensional operators are suppressed by powers of an energy scale M, which we will assume to be of the order of the reduced Planck mass, $M \sim M_{\rm p}$, and they are therefore expected to be negligible at sufficiently small momenta, or sufficiently long wavelengths. Note however that this does not imply that we can simply discard high-momentum modes from the low-energy theory. In a gauge theory in flat space for instance, a momentum cut-off breaks gauge invariance and is thus incompatible with the symmetries of the theory. Similarly, in a curved spacetime, the definition of properly renormalized generally covariant field operators requires subtractions that involve all the momentum modes of the fields [184]. The effective theory is a useful low-energy approximation simply because, on dimensional grounds, the corrections to any observable introduced by the higher-dimensional operators must be proportional to ratios of the external momenta or energies that characterize the process to the energy scale M. The goal of this chapter is to determine the three-momentum scale Λ at which such higher-dimensional operators significantly modify the dispersion relation of cosmological modes. Beyond that scale, we cannot trust the free sector of the theory, and cosmological perturbation theory breaks down. Since the dispersion relation of a mode is what sets its mean square amplitude, we identify such a breakdown with the point at which the corrections to the power spectrum caused by higher-dimensional operators become dominant.

In Minkowski spacetime, the scale at which effective corrections to observable quantities become important roughly coincides with the scale that suppresses the non-renormalizable operators in the effective action. For instance, in the presence of such terms, the propagator of a massless particle with (off-shell) momentum k^μ can be cast as an expansion of the form [54]

$$\Delta(k_\mu, k'_\mu) = \frac{1}{k_\mu k^\mu}\left(c_0 + c_2\frac{k_\mu k^\mu}{M^2} + c_4\frac{(k_\mu k^\mu)^2}{M^4} + \cdots\right)\delta(k_\mu - k'_\mu), \tag{5.1}$$

where the c_n are coefficients of order one that typically depend on logarithms of $k_\mu k^\mu$.

Lorentz-invariance implies that the corrections must be a function of the scalar $k_\mu k^\mu$, while Poincare symmetry implies that they must conserve four-momentum. From the structure of the corrections, it is clear that the expansion breaks down around $k_\mu k^\mu = M^2$.

On the other hand, it is crucial to realize that the three-momentum scale Λ at which corrections to the power spectrum become dominant does not need to equal the fundamental scale M. On short time-scales and distances, an inflating spacetime can be regarded as flat. Hence, our previous result in Minkowski space suggests that cosmological perturbation theory is valid as long as $k_\mu k^\mu \ll M^2 = M_P^2$. On shell, the four-momenta of cosmological perturbations are light-like, $k_\mu k^\mu \equiv -k_0^2 + \mathbf{k}\cdot\mathbf{k} = 0$. Thus, substituting in equation (5.1) we find that corrections are not only independent of the three-momentum $\mathbf{k}$, but also that they are actually zero. As we shall see though, the evolution of the inflaton leads to small but finite violations of the Lorentz symmetry even in the short-wavelength limit, which are imprinted on the power spectrum as $\mathbf{k}$-dependent corrections.

The phenomenological imprints of trans-Planckian physics on the primordial spectrum of perturbations, and the implications of a finite cut-off Λ on the spatial momentum of cosmological modes have been extensively studied [97, 185–210]. These articles mostly study corrections to the power spectrum in the long-wavelength limit $|\mathbf{k}/a| \equiv |\mathbf{k}_{ph}| \ll H$, at late times, which is the regime directly accessible by experimental probes. In this chapter we focus instead on the short-wavelength regime $|\mathbf{k}_{ph}| \gg H$, at early times, since we are interested in determining how far into the ultraviolet cosmological linear perturbation theory applies. At short wavelengths, the power spectrum can be cast again as a derivative expansion of the form

$$\langle \delta\varphi^*(\mathbf{k})\delta\varphi(\mathbf{k})\rangle = \frac{1}{2|\mathbf{k}|}\left(\alpha_0 + \alpha_2\frac{\mathbf{k}_{ph}\cdot\mathbf{k}_{ph}}{M^2} + \alpha_4\frac{(\mathbf{k}_{ph}\cdot\mathbf{k}_{ph})^2}{M^4} + \cdots\right), \qquad (5.2)$$

with coefficients α_i that depend on slow-roll parameters and the dimensionless ratio H/M_p. The analytic corrections to the leading result $1/2|\mathbf{k}|$ arise from tree-level diagrams with vertices from higher-dimensional operators. We only consider tree-level diagrams here, since we expect loop diagrams to simply introduce a logarithmic dependence of the dimensionless coefficients α_i on scale, though we have not verified this explicitly. Cosmological perturbation theory fails (in our restricted sense) when the expansion in powers of $|\mathbf{k}|$ breaks down, namely, when all the terms become of the same order,

$$|\mathbf{k}_{ph}| \approx M_{\rm p}\sqrt{\frac{\alpha_{2n}}{\alpha_{2n+2}}} \equiv \Lambda\ . \tag{5.3}$$

As we shall show, the ratios $\alpha_{2n}/\alpha_{2n+2}$ are all quite large and of the same order, so the effective cut-off Λ significantly differs from $M_{\rm p}$. In a slightly different context, a similar analysis has been applied to the bispectrum in [211]. The terms that yield the leading (momentum-independent) corrections to the primordial spectrum have been recently discussed in [80]. Note by the way that there are many different ways in which perturbation could break down. The authors of [212] argue for instance that in a nearly de Sitter universe certain second order perturbations may be as important as linear ones, which also implies a failure of linear perturbation theory.

The structure of this chapter, which is based on the paper [213], is as follows. In the next section we describe the relevant background to our problem and introduce the in-in formalism necessary to calculate corrections to the 2-point function of cosmological perturbations. In section 5.3 we compute the squared amplitude of tensor perturbations and we derive the results mentioned above. In section 5.4 we apply a similar analysis to the case of scalar perturbations, and obtain similar results. We conclude and discuss possible implications of our results in section 5.5.

5.2 Cosmological Perturbation Theory

5.2.1 The Inflating Background

Our starting point is a standard single-field inflation model. At sufficiently late times, the inflaton and gravity must be described by a low-energy effective action, whose leading terms are dictated by general covariance and the field content,

$$S_0 = \int d^4x\sqrt{-g}\left[\frac{M_{\rm p}^2}{2}R - \frac{1}{2}\partial_\mu\varphi\,\partial^\mu\varphi - V(\varphi)\right] . \tag{5.4}$$

In an EFT context, the action should also contain additional terms suppressed by powers of a dimensionful scale, which we assume to be of the order of the Planck mass M_{p}. Our goal is to determine the point beyond which such higher-dimensional operators produce corrections to the two-point function of cosmological perturbations that cannot be neglected. Our considerations can be readily generalized to cases in which the suppression scale of the higher-dimensional operators is not the Planck mass, but any other scale.

If the potential $V(\varphi)$ is sufficiently flat, at least in a certain region in field space, there exist inflationary solutions, along which a homogeneous scalar field $\varphi(\tau, \mathbf{x}) = \varphi_0(\tau)$ slowly rolls down the potential and spacetime is spatially homogeneous, isotropic and flat[1] ,

$$g^{(0)}_{\mu\nu} \equiv a^2(\tau)\eta_{\mu\nu}, \tag{5.5}$$

where $\eta_{\mu\nu}$ is the Minkowski metric and τ denotes conformal time. A model-independent measure of the slowness of the inflation is given by the slow-roll parameter

$$\varepsilon \equiv -\frac{H'}{aH^2}, \tag{5.6}$$

where $H \equiv a'/a^2$ is the Hubble parameter and a prime denotes a derivative with respect to conformal time. During slow-roll, ε is nearly constant, and to lowest order in slow-roll parameters its time derivative can be neglected. Throughout this chapter we work to leading non-vanishing order in the slow-roll expansion.

5.2.2 Cosmological Perturbations

Let us now consider cosmological perturbations around the homogeneous and isotropic background described above. Writing $\varphi = \varphi_0 + \delta\varphi$ and $g_{\mu\nu} = g^{(0)}_{\mu\nu}(\tau) + \delta g_{\mu\nu}(\tau, \mathbf{x})$, and substituting into equation (5.4), we can expand the action S_0 up to the desired order in the fluctuations $\delta\varphi$ and $\delta g_{\mu\nu}$,

$$S_0[\varphi, g_{\mu\nu}] = \delta_0 S_0 + \delta_1 S_0 + \delta_2 S_0 + \cdots . \tag{5.7}$$

[1]Strictly speaking, inflation generates an *almost* perfectly flat spacetime. However, tiny departures from perfect flatness will not play any role in what follows, since we will be interested in the small-scale regime at which even a spatially curved spacetime looks Euclidean.

The lowest order term $\delta_0 S_0$ does not contain any fluctuations and describes the inflating background; the linear term $\delta_1 S_0$ vanishes because it corresponds to the first variation of the action along the background solution, and the quadratic part of the action $\delta_2 S_0$ describes the free dynamics of the perturbations. The latter is what we need in order to calculate the primordial spectrum of fluctuations. As we pointed out in Sec.2.5, tensor and scalar perturbations are decoupled to quadratic order, so we may study them separately. It follows from our analysis in Sec. 2.5 that, to leading order in the slow-roll expansion, both scalar and tensor modes are described by the same action, namely

$$\delta_2 S_0 = \frac{1}{2} \int d\tau d^3x \left[(v')^2 - (\partial_i v)^2 + \frac{a''}{a} v^2 \right] . \tag{5.8}$$

Therefore, the mode functions $v_{\mathbf{k}}$ of both scalar and tensor perturbations satisfy the same equation of motion during inflation to leading order in the slow-roll expansion, namely

$$v_{\mathbf{k}}'' + \left[k^2 - \frac{a''}{a} \right] v_{\mathbf{k}} = 0 . \tag{5.9}$$

This equation has a unique solution for appropriate initial conditions. The conventional choice is the Bunch-Davies or adiabatic vacuum, whose mode functions obey

$$v_{\mathbf{k}}(\tau) \quad \xrightarrow{|k\tau| \gg 1} \quad \frac{e^{-ik\tau}}{\sqrt{2k}} \left[1 + \mathcal{O}\left(\frac{1}{k\tau} \right) \right] . \tag{5.10}$$

Because we are only interested in the sub-horizon limit, this is all we need to know about the mode functions. In particular, because the behavior of the mode functions in the short-wavelength limit does not depend on the details of inflation, our results are also insensitive to the particular form of the inflaton potential.

