Master's Thesis
Theoretical Physics

Higgs inflation and higher-order gravity in Palatini formulation

Jaakko Annala

April 2020

Supervisor(s): Syksy Räsänen

Examiner(s): Syksy Räsänen
Mark Hindmarsh

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty		Koulutusohjelma — Utbildningsprogram — Degree programme	
Faculty of Science		Theoretical Physics	
Tekijä — Författare — Author			
Jaakko Annala			
Työn nimi — Arbetets titel — Title			
Higgs inflation and higher-order gravity in Palatini formulation			
Työn laji — Arbetets art — Level	Aika — Datum — Month and year		Sivumäärä — Sidantal — Number of pages
Master's Thesis	April 2020		69

Tiivistelmä — Referat — Abstract

We study how higher-order gravity affects Higgs inflation in the Palatini formulation. We first review the metric and Palatini formulations in comparative manner and discuss their differences. Next cosmic inflation driven by a scalar field and inflationary observables are discussed. After this we review the Higgs inflation and compute the inflationary observables both in the metric and Palatini formulations. We then consider adding higher-order terms of the curvature to the action. We derive the equations of motion for the most general action quadratic in the curvature that does not violate parity in both the metric and Palatini formulations. Finally we present a new result. We analyse Higgs inflation in the Palatini formulation with higher-order curvature terms. We consider a simplified scenario where only terms constructed from the symmetric part of the Ricci tensor are added to the action. This implies that there are no new gravitational degrees of freedom, which makes the analysis easier. As a new result we found out that the scalar perturbation spectrum is unchanged, but the tensor perturbation spectrum is suppressed by the higher-order curvature couplings.

Avainsanat — Nyckelord — Keywords

cosmology, inflation, Higgs, modifications of gravity, Palatini formulation

Säilytyspaikka — Förvaringsställe — Where deposited

Muita tietoja — Övriga uppgifter — Additional information

Contents

7   Conclusions   **54**

Appendices   **56**

A   Useful identities   **56**

B   Equations of motion in the Palatini formulation   **58**

C   Details of the series solution   **61**

Notation and conventions

We follow the conventions of [1]. The metric tensor with the signature $(-+++)$ is denoted by $g_{\mu\nu}$ and its determinant by

$$g = \det(g_{\mu\nu}). \tag{0.1}$$

Covariant derivative with respect to a general connection is denoted by ∇_μ and its connection coefficients by $\Gamma^\sigma_{\mu\nu}$. The connection coefficients are defined such that the covariant derivative is

$$\nabla_\mu T^\nu = \partial_\mu T^\nu + \Gamma^\nu_{\mu\lambda} T^\lambda, \tag{0.2a}$$

$$\nabla_\mu T_\nu = \partial_\mu T_\nu - \Gamma^\lambda_{\mu\nu} T_\lambda. \tag{0.2b}$$

When it is more convenient we use the comma notation for partial derivatives, $\frac{\partial \phi}{\partial x^\mu} = \partial_\mu \phi = \phi_{,\mu}$, and a semicolon for covariant derivatives, $\nabla_\mu T^{\cdots}_{\cdots} = T^{\cdots}_{\cdots;\mu}$. The Levi–Civita connection, which is metric compatible and symmetric, is denoted by a circle over the symbol and the coefficients are

$$\mathring{\Gamma}^\sigma_{\mu\nu} = \frac{1}{2} g^{\sigma\rho} \left(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu} \right). \tag{0.3}$$

Furthermore, quantities constructed from the Levi–Civita connection are denoted by a circle over the symbols, e.g. $\mathring{\nabla}_\mu$. The Riemann tensor is defined as

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda} \Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda} \Gamma^\lambda_{\mu\sigma}. \tag{0.4}$$

The symmetrization of indices is defined as

$$T^{(\mu_1 \ldots \mu_n)} = \frac{1}{n!} \sum_{\sigma \in S_n} T^{\mu_{\sigma(1)} \ldots \mu_{\sigma(n)}}, \tag{0.5}$$

where S_n is the symmetric group of n elements. The anti-symmetrization of indices is defined as

$$T^{[\mu_1 \ldots \mu_n]} = \frac{1}{n!} \sum_{\sigma \in S_n} \text{sgn}(\sigma) T^{\mu_{\sigma(1)} \ldots \mu_{\sigma(n)}}, \tag{0.6}$$

where $\text{sgn}(\sigma)$ gives the sign of the permutation σ: $+1$ for even and -1 for odd permutation. The torsion tensor is defined as

$$T^\sigma{}_{\mu\nu} = 2\Gamma^\sigma_{[\mu\nu]}. \tag{0.7}$$

The Weyl tensor is defined in dimension $d > 2$ as

$$\begin{aligned} W_{\rho\sigma\mu\nu} = R_{\rho\sigma\mu\nu} &- \frac{2}{(d-2)} \left(g_{\rho[\mu} R_{\nu]\sigma} - g_{\sigma[\mu} R_{\nu]\rho} \right) \\ &+ \frac{2}{(d-1)(d-2)} g_{\rho[\mu} g_{\nu]\sigma} R. \end{aligned} \tag{0.8}$$

The Levi–Civita symbol in right handed coordinate systems is defined

$$\epsilon_{\mu_1\mu_2...\mu_n} = \begin{cases} +1 & \text{if} \quad \mu_1...\mu_n \quad \text{is an even permutation of} \quad 0,1,...,(n-1) \\ -1 & \text{if} \quad \mu_1...\mu_n \quad \text{is an odd permutation of} \quad 0,1,...,(n-1) \\ 0 & \text{otherwise.} \end{cases} \tag{0.9}$$

The generalized Kronecker delta is defined

$$\delta^{\mu_1\mu_2...\mu_n}_{\nu_1\nu_2...\nu_n} \equiv n! \delta^{[\mu_1}_{\nu_1} \delta^{\mu_2}_{\nu_2} ... \delta^{\mu_n]}_{\nu_n}. \tag{0.10}$$

We use units where

$$\hbar = c = 1. \tag{0.11}$$

Additionally, unless otherwise stated, we set the Planck mass to unity

$$M_p \equiv \frac{1}{\sqrt{8\pi G_N}} = 1, \tag{0.12}$$

where G_N is the Newton's gravitational constant.

Chapter 1

Introduction

Inflation is a hypothetical period of exponential expansion in the very early universe, which is able to explain problems that the standard Hot Big Bang model cannot answer. Most notably, it can explain the origin of the primordial perturbations that ultimately give rise to the structure formation of our universe and predict the form of the fluctuations we observe in the cosmic microwave background (CMB). Inflation is a paradigm, and there is a plethora of different models that have various ways of producing the exponential expansion of the universe, see e.g. [2]. In Higgs inflation the Standard Model (SM) Higgs is non-minimally coupled to gravity and drives inflation. The advantage of this scenario is that it is very minimalistic: no new fields are added to the SM and only one new parameter is added, namely the non-minimal coupling to gravity. Higgs inflation is reviewed in chapter 4.

In higher-order gravity, higher-order terms in the curvature are present in the gravity sector. The appearance of these terms can be motivated in several ways, more on this in chapter 5. Modified gravity sector (even just the addition of non-minimal coupling in Higgs inflation) makes the question: which are the fundamental degrees of freedom in General Relativity (GR), even more relevant. GR can be formulated in different ways by starting with different independent variables and applying the variational principle to the Einstein–Hilbert action. In the metric formulation of GR the spacetime manifold is described by a metric and the Levi–Civita connection, which is completely described by the metric. The metric is assumed to be the only degree of freedom of gravity and varying the Einstein–Hilbert action with respect to the metric gives the Einstein field equations. However, a priori there is no reason for the connection to be described purely by the metric. In the Palatini formulation[1] the metric and a general connection are treated as independent degrees of freedom. In the case of the Einstein–Hilbert action, variation with respect to the connection makes the connection become the Levi–Civita connection, and thus in the case of the Einstein–Hilbert action these formulations are equivalent. However, when the gravity sector is modified the two formulations differ. These two formulations lead to different inflationary predictions. The aim of this thesis is to investigate how higher-order curvature terms in the Palatini formulation change the predictions of Higgs inflation.

The structure of this thesis is as follows. First, in chapter 2 we introduce the metric and Palatini formulations of gravity. We show that the formulations are equivalent with the Einstein–Hilbert action, however with subtle differences. When introducing the metric formulation we discuss non-trivial boundary terms. In Pala-

---

[1]In fact Einstein was the first to consider the Palatini formulation [3].

tini formulation we consider the subtleties arising with the most general connection.

In chapter 3, we introduce the topics of cosmology relevant for this thesis, namely the FLRW model and inflation. In the case of inflation we focus on the generation of primordial perturbations and inflationary observables.

In chapter 4, we review Higgs inflation in the metric and Palatini formulations. We also discuss conformal transformations that make the analysis easier.