5.2.3 Quantum Fluctuations and the *in-in* Formalism

In order to study the properties of cosmological modes in the short-wavelength regime, we concentrate on the two-point function of the field v,

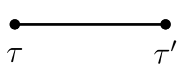
$$\langle v^*(\tau, \mathbf{k}) v(\tau, \mathbf{k}) \rangle \equiv \langle 0, in | v^*(\tau, \mathbf{k}) v(\tau, \mathbf{k}) | 0, in \rangle \ , \tag{5.11}$$

where $|0, in\rangle$ is the quantum state of the perturbations, which we assume to be the Bunch-Davies vacuum. The two-point function characterizes the mean square amplitude of cosmological perturbation modes, and differs from the power spectrum just by a normalization factor. Note that in an infinite universe, the two-point function is proportional to a momentum-conserving delta function, which in a spatially compact universe is replaced by a Kronecker delta.

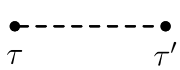
In the *in-in* formalism (see [214] for a clear and detailed exposition) the two-point function can be expressed as a path integral,

$$\langle v^*(\tau, \mathbf{k}) v(\tau, \mathbf{k}) \rangle = \int \mathcal{D}v_+ \mathcal{D}v_- \, v_+^*(\tau, \mathbf{k}) v_-(\tau, \mathbf{k}) \exp\left(iS_{\text{free}}[v_+, v_-]\right) \exp\left(iS_{\text{int}}[v_+]\right) \exp\left(-iS_{\text{int}}[v_-]\right) \ , \tag{5.12}$$

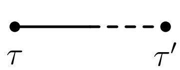
where S_{free} is quadratic in the fields, and S_{int} contains not just the remaining cubic and higher order terms in the action, but also any other quadratic terms we may decide to regard as perturbations. Note that there are two copies of the integration fields v_- and v_+, because we are calculating expectation values, rather than *in-out* matrix elements. This path integral expression is very useful to perturbatively expand the expectation value in powers of any interaction. In particular, each contribution can be represented by a Feynman diagram, with vertices drawn from the terms in S_{int} and propagators determined by the free action S_{free}. In our case, the latter are given by



$$= \int \mathcal{D}v_+ \mathcal{D}v_- \, v_+^*(\tau, \mathbf{k}) v_+(\tau', \mathbf{k}) \exp(iS_{\text{free}}) \approx \frac{e^{-ik|\tau-\tau'|}}{2k}, \tag{5.13a}$$



$$= \int \mathcal{D}v_+ \mathcal{D}v_- \, v_-^*(\tau, \mathbf{k}) v_-(\tau', \mathbf{k}) \exp(iS_{\text{free}}) \approx \frac{e^{ik|\tau-\tau'|}}{2k}, \tag{5.13b}$$



$$= \int \mathcal{D}v_+ \mathcal{D}v_- \, v_+^*(\tau, \mathbf{k}) v_-(\tau', \mathbf{k}) \exp(iS_{\text{free}}) \approx \frac{e^{ik(\tau-\tau')}}{2k}, \tag{5.13c}$$

which we quote here just in the sub-horizon limit. Note that to first order in S_{int} there are two vertices, one that contains powers of v_+ and one that contains powers of v_-; the associated coefficients just differ by an overall sign.[2]

[2] The quadratic action S_{free} enforces $v_+(\mathbf{k}) = v_-(\mathbf{k})$ at time τ. Hence, we could replace $v_+^*(\tau, \mathbf{k}) v_-(\tau, \mathbf{k})$ by $v_+^*(\tau, \mathbf{k}) v_+(\tau, \mathbf{k})$ or $v_-^*(\tau, \mathbf{k}) v_-(\tau, \mathbf{k})$ inside the path integral (5.12). Our choice re-

As a simple example, let us calculate the value of the two-point function in the short-wavelength limit to zeroth order in the interactions. Using the definition (5.12) and equation (5.13c), we find

$$\langle v^*(\tau, \mathbf{k}) v(\tau, \mathbf{k}) \rangle = \underset{\tau}{\bullet}\!\!\text{———}\,\text{- - -}\!\!\underset{\tau}{\bullet} \approx \frac{1}{2k} \qquad (|k\tau| \gg 1), \tag{5.14}$$

which is the well-known and standard short-wavelength limit result. In this regime, the two-point function is hence the inverse of the dispersion relation, since the latter determines the appropriate boundary conditions for the mode functions, as in equation (5.10).

In the next two sections we use the path integral (5.12) to calculate the corrections to the two-point function coming from higher-order operators in the action. These can be interpreted as corrections to the dispersion relation, even though in the presence of such terms the mode equations generally contain higher order time derivatives. In any case, a significant disagreement between the calculated two-point function and the lowest order result (5.14) points to the lack of self-consistency of our quantization procedure, and signals the breakdown of cosmological perturbation theory.

5.3 The Limits of Perturbation Theory: Tensors

The lowest order action (5.4) contains the leading terms that describe the dynamics of the inflaton and its perturbations. However, as we have noted, in an EFT approach the action generically contains all possible terms compatible with general covariance and any other symmetry of the theory. Here, for simplicity, we assume invariance under parity, an approximate shift symmetry of the inflaton, and a discrete $\mathbb{Z}_2$ symmetry $\varphi \to -\varphi$. Thus, all possible effective corrections to the action (5.4) can be built from the metric $g_{\mu\nu}$, the Riemann tensor $R_{\mu\nu\lambda\rho}$, the covariant derivative ∇_μ and an even number of scalar fields φ. In what follows, we consider these additional

---

moves the apparently ill-defined corrections we otherwise obtain when higher-order time derivatives act on the time-ordered products in equations (5.13a) and (5.13b). These ill-defined corrections can also be eliminated by field-redefinitions, a procedure that leads to the same corrections we find using our choice of field insertions.

terms and compute the corrections they induce on the two-point function of tensor perturbations in the short-wavelength limit. This allows us to determine the regime in which additional terms in the action cannot be neglected, and hence, the range over which cosmological perturbation theory is applicable. The reader not interested in technical details may skip directly to section 5.3.3, where we collect and summarize our results.

5.3.1 Dimension Four Operators

We begin our analysis by considering all dimension four operators, which appear in the action multiplied by dimensionless coefficients. On dimensional grounds, we expect these to yield corrections to the two-point function that are suppressed by only two powers[3] of $M_{\rm p}$. These operators will also help us to illustrate our formalism and discuss some of the important issues related to our calculation.

Any generally covariant dimension four effective correction must be of the form

$$S_1 \equiv S_\alpha + S_\beta = \int \sqrt{-g}\left(\alpha R^2 + \beta C^2\right), \tag{5.15}$$

where C^2 is the square of the Weyl tensor,

$$C^2 = R_{\mu\nu\lambda\rho}R^{\mu\nu\lambda\rho} - 2R_{\mu\nu}R^{\mu\nu} + \frac{1}{3}R^2, \tag{5.16}$$

and the dimensionless couplings α and β are assumed to be of order one. Note that we have ignored total derivatives like the Gauss-Bonnet term, since they do not lead to any corrections in perturbation theory. The Levi-Civita tensor cannot appear in the action because we assume invariance under parity.

We start by substituting the perturbed metric (2.32) into equation (5.15) and expanding up to second order in h_{ij}. Using the modified background equations and the relation $v^s_{\mathbf{k}} \equiv aM_{\rm p}h^s_{\mathbf{k}}/2$ to express the tensor perturbations in terms of the variable

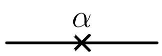
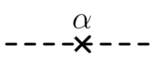
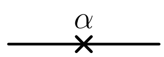
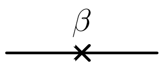
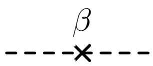
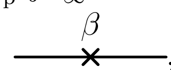
[3]Dimension six operators quadratic in φ also contribute at this order; we consider them later.

v, we obtain in the sub-horizon limit

$$\delta_2 S_\alpha = \frac{\alpha}{2M_\mathrm{p}^2}\sum_{\mathbf{k}}\int d\tau'\left[-\frac{6a''}{a^3}v_{\mathbf{k}}(v''_{-\mathbf{k}}+k^2 v_{-\mathbf{k}})-\frac{6a''}{a^3}(v''_{\mathbf{k}}+k^2 v_{\mathbf{k}})v_{-\mathbf{k}}\right], \tag{5.17}$$

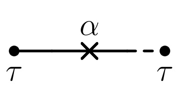
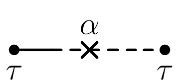
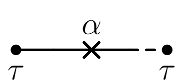
$$\delta_2 S_\beta = \frac{\beta}{M_\mathrm{p}^2}\sum_{\mathbf{k}}\int d\tau'\left[\frac{1}{a}(v''_{\mathbf{k}}+k^2 v_{\mathbf{k}})-2\,aH\left(\frac{v_{\mathbf{k}}}{a}\right)'\right]\times$$
$$\times\left[\frac{1}{a}(v''_{-\mathbf{k}}+k^2 v_{-\mathbf{k}})-2\,aH\left(\frac{v_{-\mathbf{k}}}{a}\right)'\right].$$

From these expressions, it is easy to derive the rules for the vertices


$$\approx \frac{i\alpha}{M_\mathrm{p}^2}\int_{-\infty}^{\tau} d\tau'\left[-\frac{6a''}{a^3}\left(\overrightarrow{\partial}^2_{\tau'}+k^2\right)-\left(\overleftarrow{\partial}^2_{\tau'}+k^2\right)\frac{6a''}{a^3}\right]$$

$$= -$$

$$\tag{5.18a}$$

$$\approx \frac{2i\beta}{M_\mathrm{p}^2}\int_{-\infty}^{\tau}\frac{d\tau'}{a^2}\left[\left(\overleftarrow{\partial}^2_{\tau'}+k^2\right)-\overleftarrow{\partial}_{\tau'}2\,aH\right]\left[\left(\overrightarrow{\partial}^2_{\tau'}+k^2\right)-2\,aH\overrightarrow{\partial}_{\tau'}\right]$$

$$= -$$

, (5.18b)

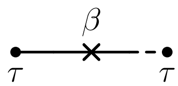
where the arrows indicate the propagator on which the derivatives act (because the vertex is quadratic, two propagators meet at the vertex.)

We are now ready to consider the correction due to the square of the Ricci scalar. The first order correction to the two-point function is given by the sum of the following two graphs,


$$\approx \frac{i\alpha}{M_\mathrm{p}^2}\int_{-\infty}^{\tau} d\tau'\left\{\frac{i\delta(\tau-\tau')}{2k}\frac{6a''}{a^3}\right\}\approx -\frac{\alpha}{2k}\frac{12H^2}{M_\mathrm{p}^2},$$

$$= \Big($$

$$\Big)^*, \tag{5.19}$$

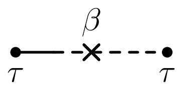
where we have used the fact that $a''/a^3 \approx 2H^2$ to lowest order in slow-roll. Notice that the operator $\left(\overrightarrow{\partial}^2_{\tau'}+k^2\right)$ acting on the time-ordered propagators (5.13a) or (5.13b) produces a delta function, since both are Green's functions. On the other hand, when the same operator acts on the propagator (5.13c) we get zero, because the latter is a regular solution of the free equation of motion (5.9) in the sub-horizon limit. This remark will turn out to be very useful when studying higher dimension operators.

We can now consider the correction due to the square of the Weyl tensor. In this case, the first order contribution is given by the sum of the following two graphs

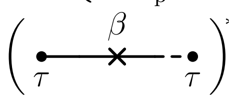


$$= -\frac{2i\beta}{M_{\rm p}^2}\int_{-\infty}^{\tau} d\tau' \left\{\delta(\tau-\tau')aH + H^2 e^{2ik(\tau'-\tau)}\right\}$$

$$\approx -\frac{\beta}{2k}\left\{\frac{4iHk_{ph}}{M_{\rm p}^2}+\frac{2H^2}{M_{\rm p}^2}\right\} \qquad (5.20)$$

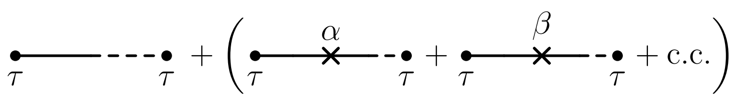


$$= \left(\text{[graph]} \right)^{*}.$$



Note that the imaginary parts cancel once we sum the two graphs. This result is quite general and ensures that only corrections with even powers of k_{ph} appear.

In conclusion, we have found that the leading corrections due to dimension four operators result in a two-point function which in the short-wavelength limit has the form



$$\approx \frac{1}{2k}\left[1-(24\alpha+4\beta)\,\frac{H^2}{M_{\rm p}^2}\right].$$

Thus, when H becomes of order $M_{\rm p}$, these corrections become as important as the leading result, and standard cosmological perturbation theory ceases to be applicable, as the reader may have expected.

5.3.2 Higher Dimension Operators

We would now like to consider a generic operator of dimension $2d+4$, suppressed by a factor of order $1/M_{\rm p}^{2d}$. However, it turns out that considering directly corrections to the action (5.8) for the perturbations is a much more efficient approach than starting from generally covariant effective terms added to the Lagrangian (5.4), particularly if we are interested in identifying the dominant corrections in the sub-horizon limit. Hence, we shall focus directly on modifications to the action for the perturbations. A related approach has been described in [98].

Dimensional analysis implies that any operator of dimension $2d+4$, quadratic in the dimensionless tensor perturbations h_{ij} and proportional to $2f$ powers of the inflaton field φ must contain $2d-2f+4$ derivatives ∂_μ acting on h_{ij}, $\varphi_0(\tau)$ and $a(\tau)$. The derivatives can be distributed and contracted using the Minkowski metric in many different ways,[4] but each of these terms can be schematically represented as

$$M_{\rm p}^{-2d-2}\left(\partial^{2n+m+p}\,[a,\varphi_0]\right)\,\left(\partial^{2q+m+r}\,v\right)\,\left(\partial^{2s+p+r}\,v\,\right), \tag{5.21}$$

where $\partial^i[a,\varphi_0]$ is just a symbol that represents any combination of i derivatives acting on a's and φ_0's. One such term would have $2n+m+p$ derivatives acting on one or more factors of a or φ_0, $2q+m+r$ derivatives acting on one field v and $2s+p+r$ acting on the other v. In particular, $2n$ of the derivatives acting on the scale factor or the background field are contracted among themselves while m and p of them are contracted with derivatives acting on, respectively, the first and second field v. The derivatives acting on the fields v are organized in a similar way.

Let us illustrate this notation by considering a term with $p=s=f=0$, $m=n=q=1$ and $r=2$. Dimensional analysis implies that $d=3$, and thus the explicit form of such a term would be

$$M_{\rm p}^{-8}\partial_{2+1+0}\,[a]\,\partial_{2+1+2}\,v\,\partial_{0+0+2}\,v = M_{\rm p}^{-8}\partial_\mu\partial^\mu\,\partial_\nu[a]\,\partial_\lambda\partial^\lambda\,\partial^\nu\,\partial_\alpha\partial_\beta\,v\,\partial^\alpha\partial^\beta\,v\ , \tag{5.22}$$

where $\partial_\mu\partial^\mu\partial_\nu[a]$ denotes all possible ways to construct a term with three derivatives of the scale factor, with the given tensor structure.

The first step to estimate the leading correction due to a term of the form (5.21) is realizing that this can always be expressed as a linear combination of terms of the form

$$M_{\rm p}^{-2d-2}\partial^{2j+l}\,[a,\varphi_0]\,\partial^{2m+l}v\,\partial_\mu\partial^\mu v, \tag{5.23}$$

plus, possibly, a term with no derivatives acting on v, which in any case gives a contribution that is always subdominant in the sub-horizon limit. For a proof that

[4]The reader may think that derivatives could be contracted not only among each other with the Minkowski metric, but also by using the additional tensor structure provided by the metric perturbations $\delta g_{\mu\nu}/a^2 = h_{ij}\eta_{\mu i}\eta_{\nu j}$. However, it turns out that $\left(\delta g_{\mu\nu}/a^2\right)\partial^\nu a = h_{ij}\eta_{\mu i}\partial^j a = 0$ and, since h_{ij} is transverse,

$$\partial^\mu\left(\delta g_{\mu\nu}/a^2\right) = \eta_{\nu j}\partial^i h_{ij} = 0\ .$$

Thus, we get a non-vanishing contribution only when we contract derivatives with the Minkowski metric while the factors $\eta_{\mu i}\eta_{\nu j}$ are contracted among each other yielding an irrelevant overall factor.

this decomposition is always possible, we refer the reader to the Appendix. In what follows, we therefore restrict ourselves to terms of the form (5.23).

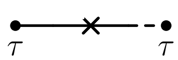
Dimensional analysis requires that the indexes j, l and m in equation (5.23) obey

$$j + l + m = d - f + 1 \ . \tag{5.24}$$

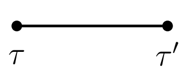
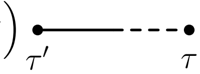
Furthermore, since $d^n\varphi_0/d\tau^n \propto \sqrt{2\varepsilon}\, M_\mathrm{p}\, a^n H^n$ and $d^n a/d\tau^n \propto a^{n+1}H^n$, each field φ_0 yields a factor of M_p, while each derivative acting on it or on the scale factor results in a factor of H to leading order in slow-roll. Finally, the l partial derivatives ∂_μ acting on v that are contracted with derivatives acting on a or φ_0 can be turned into derivatives with respect to τ only. Thus, (5.23) can be re-written as

$$M_\mathrm{p}^{-2d+2f-2} f(a) H^{2j+l}\, \Box^m \partial_\tau^l v\, \Box\, v \ , \tag{5.25}$$

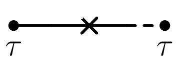
where we have defined $\Box \equiv \partial_\mu \partial^\mu$; the corresponding correction to the two-point function is schematically given by



$$= \frac{i}{M_\mathrm{p}^{2d-2f+2}} \int_{-\infty}^{\tau} d\tau' f(a) H^{2j+l} \times \tag{5.26}$$

$\times$  $\left(\overleftarrow{\Box}^m \overleftarrow{\partial}^{\,l}_{\tau'} \overrightarrow{\Box} + \overleftarrow{\Box} \overrightarrow{\partial}^{\,l}_{\tau'} \overrightarrow{\Box}^m \right)$ 

plus the complex conjugate of this graph. Because (5.13c) satisfies the free equation of motion, this correction is non–vanishing only when the index m is equal to zero, and in this case we obtain



$$= \frac{1}{M_\mathrm{p}^{2d-2f+2}} \int\limits_{-\infty}^{\tau} d\tau' f(a) H^{2j+l} \delta(\tau - \tau') \frac{(ik)^l}{2k} = \frac{f(a)}{2k} \frac{H^{2j+l}(ik)^l}{M_\mathrm{p}^{2d-2f+2}}. \tag{5.27}$$

The leading correction in the short-wavelength limit is the one with the maximum number of powers of k. According to equation (5.24), this maximum number simply equals $d - f + 1 \equiv l_{max}$, and it corresponds to the case in which $j = m = 0$. Thus, if $d - f$ is odd, l_{max} is even and the leading correction is simply given by

$$\delta\langle v^*(\mathbf{k}) v(\mathbf{k})\rangle \propto \frac{1}{2k} \left(\frac{H}{M_\mathrm{p}}\right)^{d-f+1} \left(\frac{k_{ph}}{M_\mathrm{p}}\right)^{d-f+1} \qquad (d - f \text{ odd}) \ , \tag{5.28}$$

since each factor of k/M_p must be accompanied by a factor of a to render the spatial momentum physical. On the other hand, if $d-f$ is even, l_{max} as defined above is odd

$d-f$	Leading correction	$d-f$	Leading correction
0	$H^2/M_{\rm p}^2$	4	$H^6k_{ph}^4/M_{\rm p}^{10}$
1	$H^2k_{ph}^2/M_{\rm p}^4$	5	$H^6k_{ph}^6/M_{\rm p}^{12}$
2	$H^4k_{ph}^2/M_{\rm p}^6$	6	$H^8k_{ph}^6/M_{\rm p}^{14}$
3	$H^4k_{ph}^4/M_{\rm p}^8$	7	$H^8k_{ph}^8/M_{\rm p}^{16}$

Table 5.1: Leading corrections to the gravitational wave two-point functions in the short-wavelength limit.

and the term with the highest number of powers of k is purely imaginary. As we have seen in the previous section, such a term disappears when we add the contribution from the complex conjugate graph. Therefore, the leading correction corresponds to the largest *even* value of l, which turns out to be $l_{max} = d-f$, and is therefore given by

$$\delta\langle v^*(\mathbf{k})v(\mathbf{k})\rangle \propto \frac{1}{2k}\left(\frac{H}{M_{\rm p}}\right)^{d-f+2}\left(\frac{k_{ph}}{M_{\rm p}}\right)^{d-f} \qquad (d-f \text{ even}) \ . \tag{5.29}$$

Equations (5.28) and (5.29) represent the main results of this section: they express the leading corrections to the two-point function (in the sub-horizon limit) associated with a generic operator of dimension $2d+4$ containing $2f$ powers of the inflaton field. Since we have assumed an approximate shift invariance of the inflaton, the total number of derivatives, $2d-2f+2$, must be greater or equal than the number of fields $2f$, which in turn implies that $d \geq f$. Thus, we can label all the possible corrections with the non-negative index $d-f$. Their magnitude is given in Table 5.1 for the first eight values of $d-f$. Note that corrections with $d-f=0$ arise from the operators identified by Weinberg in [80]. The leading momentum-dependent corrections are given by operators with $d-f=1$.