In chapter 5 we review the basic properties of higher-order gravity in the metric and Palatini formulations. By considering $f(R)$ gravity we illustrate the differences between the two formulations. We also present the general parity preserving action that is quadratic in the curvature tensor in the metric and Palatini formulation and derive the equations of motion for these terms.

Finally, in chapter 6 we present a new result. We tackle the question how the higher-order curvature terms change Higgs inflation. We consider a simple case where only the symmetric part of the Ricci tensor appears in the action. We also discuss disformal transformations, which turn out to be relevant for our analysis. We then derive the slow-roll equations and finally compute how the inflationary predictions are modified by the added higher-order curvature terms.

Chapter 2

Variational principle in general relativity

The variational principle has an important role in physics. Most field theories are formulated by writing down an action in terms of the fields of the theory. GR is no exception and it can be formulated using an action and applying the variational principle.

The formulation of gravity that we are focusing in this thesis is the Palatini formulation (sometimes also called the first order formalism) in which the metric tensor and the connection are independent degrees of freedom. In this chapter we review the main features of the metric and the Palatini formulations and see how the procedure works with the Einstein–Hilbert action. We also consider the subtle differences of these formulations.

2.1 Metric formulation

The assumption in metric formulation is that the only degree of freedom describing the spacetime is the metric $g_{\mu\nu}$ and hence the connection on the manifold is metric compatible $\mathring{\nabla}_\sigma g^{\mu\nu} = 0$ (parallel transport preserves the metric) and torsion free $\mathring{\Gamma}^\sigma_{\mu\nu} = \mathring{\Gamma}^\sigma_{\nu\mu}$. Thus, the connection coefficients are

$$\mathring{\Gamma}^\sigma_{\mu\nu} = \frac{1}{2} g^{\sigma\rho} \left(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu} \right). \tag{2.1}$$

The formulation proceeds by defining a Lagrangian density $\mathcal{L}$, which is a functional depending purely on the metric and its derivatives: $\mathcal{L} = \mathcal{L}\left(g_{\mu\nu}, \partial g_{\mu\nu}, \partial\partial g_{\mu\nu}, ...\right)$. The Lagrangian has to be Lorentz invariant, also other mathematical or physical constraints may be imposed e.g. simplicity and invariance under other transformations. The action integral is the Lagrangian integrated over a compact region of the spacetime Σ with respect to the invariant spacetime volume element $\sqrt{-g}d^n x$. The action reads

$$S[g_{\mu\nu}] = \int_\Sigma d^n x \sqrt{-g}\mathcal{L}. \tag{2.2}$$

Then the classical variational principle is applied: requiring that an arbitrary variation of the action with respect to the metric should vanish identically, while the variation on the boundary $\partial\Sigma$ vanishes. In other words, the functional derivative of the Lagrangian $\mathcal{L}$ with respect to the metric vanishes,

$$\frac{\delta\mathcal{L}}{\delta g^{\mu\nu}} = 0. \tag{2.3}$$

The simplest Lagrangian constructed only from the metric and its derivatives, that leads to the Einstein field equations, is the Ricci scalar. This gives the Einstein–Hilbert action

$$S_{EH} = \frac{1}{2} \int d^4x \sqrt{-g} \mathring{R}, \tag{2.4}$$

where $\mathring{R} = g^{\mu\nu} \mathring{R}^{\lambda}{}_{\mu\lambda\nu}$ is the Ricci scalar. Variation of this action produces Einstein field equations in vacuum. One may add in to the action a cosmological constant Λ and a matter part $\mathcal{L}_M$. The variation of the matter part will yield the stress-energy tensor

$$T_{\mu\nu} = \frac{-2}{\sqrt{-g}} \frac{\delta S_M}{\delta g^{\mu\nu}}, \tag{2.5}$$

where S_M is the matter part of the action containing $\mathcal{L}_M$. This leads to the Einstein field equations with a cosmological constant and a source term

$$\mathring{R}_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \mathring{R} + \Lambda g_{\mu\nu} = T_{\mu\nu}. \tag{2.6}$$

There is, however, a slight caveat in the derivation that concerns the boundary terms. When comparing different formulations it is useful to take the boundaries into account, since the resulting boundary conditions might be different.