So far, we have calculated the largest possible corrections to the two-point function in the sub-horizon limit given a certain value of $d-f$. However, the reader might still wonder whether such terms can be actually obtained from a covariant action. Employing the same technique we used to study the impact of the lowest order terms, it is indeed possible to show—after some rather lengthy calculations—that the following family of covariant terms generates this kind of contributions,

$$\begin{aligned} d-f=0: &\qquad R^{\mu\nu}\,R_{\mu\nu} \\ d-f=1: &\qquad (\nabla^{\alpha}R^{\mu\nu})\,\nabla_{\alpha}R_{\mu\nu} \\ d-f=2: &\qquad (\nabla^{\alpha}\nabla^{\beta}R^{\mu\nu})\,\nabla_{\alpha}\nabla_{\beta}R_{\mu\nu} \\ \vdots\qquad &\qquad\qquad \vdots \qquad\qquad . \end{aligned} \tag{5.30}$$

It can be also verified that the $d-f=1$ term yields a correction to the two-point function proportional to the slow-roll parameter ε, and given the structure of this family of operators, we anticipate the remaining terms to share the same slow-roll suppression.

5.3.3 The Three-Momentum Scale Λ

The corrections to the two-point function are functions of two dimensionful parameters, H and k_{ph}. For our purposes, it is convenient to organize these corrections in powers of k_{ph}. Thus, following Table 5.1, and reintroducing the subleading terms that we previously neglected, we find that the two-point function is

$$\langle v^*(\mathbf{k})v(\mathbf{k})\rangle \approx \frac{1}{2k}\Bigg[\left(1+\alpha_{20}\frac{H^2}{M_{\mathrm{p}}^2}+\cdots\right)+\left(\alpha_{22}\frac{H^2}{M_{\mathrm{p}}^2}+\alpha_{42}\frac{H^4}{M_{\mathrm{p}}^4}+\cdots\right)\frac{k_{ph}^2}{M_{\mathrm{p}}^2}+ \\ +\left(\alpha_{44}\frac{H^4}{M_{\mathrm{p}}^4}+\alpha_{64}\frac{H^6}{M_{\mathrm{p}}^6}+\cdots\right)\frac{k_{ph}^4}{M_{\mathrm{p}}^4}+\cdots\Bigg]. \tag{5.31}$$

The coefficient α_{20} is of order one, while all the α_{nn} with $n\geq 2$ are of order ε, as the family of covariant terms (5.31) suggests. At the end of Section 5.4 we provide further evidence supporting this claim.

In order for Equation (5.31) to be a valid perturbative expansion, every correction term must be much smaller than one. Because α_{20} is of order one, this implies the condition

$$\frac{H}{M_{\mathrm{p}}}\ll 1, \tag{5.32}$$

which must hold for all values of k_{ph}. Equation (5.31) then shows that if condition (5.32) is satisfied, the corrections to the two-point function remain small even for

$k_{ph} \approx M_{\rm p}$. In fact, to leading order in $H/M_{\rm p}$ we can rewrite equation (5.31) as

$$\langle v^*(\mathbf{k})v(\mathbf{k})\rangle \approx \frac{1}{2k}\left[1+\alpha_{22}\frac{k_{ph}^2}{\Lambda^2}+\alpha_{44}\frac{k_{ph}^4}{\Lambda^4}+\cdots\right] , \tag{5.33}$$

where we have introduced the effective cut-off

$$\Lambda \approx \frac{M_{\rm p}^2}{H}. \tag{5.34}$$

Equations (5.33) and (5.34) are the main result of these chapter and were first derived in the paper [213]. For $k_{ph} \ll \Lambda$, all the corrections are strongly suppressed and can thus be neglected. However, at $k_{ph} \approx \Lambda$, all the corrections become of order ε, the asymptotic series breaks down, and the effective theory ceases to be valid. As we discuss below, this value of Λ should be understood as an upper limit on the validity of cosmological perturbation theory.

To conclude this section, let us briefly comment on the effects of terms that break the shift symmetry. Because the only difference is that these terms contain undifferentiated scalars, any such correction can be cast as a generally covariant term that respects the symmetry, multiplied by a power of the dimensionless ratio $\varphi/M_{\rm p}$. Hence, these terms introduce corrections to the two-point function of the form we have already discussed, but with coefficients α_{ij} that can now depend on arbitrary powers of the background field φ_0,

$$\alpha_{ij} = \alpha_{ij}^{(0)} + \alpha_{ij}^{(1)}\frac{\varphi_0}{M_{\rm p}} + \alpha_{ij}^{(2)}\left(\frac{\varphi_0}{M_{\rm p}}\right)^2 + \cdots . \tag{5.35}$$

Therefore, in the absence of any mechanism or symmetry that keeps the coefficients $\alpha_{ij}^{(1)}, \alpha_{ij}^{(2)}, \ldots$ small (e.g. an approximate shift symmetry), such an expansion looses its validity for $\varphi_0 > M_{\rm p}$, regardless of the value of k_{ph}. If, on the other hand, equation (5.35) is a sensible expansion, and $\alpha_{ij}^{(0)}$ is much greater than $\alpha_{ij}^{(1)}, \alpha_{ij}^{(2)}$, ..., then we can effectively assume that the shift-symmetry is exact, and perturbations theory breaks down again at $k_{ph} \approx \Lambda$.

5.3.4 Loop Diagrams and Interactions

Our analysis so far has concentrated only on tree-level corrections to the two-point function, which arise just from the quadratic terms in the action. Cubic and higher order interactions also contribute to the two-point function, but their contribution is obscured by the appearance of divergent momentum integrals in loops. Even in a non-renormalizable theory like GR, at any order in the derivative expansion it is still possible to cancel these divergences by renormalizing a finite number of parameters, provided that all terms consistent with the symmetries of the theory are included in the Lagrangian [54]. In practice, this cancellation is due to the presence of appropriate counter-terms in the Lagrangian. For this reason, divergent integrals in loop diagrams are rather harmless. They yield corrections of the same structure as tree-level diagrams, modulo a (mild) logarithmic running of their values with scale [215]. Hence, we do not expect this type of contributions to drastically change our conclusions, though we should emphasize that this is just an expectation. Thus, strictly speaking, equation (5.34) is just an upper limit for the scale beyond which the corrections to the two-point function remain small and one can trust cosmological perturbation theory.

5.4 The Limits of Perturbation Theory: Scalars

We now turn our attention to corrections to the two-point function of scalar perturbations. Despite some complications that are particular to the this sector, the method developed in the previous section can be easily extended to scalars.

To this end, let us consider the action $S = S_0 + \lambda S_1$, where S_0 is the action (5.4) describing a scalar field minimally coupled to Einstein gravity, while S_1 is a generic generally covariant correction suppressed by a coupling $\lambda \sim 1/M_\mathrm{p}^{2d}$. As we pointed out in the previous section, S_1 generically involves contractions of the Riemann tensor $R_{\mu\nu\lambda\rho}$ and the covariant derivative ∇_μ as well as the scalar field φ. In order to compute the resulting first order contribution to the two-point function for v, we insert the perturbed metric and the perturbed inflaton field into the action S and expand up to second order in v.

Expanding the leading action S_0 to quadratic order in the perturbations, we obtain the free action (5.8). The additional quadratic terms stemming from S_1 must be appropriately contracted expressions containing partial derivatives of the perturbation variable v, the scale factor a and the background field φ_0. In what follows, it will be convenient to work in *spatially flat gauge* such that the perturbed metric reads

$$ds^2 = a^2(\eta)\left[-(1+2\phi)d\eta^2 + 2\partial_i B dx^i d\eta + \delta_{ij}dx^i dx^j\right] \ . \tag{5.36}$$

In this gauge the variable v is simply proportional to the inflation fluctuation, i.e. $v = a\delta\varphi$ (see Eq. (2.35)). By solving the constraint imposed by Einstein equations, one can express both ϕ and B in terms of v. To leading order in the slow-roll expansion, the solutions for the Fourier modes are (see e.g. [216])

$$\phi_k = \sqrt{\frac{\varepsilon}{2}}\frac{v_{\mathbf{k}}}{aM_{\rm p}} \ , \tag{5.37}$$

$$B_k = \sqrt{\frac{\varepsilon}{2}}\frac{1}{M_{\rm p}k^2}\left(\frac{v_{\mathbf{k}}}{a}\right)' \ . \tag{5.38}$$

In the case of scalar perturbations, the field v can arise from fluctuations of the scalar field, $\delta\varphi = v/a$, or from fluctuations of the metric,

$$\delta g^{\mu\nu} = -\frac{a\sqrt{2\varepsilon}}{M_{\rm p}}\left[2v_{\mathbf{k}}\delta^\mu{}_0\delta^\nu{}_0 + \frac{ik_j}{k^2}\left(v_{\mathbf{k}}' - aHv_{\mathbf{k}}\right)\left(\delta^\mu{}_j\delta^\nu{}_0 + \delta^\mu{}_0\delta^\nu{}_j\right)\right] \equiv \frac{a\sqrt{2\varepsilon}\,\mathcal{V}^{\mu\nu}}{M_{\rm p}} \ . \tag{5.39}$$

Thus, unlike the case of tensor perturbations, $\delta g^{\mu\nu}$ provides an additional tensor structure that can be used to contract derivatives. We now show that such contractions yield terms where the derivatives acting on v or a are contracted with $\eta^{\mu\nu}$. This means that the argument in the previous section can be applied to scalar perturbations as well, yielding essentially the same results. We note that terms which contain only fluctuations coming directly from the scalar field do not present this problem, and can be easily written as in equation (5.21).

Let us first consider terms with only one factor of $\mathcal{V}^{\mu\nu}$. In this case, $\mathcal{V}^{\mu\nu}$ can be contracted either with $\eta_{\mu\nu}$, leading to $\mathcal{V}^{\mu\nu}\eta_{\mu\nu} = 2\,v$, or with two derivatives $\partial_\mu\partial_\nu$, resulting in

$$\partial_\mu\partial_\nu\mathcal{V}^{\mu\nu} = 2\left(\frac{\partial_\mu\partial^\mu a}{a}v - \frac{\partial_\mu a\,\partial^\mu a}{a^2}v + \frac{\partial_\mu a\,\partial^\mu v}{a}\right), \tag{5.40a}$$

$$\partial_\mu a\,\partial_\nu\mathcal{V}^{\mu\nu} = \partial_\mu a\,\partial^\mu v + \frac{\partial_\mu a\,\partial^\mu a}{a}v, \tag{5.40b}$$

$$\partial_\mu\partial_\nu[a,\varphi_0]\,\mathcal{V}^{\mu\nu} = 2\left(\partial_\mu\partial^\mu[a,\varphi_0]\right)v \ , \tag{5.40c}$$

where, again, the square brackets in the last line mean that the derivatives can act on *one or more* factors of a or φ_0. Thus, terms with only one factor of $\mathcal{V}_{\mu\nu}$ do not present any problem since, as anticipated, all the derivatives are contracted with the inverse of the Minkowski metric.

Corrections which contain two factors of $\mathcal{V}^{\mu\nu}$, and are not products of terms in (5.40b), can always be recast as

$$\mathcal{V}_{\mu\nu}\mathcal{V}^{\mu\nu} = 2v^2 + \frac{2}{k^2}\left[\partial_\mu v\,\partial^\mu v - \frac{\partial_\mu a}{a}\,\partial^\mu(v^2) + \frac{\partial_\mu a\,\partial^\mu a}{a^2}v^2\right], \tag{5.41a}$$

$$\partial^\mu\mathcal{V}_{\mu\nu}\partial_\lambda\mathcal{V}^{\lambda\nu} = -\frac{1}{k^2}\partial_\mu v'\partial^\mu v' + \frac{1}{k^2}\left[-2\left(\frac{\partial_\mu\partial_\nu a}{a} - \frac{\partial_\mu a\,\partial_\nu a}{a^2}\right)v\,\partial^\mu\partial^\nu v\right. \tag{5.41b}$$
$$\left. 2\frac{\partial_\mu a\,\partial^\mu v}{a}\partial_\nu\partial^\nu v - \frac{\partial_\mu a\,\partial^\mu a}{a^2}\partial_\nu v\,\partial^\nu v + \left(\frac{\partial_\mu\partial^\mu a}{a} - \frac{\partial_\mu a\,\partial^\mu a}{a^2}\right)\frac{\partial_\nu a}{a}\,\partial^\nu(v^2)\right].$$

All the terms inside the square brackets become negligible in the sub-horizon limit, since their contribution is suppressed by an extra factor of $1/k^2$. The only term in which some derivatives are not contracted with the Minkowski inverse metric is the first one in equation (5.41b). However, the two derivatives with respect to conformal time result in a factor of k^2 which is precisely canceled by the extra factor $1/k^2$, and for all practical purposes such a term is equivalent to $\partial_\mu v\,\partial^\mu v$.

Therefore, we have demonstrated that terms quadratic in the scalar fluctuations can be schematically written as in equation (5.21). The remainder of the analysis then proceeds as for tensor perturbations, and effective corrections to scalar perturbations are thus also subdominant in the regime

$$H \ll M_{\mathrm{p}} \quad \text{and} \quad k_{ph} \ll \Lambda \sim \frac{M_{\mathrm{p}}^2}{H}\ . \tag{5.42}$$

Before concluding, we would like to address again whether the operators that we have considered can be actually obtained from generally covariant terms. In the case of scalar perturbations, it is indeed possible to show—after further rather lengthy calculations—that the following family of covariant terms generates the kind of corrections shown in Table 5.1,

$$\begin{aligned}
d-f=0: &\qquad R^{\mu\nu}\,(\nabla_\mu\varphi)\,\nabla_\nu\varphi\\
d-f=1: &\qquad R^{\mu\nu}\,(\nabla_\mu\varphi)\,\nabla_\nu\,\nabla_\gamma\,\nabla^\gamma\,\varphi\\
d-f=2: &\qquad (\nabla^\alpha\nabla^\beta R^{\mu\nu})\,(\nabla_\alpha\nabla_\mu\varphi)\,\nabla_\beta\nabla_\nu\,\varphi\\
d-f=3: &\qquad (\nabla^\alpha\nabla^\beta R^{\mu\nu})\,(\nabla_\alpha\nabla_\mu\varphi)\,\nabla_\beta\nabla_\nu\,\nabla_\gamma\,\nabla^\gamma\,\varphi\\
\vdots\quad &\qquad\qquad\vdots \qquad\qquad\qquad\qquad .
\end{aligned} \tag{5.43}$$

In order to illustrate how this happens, let us consider for example the $d-f=1$ term. It contains, among many other terms a factor

$$a^2\,R^{\mu\nu}\,\partial_\mu\,\delta\varphi\,\partial_\nu\,\partial_\gamma\,\partial^\gamma\,\delta\varphi \supset \frac{2\varepsilon}{a^6}\,\partial^\mu a\,\partial^\nu a\,\partial_\mu\,v\,\partial_\nu\,\partial_\gamma\,\partial^\gamma\,v \sim -\frac{2\varepsilon}{a^6}\,\partial^\mu a\,\partial^\nu a\,\partial_\mu\,\partial_\nu\,v\,\Box\,v+... \tag{5.44}$$

where, in the last step, we have neglected a subdominant contribution in the short-wavelength limit. The last term in (5.44) indeed generates a correction proportional to $H^2k_{ph}^2/M_{\rm p}^4$ and it is suppressed by one factor of the slow-roll parameter. It is relatively easy to verify that the corrections generated by the other members of the family (5.43) have the same slow-roll suppression, which strongly supports the assumption we made in the context of tensor perturbations.

5.5 Summary

The connection, through cosmological inflation, between physics on the smallest scales, described by quantum field theory, and that on the largest scales in the universe is one of the most profound aspects of modern cosmology. However, since inflation takes place at such early epochs, and magnifies fluctuations of such small wavelengths, it is important to establish the regime of validity of the usual formalism—that of semiclassical gravity, with quantum field theory assumed valid, and coupled to the minimal Einstein-Hilbert action—at those scales.

On general grounds, we expect the canonical approach to break down at ultra-short distances, where the operators that arise in an EFT treatment of the coupled metric-inflaton system become relevant. In this chapter we have calculated the impact of these higher-dimensional operators on the power spectrum at short wavelengths. In this way, we have been able to probe the regime in which the properties of the perturbations deviate from what is conventionally assumed. From a purely theoretical standpoint, these considerations are important if we are to understand the limits of applicability of cosmological perturbation theory. From an observational standpoint, cosmic microwave background measurements are becoming so precise that we may hope to use them to identify the signatures of new gravitational or field theoretic physics.

Our analysis has focused on tree-level corrections to the spectrum. Because we have essentially considered all possible generally covariant terms in the effective action, we expect to have unveiled the form of all possible corrections that are compatible with the underlying symmetries of the theory. It is however possible that loop diagrams yield additional corrections that we have not considered. In any case, our results indicate that cosmological perturbation theory does not apply all the way to infinitesimally small distances, $k_{ph} \to \infty$, and that, indeed, there is a physical spatial momentum Λ (or a physical length $1/\Lambda$) beyond which cosmological perturbation ceases to be valid. The scale at which perturbation theory breaks down has to be lower than

$$\Lambda \sim \frac{M_\mathrm{p}^2}{H}, \tag{5.45}$$

which, because of existing limits on the scalar to tensor power spectrum ratio [7], is at least 10^4 times the Planck scale.

These results have significant implications for the impact of trans-Planckian on the primordial spectrum of primordial perturbations, which typically is at most of order H/Λ [210]. Substituting the upper limit of Λ we have found, we obtain corrections of the order of H^2/M_p^2, which are likely to remain unobservable [201]. This value of Λ also solves a problem that was noticed in [217], namely, that in the presence of a Planckian cut-off, cosmological perturbations do not tend to decay into the Bunch-Davis vacuum (or similar states). In particular, to lowest order in perturbation theory, the transition probability from an excited state into the Bunch-Davis vacuum is significantly less than one for $\Lambda = M_{\rm p}$, but proportionally larger if Λ is given by (5.45). Ultimately, a large decay probability is what justifies the choice of the Bunch-Davies vacuum as the preferred initial state for the perturbations at scales below the cut-off, since, as we have found, our theories certainly lose their validity at momentum scales above the spatial momentum Λ.

Appendix 5.A Derivation of Equation (5.23)

In this appendix, we show how to integrate by parts every term of the form

$$\partial^{2n+m+p}\,[a,\varphi_0]\,\partial^{2q+m+r}\,v\,\partial^{2s+p+r}\,v \tag{5.46}$$

in order to express it as a linear combination of terms like

$$\partial^{2j+l}\,[a,\varphi_0]\,\partial^{2m+l}v\,\Box\,v \tag{5.47}$$

plus, possibly, a term with no derivatives acting on v. Notice that, for notational convenience, we have defined $\Box \equiv \partial_\mu \partial^\mu$. Of course, if the index q (or s) in equation (5.46) is not zero, we can easily integrate by parts $2q+m+r-2$ $(2s+p+r-2)$ times to get only terms of the form of that in equation (5.47). Therefore, in what follows we only consider terms with $q = s = 0$. In this case, we can always integrate by parts an appropriate number of times to get only terms for which $m = p$. Thus, without loss of generality, we can restrict ourselves to considering terms of the form

$$\partial^{2n+m+p}\,[a,\varphi_0]\,\partial^{m+r}\,v\,\partial^{p+r}\,v, \qquad (m=p)\ . \tag{5.48}$$

The derivatives acting on v that are contracted with derivatives acting on a or φ_0 can be systematically eliminated by repeated integrations by parts:

$$\begin{aligned}\partial^{2n+m+p}\,[a,\varphi_0]\,\partial^{m+r}\,v\,\partial^{p+r}\,v &\sim -\partial^{2n+m+(p-1)}\,[a,\varphi_0]\,\partial^{m+r}\,v\,\partial^{(p-1)+r}\,\Box\,v \\ &+\frac{1}{2}\partial^{2(n+1)+(m-1)+(p-1)}\,[a,\varphi_0]\,\partial^{(m-1)+(r+1)}\,v\,\partial^{(p-1)+(r+1)}\,v\ ,\end{aligned} \tag{5.49}$$

where we have denoted equivalence up to integration by parts with $\sim$. The first term on the right hand side can be cast in the form (5.47) by integrating by parts $(p-1)+r$

times, while the second one is of the form (5.48) with n and r (p and m) increased (decreased) by one. By iterating this procedure, we eventually obtain terms of the form

$$\partial^{2n}\,[a,\varphi_0]\,\partial^{r}\,v\,\partial^{r}\,v\ , \tag{5.50}$$

where now n and r have changed. Again, we can integrate by parts and obtain

$$\partial^{2n}\,[a,\varphi_0]\,\partial^{r}\,v\,\partial^{r}\,v \sim -\partial^{2n}\,[a,\varphi_0]\,\partial^{r-1}\,v\,\partial^{r-1}\,\Box\,v + \frac{1}{2}\partial^{2(n+1)}\,[a,\varphi_0]\,\partial^{(r-1)}\,v\,\partial^{(r-1)}\,v\ .$$

The first term on the right hand side can be re-written as (5.47) after $r-1$ integrations by parts, while the second term has the form (5.50) with n (r) increased (decreased) by one. Thus, by repeating this procedure we obtain many terms of the form (5.47) and we are eventually left with a term without derivatives acting on v. This completes our proof.

Chapter 6

Conclusions

In this thesis, we have applied EFT methods to the study of modified theories of gravity and the spectrum of primordial perturbations produced during inflation. In the first case, we pointed out that any modification of GR requires either violations of Lorentz invariance or additional degrees of freedom in the gravitational sector. Thus, in chapter 3 we extended the coset construction of Callan, Coleman, Wess and Zumino [19] to the describe gravitational theories in which the local Lorentz group is spontaneously broken down to any of its subgroups. We provided an explicit illustration of this formalism by considering the case in which rotations remain unbroken, and we proved that the Einstein-aether theory [75] is the most general low-energy effective theory of gravity which preserves local rotations.

In chapter 4, we considered instead the simplest possible modification of gravity which preserves Lorentz invariance and features a single additional scalar degree of freedom. We showed that gravitational interactions mediated by the additional scalar are bound to violate the weak equivalence principle (WEP), even if the classical action is chosen in such a way that point-like particles experience the same gravitational acceleration at the classical level. In this case, violations of the WEP are generated by loop corrections with at least one scalar or one graviton running in the loop. Therefore, quantum WEP violations are suppressed by at least two powers of the ratio $m/M_{\rm p}$ where m is the mass of the point-like particle. Although we have not worked out the implications of this result for macroscopic bodies, we conclude that quantum WEP violations are likely too small to be detectable.

Finally, in chapter 5 we turned to the study of primordial scalar and tensor perturbations generated during single-field inflation. We showed that Planck-suppressed irrelevant operators will introduce modifications of the tree-level spectrum of primordial perturbations which become relevant only when the physical wavenumber of the perturbations becomes of order $M_{\rm p}^2/H \gg M_{\rm p}$. This value is likely to lie beyond the regime of validity of the effective theory, implying that the impact of high energy physics on the spectrum of primordial perturbations is likely to be negligible.

In this dissertation, I have applied EFT techniques to study models of gravity and inflation which can equally describe the background evolution as well as the behavior of perturbations around it. However, given that many cosmological observations (e.g. CMB anisotropies, large scale structures, gravitational waves, ...) actually refer to properties of fluctuations around a given background, in future work we will turn to the study of EFTs of perturbations. This approach has been already pursued extensively to study perturbations generated during inflation [98], and we think that it could also be used to study fluctuations around more generic backgrounds in modified theories of gravity. For instance, this method could be used to examine fluctuations around spherically symmetric backgrounds in scalar-tensor theories.

The main idea behind this approach is that the scalar degree of freedom provides an additional geometrical structure which can be used to define a preferred coordinate system. In the case of spherically symmetric backgrounds, the hypersurfaces on which the scalar field remains constant define a preferred radial coordinate, thus leading to a spontaneous breaking of radial diffeomorphisms. Hence, the large-distance phenomenology of any modified theory of gravity which involves a single additional scalar degree of freedom can be captured by an effective action in which radial diffeomorphisms are broken. Within this framework, theoretical issues such as quantum and classical stability of fluctuations can be addressed in a model-independent way, leading to constraints on the pool of modified gravity models which are theoretically consistent.

Bibliography

[1] E. Hubble, Proc.Nat.Acad.Sci. **15**, 168 (1929).

[2] G. Gamow, Phys.Rev. **70**, 572 (1946).

[3] R. Alpher, H. Bethe, and G. Gamow, Phys.Rev. **73**, 803 (1948).

[4] R. A. Alpher and R. C. Herman, Phys.Rev. **74**, 1737 (1948).

[5] G. Gamow, Nature **162**, 680 (1948).

[6] G. Gamow, Phys.Rev. **74**, 505 (1948).

[7] E. Komatsu et al. (WMAP Collaboration), Astrophys.J.Suppl. **180**, 330 (2009), 0803.0547.

[8] G. F. Smoot, C. Bennett, A. Kogut, E. Wright, J. Aymon, et al., Astrophys.J. **396**, L1 (1992).

[9] URL http://www.mso.anu.edu.au/2dFGRS/.

[10] URL http://www.sdss.org/.

[11] V. C. Rubin and J. Ford, W. Kent, Astrophys.J. **159**, 379 (1970).

[12] D. Clowe, A. Gonzalez, and M. Markevitch, Astrophys.J. **604**, 596 (2004), astro-ph/0312273.

[13] D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, et al., Astrophys.J. **648**, L109 (2006), astro-ph/0608407.

[14] F. Zwicky, Helv.Phys.Acta **6**, 110 (1933).

[15] D. Bailin and A. Love, *Cosmology in Gauge Field Theory and String Theory* (IOP Publishing, 2004).

[16] A. G. Riess et al. (Supernova Search Team), Astron.J. **116**, 1009 (1998), `astro-ph/9805201`.

[17] S. Perlmutter et al. (Supernova Cosmology Project), Astrophys.J. **517**, 565 (1999), `astro-ph/9812133`.

[18] S. M. Carroll, V. Duvvuri, M. Trodden, and M. S. Turner, Phys.Rev. **D70**, 043528 (2004), `astro-ph/0306438`.

[19] J. Callan, Curtis G., S. R. Coleman, J. Wess, and B. Zumino, Phys.Rev. **177**, 2247 (1969).

[20] H. Georgi, Ann.Rev.Nucl.Part.Sci. **43**, 209 (1993).

[21] A. V. Manohar (1996), `hep-ph/9606222`.

[22] A. Pich (1998), `hep-ph/9806303`.

[23] I. Z. Rothstein (2003), `hep-ph/0308266`.

[24] D. B. Kaplan (2005), `nucl-th/0510023`.

[25] C. Burgess, Ann.Rev.Nucl.Part.Sci. **57**, 329 (2007), `hep-th/0701053`.

[26] J. Polchinski (1992), `hep-th/9210046`.

[27] J. F. Donoghue (1995), `gr-qc/9512024`.

[28] C. Burgess, Living Rev.Rel. **7**, 5 (2004), `gr-qc/0311082`.

[29] S. Weinberg, in *Conceptual foundations of quantum field theory* (1996), pp. 241–251, `hep-th/9702027`.

[30] M. Srednicki, *Quantum Field Theory* (Cambridge University Press, 2007).

[31] B. Delamotte (2007), cond-mat/0702365.

[32] G. Dvali, G. F. Giudice, C. Gomez, and A. Kehagias, JHEP **1108**, 108 (2011), 1010.1415.

[33] S. Weinberg, in *Understanding the Fundamental Constituents of Matter*, edited by A. Zichichi (Plenum Press, 1977).

[34] G. 't Hooft, NATO Adv.Study Inst.Ser.B Phys. **59**, 135 (1980).

[35] I. Low and A. V. Manohar, Phys.Rev.Lett. **88**, 101602 (2002), hep-th/0110285.

[36] S. Weinberg, *The quantum theory of fields. Vol. 2: Modern applications* (Cambridge University Press, 1996).

[37] P. Binetruy, *Supersymmetry: Theory, experiment and cosmology* (Oxford University Press, 2006).

[38] S. Weinberg, Rev.Mod.Phys. **61**, 1 (1989).

[39] S. Weinberg, Phys.Rev.Lett. **59**, 2607 (1987).

[40] A. Silvestri and M. Trodden, Rept.Prog.Phys. **72**, 096901 (2009), 0904.0024.

[41] S. Weinberg, *Cosmology* (Oxford University Press, 2008).

[42] V. F. Mukhanov and G. Chibisov, JETP Lett. **33**, 532 (1981).

[43] A. A. Starobinsky, Phys.Lett. **B117**, 175 (1982).

[44] A. H. Guth and S. Pi, Phys.Rev.Lett. **49**, 1110 (1982).

[45] S. Hawking, Phys.Lett. **B115**, 295 (1982), revised version.

[46] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, Phys.Rev. **D28**, 679 (1983).

[47] J. Khoury and G. E. Miller, Phys.Rev. **D84**, 023511 (2011), 1012.0846.

[48] D. Baumann, L. Senatore, and M. Zaldarriaga, JCAP **1105**, 004 (2011), 1101.3320.

[49] A. Joyce and J. Khoury, Phys.Rev. **D84**, 083514 (2011), 1107.3550.

[50] G. Geshnizjani, W. H. Kinney, and A. M. Dizgah, JCAP **1111**, 049 (2011), 1107.1241.

[51] S. Weinberg, Phys.Rev. **138**, B988 (1965).

[52] K. Hinterbichler (2011), 1105.3735.

[53] R. Feynman, F. Morinigo, W. Wagner, and e. Hatfield, B., *Feynman lectures on gravitation* (Addison-Wesley, 1996).

[54] S. Weinberg, *The quantum theory of fields. Vol. 1: Foundations* (Cambridge University Press, 1995).

[55] F. A. Berends, G. Burgers, and H. Van Dam, Z.Phys. **C24**, 247 (1984).

[56] N. Boulanger, T. Damour, L. Gualtieri, and M. Henneaux, Nucl.Phys. **B597**, 127 (2001), hep-th/0007220.

[57] N. Boulanger, Fortsch.Phys. **50**, 858 (2002), hep-th/0111216.

[58] S. C. Anco, Class.Quant.Grav. **19**, 6445 (2002), gr-qc/0303033.

[59] Y. Zinoviev, Nucl.Phys. **B770**, 83 (2007), hep-th/0609170.

[60] S. Hassan and R. A. Rosen, JHEP **1202**, 126 (2012), 11 pages, 1109.3515.

[61] G. Dvali, G. Gabadadze, and M. Porrati, Phys.Lett. **B485**, 208 (2000), hep-th/0005016.

[62] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, Phys.Rept. **513**, 1 (2012), 312 pages, 15 figures, 1106.2476.

[63] C. Brans and R. H. Dicke, Phys.Rev. **124**, 925 (1961).

[64] C. M. Will, Living Rev.Rel. **9**, 3 (2006), gr-qc/0510072.

[65] J. Khoury (2010), 1011.5909.

[66] J. Khoury and A. Weltman, Phys.Rev.Lett. **93**, 171104 (2004), astro-ph/0309300.

[67] J. Khoury and A. Weltman, Phys.Rev. **D69**, 044026 (2004), astro-ph/0309411.

[68] M. Pietroni, Phys.Rev. **D72**, 043535 (2005), astro-ph/0505615.

[69] K. A. Olive and M. Pospelov, Phys.Rev. **D77**, 043524 (2008), 0709.3825.

[70] K. Hinterbichler and J. Khoury, Phys.Rev.Lett. **104**, 231301 (2010), 1001.4525.

[71] A. Vainshtein, Phys.Lett. **B39**, 393 (1972).

[72] K. Nordtvedt, Phys.Rev. **169**, 1014 (1968).

[73] L. Hui, A. Nicolis, and C. Stubbs, Phys.Rev. **D80**, 104002 (2009), 0905.2966.

[74] N. Arkani-Hamed, H.-C. Cheng, M. A. Luty, and S. Mukohyama, JHEP **0405**, 074 (2004), hep-th/0312099.

[75] T. Jacobson, PoS **QG-PH**, 020 (2007), 0801.1547.

[76] V. Kostelecky and N. Russell, Rev.Mod.Phys. **83**, 11 (2011), 0801.0287.

[77] A. H. Guth, Phys.Rev. **D23**, 347 (1981).

[78] A. D. Linde, Phys.Lett. **B108**, 389 (1982).

[79] A. Albrecht and P. J. Steinhardt, Phys.Rev.Lett. **48**, 1220 (1982).

[80] S. Weinberg, Phys.Rev. **D77**, 123541 (2008), 0804.4291.

[81] J. M. Bardeen, Phys.Rev. **D22**, 1882 (1980).

[82] V. F. Mukhanov, H. Feldman, and R. H. Brandenberger, Phys.Rept. **215**, 203 (1992).

[83] N. Birrell and P. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, 1984).

[84] C. Armendariz-Picon, T. Damour, and V. F. Mukhanov, Phys.Lett. **B458**, 209 (1999), hep-th/9904075.

[85] J. M. Maldacena, JHEP **0305**, 013 (2003), astro-ph/0210603.

[86] D. H. Lyth, C. Ungarelli, and D. Wands, Phys.Rev. **D67**, 023503 (2003), astro-ph/0208055.

[87] D. H. Lyth, Phys.Rev.Lett. **78**, 1861 (1997), hep-ph/9606387.

[88] D. Baumann et al. (CMBPol Study Team), AIP Conf.Proc. **1141**, 10 (2009), 0811.3919.

[89] K. Freese, J. A. Frieman, and A. V. Olinto, Phys.Rev.Lett. **65**, 3233 (1990).

[90] F. C. Adams, J. Bond, K. Freese, J. A. Frieman, and A. V. Olinto, Phys.Rev. **D47**, 426 (1993), hep-ph/9207245.

[91] N. Arkani-Hamed, H.-C. Cheng, P. Creminelli, and L. Randall, JCAP **0307**, 003 (2003), hep-th/0302034.

[92] D. E. Kaplan and N. J. Weiner, JCAP **0402**, 005 (2004), hep-ph/0302014.

[93] R. Kallosh, A. D. Linde, D. A. Linde, and L. Susskind, Phys.Rev. **D52**, 912 (1995), hep-th/9502069.

[94] E. J. Copeland, A. R. Liddle, D. H. Lyth, E. D. Stewart, and D. Wands, Phys.Rev. **D49**, 6410 (1994), astro-ph/9401011.

[95] D. Baumann and D. Green, JHEP **1104**, 071 (2011), 1009.3032.

[96] R. H. Brandenberger (1999), hep-ph/9910410.

[97] J. Martin and R. H. Brandenberger, Phys.Rev. **D63**, 123501 (2001), hep-th/0005209.

[98] C. Cheung, P. Creminelli, A. Fitzpatrick, J. Kaplan, and L. Senatore, JHEP **0803**, 014 (2008), 0709.0293.

[99] L. Senatore and M. Zaldarriaga (2010), arXiv:1009.2093 [hep-th], `1009.2093`.

[100] T. Pavlopoulos, Phys.Rev. **159**, 1106 (1967).

[101] D. Colladay and V. A. Kostelecky, Phys.Rev. **D58**, 116002 (1998), `hep-ph/9809521`.

[102] S. R. Coleman and S. L. Glashow, Phys.Rev. **D59**, 116008 (1999), `hep-ph/9812418`.

[103] V. A. Kostelecky, Phys.Rev. **D69**, 105009 (2004), `hep-th/0312310`.

[104] S. Weinberg, Phys.Rev. **166**, 1568 (1968).

[105] S. R. Coleman, J. Wess, and B. Zumino, Phys.Rev. **177**, 2239 (1969).

[106] C. Isham, A. Salam, and J. Strathdee, Phys.Lett. **B31**, 300 (1970).

[107] C. Isham, A. Salam, and J. Strathdee, Annals Phys. **62**, 98 (1971).

[108] D. V. Volkov, Fiz.Elem.Chast.Atom.Yadra **4**, 3 (1973).

[109] A. Borisov and V. Ogievetsky, Theor.Math.Phys. **21**, 1179 (1975).

[110] V. I. Ogievetsky, in *X-th winter school of theoretical physics in Karpacz, Poland* (1974).

[111] V. A. Kostelecky and S. Samuel, Phys.Rev.Lett. **63**, 224 (1989).

[112] V. A. Kostelecky and S. Samuel, Phys.Rev. **D40**, 1886 (1989).

[113] M. D. Seifert, Phys.Rev. **D81**, 065010 (2010), `0909.3118`.

[114] T. Jacobson and D. Mattingly, Phys.Rev. **D64**, 024028 (2001), `gr-qc/0007031`.

[115] B. Gripaios, JHEP **0410**, 069 (2004), `hep-th/0408127`.

[116] M. Libanov and V. Rubakov, JHEP **0508**, 001 (2005), `hep-th/0505231`.

[117] M. L. Graesser, A. Jenkins, and M. B. Wise, Phys.Lett. **B613**, 5 (2005), hep-th/0501223.

[118] B. Altschul, Q. G. Bailey, and V. A. Kostelecky, Phys.Rev. **D81**, 065028 (2010), 0912.4852.

[119] A. Perez and D. Sudarsky, Int.J.Mod.Phys. **A26**, 1493 (2011), 0811.3181.

[120] A. Zee, *Quantum field theory in a nutshell* (Princeton University Press, 2010).

[121] S. Elitzur, Phys.Rev. **D12**, 3978 (1975).

[122] P. Horava, Phys.Rev. **D79**, 084008 (2009), 0901.3775.

[123] N. Arkani-Hamed, H.-C. Cheng, M. Luty, and J. Thaler, JHEP **0507**, 029 (2005), hep-ph/0407034.

[124] D. Blas, O. Pujolas, and S. Sibiryakov, JHEP **0910**, 029 (2009), 0906.3046.

[125] N. Arkani-Hamed, H. Georgi, and M. D. Schwartz, Annals Phys. **305**, 96 (2003), hep-th/0210184.

[126] S. Dubovsky, JHEP **0410**, 076 (2004), hep-th/0409124.

[127] V. Rubakov and P. Tinyakov, Phys.Usp. **51**, 759 (2008), 0802.4379.

[128] A. H. Chamseddine and V. Mukhanov, JHEP **1008**, 011 (2010), 1002.3877.

[129] S. M. Carroll, A. De Felice, V. Duvvuri, D. A. Easson, M. Trodden, et al., Phys.Rev. **D71**, 063513 (2005), astro-ph/0410031.

[130] L. Ford, Phys.Rev. **D40**, 967 (1989).

[131] M. Bento, O. Bertolami, P. Moniz, J. Mourao, and P. Sa, Class.Quant.Grav. **10**, 285 (1993), gr-qc/9302034.

[132] S. M. Carroll and E. A. Lim, Phys.Rev. **D70**, 123525 (2004), hep-th/0407149.

[133] J. Beltran Jimenez and A. L. Maroto, Phys.Rev. **D78**, 063005 (2008), 0801.1486.

[134] A. Golovnev, V. Mukhanov, and V. Vanchurin, JCAP **0806**, 009 (2008), 6 pages, 0802.2068.

[135] T. Koivisto and D. F. Mota, JCAP **0808**, 021 (2008), 0805.4229.

[136] B. Himmetoglu, C. R. Contaldi, and M. Peloso, Phys.Rev.Lett. **102**, 111301 (2009), 0809.2779.

[137] C. Armendariz-Picon and A. Diez-Tejedor, JCAP **0912**, 018 (2009), 0904.0809.

[138] C. Armendariz-Picon, A. Diez-Tejedor, and R. Penco, JHEP **1010**, 079 (2010), 1004.5596.

[139] J. Bjorken, Annals Phys. **24**, 174 (1963).

[140] P. Kraus and E. Tomboulis, Phys.Rev. **D66**, 045015 (2002), hep-th/0203221.

[141] F. Sannino, Phys.Rev. **D67**, 054006 (2003), hep-ph/0211367.

[142] M. Bando, T. Kugo, and K. Yamawaki, Phys.Rept. **164**, 217 (1988).

[143] T. Appelquist and J. Carazzone, Phys.Rev. **D11**, 2856 (1975).

[144] B. Withers, Class.Quant.Grav. **26**, 225009 (2009), 0905.2446.

[145] R. Utiyama, Phys.Rev. **101**, 1597 (1956).

[146] S. M. Carroll, *Spacetime and geometry: An introduction to general relativity* (Benjamin Cummings, 2004).

[147] R. Percacci, Nucl.Phys. **B353**, 271 (1991), 0712.3545.

[148] I. Kirsch, Phys.Rev. **D72**, 024001 (2005), hep-th/0503024.

[149] M. Leclerc, Annals Phys. **321**, 708 (2006), gr-qc/0502005.

[150] R. Bluhm and V. A. Kostelecky, Phys.Rev. **D71**, 065008 (2005), hep-th/0412320.

[151] Q. G. Bailey and V. A. Kostelecky, Phys.Rev. **D74**, 045001 (2006), gr-qc/0603030.

[152] R. Bluhm, S.-H. Fung, and V. A. Kostelecky, Phys.Rev. **D77**, 065020 (2008), 0712.4119.

[153] V. A. Kostelecky and R. Potting, Phys.Rev. **D79**, 065018 (2009), 0901.0662.

[154] H. P. Nilles, Phys.Lett. **B115**, 193 (1982).

[155] A. H. Chamseddine, R. L. Arnowitt, and P. Nath, Phys.Rev.Lett. **49**, 970 (1982).

[156] R. Barbieri, S. Ferrara, and C. A. Savoy, Phys.Lett. **B119**, 343 (1982).

[157] S. Schlamminger, K.-Y. Choi, T. Wagner, J. Gundlach, and E. Adelberger, Phys.Rev.Lett. **100**, 041101 (2008), 0712.0607.

[158] T. Damour and J. F. Donoghue, Class.Quant.Grav. **27**, 202001 (2010), 1007.2790.

[159] K. Nordtvedt, Phys.Rev. **169**, 1017 (1968).

[160] P. Jordan, Z.Phys. **157**, 112 (1959).

[161] M. Fierz, Helv.Phys.Acta **29**, 128 (1956).

[162] P. Jordan, *Schwerkraft und Weltall* (Friedrich Vieweg und Sohn, 1955).

[163] R. V. Wagoner, Phys.Rev. **D1**, 3209 (1970).

[164] Y. Fujii, Mod.Phys.Lett. **A9**, 3685 (1994), gr-qc/9411068.

[165] Y. Fujii, Mod.Phys.Lett. **A12**, 371 (1997), gr-qc/9610006.

[166] Y. Cho, Class.Quant.Grav. **14**, 2963 (1997).

[167] L. Hui and A. Nicolis, Phys.Rev.Lett. **105**, 231101 (2010), `1009.2520`.

[168] C. Armendariz-Picon and R. Penco, Phys.Rev. **D85**, 044052 (2012), `1108.6028`.

[169] S. Weinberg, Phys.Rev. **135**, B1049 (1964).

[170] R. Brout and F. Englert, Phys.Rev. **141**, 1231 (1966).

[171] B. S. DeWitt, Phys.Rev. **162**, 1239 (1967).

[172] R. Dicke, Phys.Rev. **125**, 2163 (1962).

[173] S. Deser and P. van Nieuwenhuizen, Phys.Rev. **D10**, 411 (1974).

[174] L. Abbott, M. T. Grisaru, and R. K. Schaefer, Nucl.Phys. **B229**, 372 (1983).

[175] S. R. Coleman and R. Jackiw, Annals Phys. **67**, 552 (1971).

[176] G. 't Hooft, Nucl.Phys. **B61**, 455 (1973).

[177] T. Damour, C.R.Acad.Sci. (2001), `gr-qc/0109063`.

[178] P. Brax, C. Burrage, A.-C. Davis, D. Seery, and A. Weltman, Phys.Lett. **B699**, 5 (2011), `1010.4536`.

[179] A. Nicolis, R. Rattazzi, and E. Trincherini, Phys.Rev. **D79**, 064036 (2009), `0811.2197`.

[180] S. R. Coleman, *Aspects of Symmetry* (Cambridge University Press, 1988).

[181] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena* (Oxford University Press, 2002), 4th ed.

[182] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover Publications, 1965).

[183] A. A. Starobinsky, Phys.Lett. **B91**, 99 (1980).

[184] S. Hollands and R. M. Wald, Gen.Rel.Grav. **36**, 2595 (2004), `gr-qc/0405082`.

[185] R. H. Brandenberger and J. Martin, Mod.Phys.Lett. **A16**, 999 (2001), astro-ph/0005432.

[186] J. C. Niemeyer, Phys.Rev. **D63**, 123502 (2001), astro-ph/0005533.

[187] A. Kempf, Phys.Rev. **D63**, 083514 (2001), astro-ph/0009209.

[188] A. A. Starobinsky, JETP Lett. **73**, 371 (2001), astro-ph/0104043.

[189] R. Easther, B. R. Greene, W. H. Kinney, and G. Shiu, Phys.Rev. **D64**, 103502 (2001), hep-th/0104102.

[190] R. Easther, B. R. Greene, W. H. Kinney, and G. Shiu, Phys.Rev. **D67**, 063508 (2003), hep-th/0110226.

[191] J. C. Niemeyer and R. Parentani, Phys.Rev. **D64**, 101301 (2001), astro-ph/0101451.

[192] A. Kempf and J. C. Niemeyer, Phys.Rev. **D64**, 103501 (2001), astro-ph/0103225.

[193] R. H. Brandenberger and J. Martin, Int.J.Mod.Phys. **A17**, 3663 (2002), hep-th/0202142.

[194] U. H. Danielsson, Phys.Rev. **D66**, 023511 (2002), hep-th/0203198.

[195] R. Easther, B. R. Greene, W. H. Kinney, and G. Shiu, Phys.Rev. **D66**, 023518 (2002), hep-th/0204129.

[196] L. Bergstrom and U. H. Danielsson, JHEP **0212**, 038 (2002), hep-th/0211006.

[197] N. Kaloper, M. Kleban, A. E. Lawrence, and S. Shenker, Phys.Rev. **D66**, 123510 (2002), hep-th/0201158.

[198] J. Martin and R. Brandenberger, Phys.Rev. **D68**, 063513 (2003), hep-th/0305161.

[199] O. Elgaroy and S. Hannestad, Phys.Rev. **D68**, 123513 (2003), astro-ph/0307011.

[200] J. Martin and C. Ringeval, Phys.Rev. **D69**, 083515 (2004), astro-ph/0310382.

[201] T. Okamoto and E. A. Lim, Phys.Rev. **D69**, 083519 (2004), astro-ph/0312284.

[202] S. Shankaranarayanan and L. Sriramkumar, Phys.Rev. **D70**, 123520 (2004), hep-th/0403236.

[203] J. Martin and C. Ringeval, JCAP **0501**, 007 (2005), hep-ph/0405249.

[204] K. Schalm, G. Shiu, and J. P. van der Schaar, JHEP **0404**, 076 (2004), hep-th/0401164.

[205] B. R. Greene, K. Schalm, G. Shiu, and J. P. van der Schaar, JCAP **0502**, 001 (2005), hep-th/0411217.

[206] K. Schalm, G. Shiu, and J. P. van der Schaar, AIP Conf.Proc. **743**, 362 (2005), hep-th/0412288.

[207] R. Easther, W. H. Kinney, and H. Peiris, JCAP **0505**, 009 (2005), astro-ph/0412613.

[208] B. Greene, K. Schalm, J. P. van der Schaar, and G. Shiu, eConf **C041213**, 0001 (2004), astro-ph/0503458.

[209] A. Kempf and L. Lorenz, Phys.Rev. **D74**, 103517 (2006), 19 pages, LaTeX, gr-qc/0609123.

[210] C. Armendariz-Picon and E. A. Lim, JCAP **0312**, 006 (2003), hep-th/0303103.

[211] R. Holman and A. J. Tolley, JCAP **0805**, 001 (2008), 0710.1302.

[212] B. Losic and W. Unruh, Phys.Rev.Lett. **101**, 111101 (2008), 0804.4296.

[213] C. Armendariz-Picon, M. Fontanini, R. Penco, and M. Trodden, Class.Quant.Grav. **26**, 185002 (2009), 0805.0114.

[214] S. Weinberg, Phys.Rev. **D72**, 043514 (2005), hep-th/0506236.

[215] J. F. Donoghue, Phys.Rev. **D50**, 3874 (1994), gr-qc/9405057.

[216] W. Hu (2004), astro-ph/0402060.

[217] C. Armendariz-Picon, JCAP **0702**, 031 (2007), astro-ph/0612288.

VITA

AUTHOR'S NAME: Riccardo Penco

PLACE OF BIRTH: Trieste, Italy

DATE OF BIRTH: November 3, 1981

DEGREES AWARDED:

Laurea Specialistica in Fisica, Università degli Studi di Trieste, April 2006, *cum laude*

Laurea in Fisica, Università degli Studi di Trieste, November 2003, *cum laude*

DISTINCTIONS:

Outstanding Teaching Assistant Award, Graduate School, Syracuse University, 2012

Summer Fellowship, Department of Physics, Syracuse University, 2008

CRUT Scholarship, Università degli Studi di Trieste, 2002–2006

Poropat Prize, Department of Physics, Università degli Studi di Trieste, 2004

Ananian Foundation Scholarship, 2002–2003

PROFESSIONAL EXPERIENCE:

Research/Teaching Assistant, Department of Physics, Syracuse University, 2006–2012