2.1.1 Einstein–Hilbert action in metric formulation

We will shortly review the process of variation with the Einstein–Hilbert action. We need to take the variation of $\sqrt{-g} g^{\mu\nu} \mathring{R}_{\mu\nu}$ with respect to the inverse metric $g^{\mu\nu}$. To accomplish this we make use of a couple of useful identities, namely:

$$\delta \sqrt{-g} = -\frac{1}{2} \sqrt{-g} g_{\mu\nu} \delta g^{\mu\nu} \tag{2.7}$$

and the Palatini identity, which in this case reads

$$\delta \mathring{R}^{\rho}{}_{\mu\lambda\nu} = \mathring{\nabla}_\lambda \delta \mathring{\Gamma}^{\rho}_{\nu\mu} - \mathring{\nabla}_\nu \delta \mathring{\Gamma}^{\rho}_{\lambda\mu}, \tag{2.8}$$

see appendix A. First of all we have:

$$\delta(\sqrt{-g} g^{\mu\nu} \mathring{R}_{\mu\nu}) = \sqrt{-g} \mathring{R}_{\mu\nu} \delta g^{\mu\nu} + \mathring{R} \delta \sqrt{-g} + \sqrt{-g} g^{\mu\nu} \delta \mathring{R}_{\mu\nu}. \tag{2.9}$$

Taking the variation of the connection (2.1) gives

$$\delta \mathring{\Gamma}^{\sigma}{}_{\mu\nu} = \frac{1}{2} \left[g_{\mu\alpha} g_{\nu\beta} \mathring{\nabla}^\sigma \delta g^{\alpha\beta} - g_{\alpha\mu} \mathring{\nabla}_\nu \delta g^{\alpha\sigma} - g_{\alpha\nu} \mathring{\nabla}_\mu \delta g^{\alpha\sigma} \right]. \tag{2.10}$$

Plugging (2.10) into the Palatini identity (2.8) and taking the contraction between the first and the third index we end up with the variation of the Ricci tensor, finally contracting this we have

$$g^{\mu\nu} \delta \mathring{R}_{\mu\nu} = \mathring{\nabla}_\lambda \left[g_{\mu\nu} \mathring{\nabla}^\lambda \delta g^{\mu\nu} - \mathring{\nabla}_\alpha \delta g^{\alpha\lambda} \right]. \tag{2.11}$$

The variation of the action then reads

$$
\begin{aligned}
\delta S_{EH} = & \int d^4x\sqrt{-g}\left[\mathring{R}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathring{R}\right]\delta g^{\mu\nu} \\
& + \int d^4x\sqrt{-g}\,\mathring{\nabla}_\lambda\left[g_{\mu\nu}\mathring{\nabla}^\lambda\delta g^{\mu\nu} - \mathring{\nabla}_\alpha\delta g^{\alpha\lambda}\right].
\end{aligned}
\tag{2.12}
$$

From this we immediately see that the first term leads to the Einstein field equations in vacuum and the second term leads to a boundary term which we will consider next.

2.1.2 York-Gibbons-Hawking boundary term

The second term in (2.12) is a total divergence and leads to a boundary term by using the Stokes' theorem

$$
\int_\Sigma d^n x\sqrt{-g}\,\mathring{\nabla}_\mu V^\mu = \int_{\partial\Sigma} d^{n-1}x\sqrt{|\gamma|}n_\mu V^\mu,
\tag{2.13}
$$

where V^μ is a vector field defined in the region Σ with a boundary $\partial\Sigma$, γ_{ab} is the induced metric on the boundary and n^μ is the unit normal to the surface. Note that this form of the Stokes' theorem only applies with the Levi–Civita connection; this turns out to be important later when we consider the Palatini formulation. Using this, the second term in (2.12) equals

$$
\int_{\partial\Sigma} d^3x\sqrt{|\gamma|}n_\lambda\left[g_{\mu\nu}\mathring{\nabla}^\lambda\delta g^{\mu\nu} - \mathring{\nabla}_\alpha\delta g^{\alpha\lambda}\right],
\tag{2.14}
$$

from which we see that the boundary term does not vanish if we only assume that the variation vanishes on the boundary. By imposing $\delta g^{\mu\nu}|_{\partial\Sigma} = 0$ we get

$$
\int_{\partial\Sigma} d^3x\sqrt{|\gamma|}n_\lambda\left[g_{\mu\nu}\partial^\lambda\delta g^{\mu\nu} - \partial_\alpha\delta g^{\alpha\lambda}\right].
\tag{2.15}
$$

We would also have to require $\partial_\lambda\delta g^{\mu\nu}|_{\partial\Sigma} = 0$ for this term to vanish. An other way is to modify the action by a term that will cancel this boundary contribution. This additional term is called the York-Gibbons-Hawking (YGH) boundary term [4, 5]. The expression (2.15) can be further simplified by using the following identity for the induced metric, assuming that the surface $\partial\Sigma$ is not null (see e.g. [6, ch. 2.7])

$$
g^{\mu\nu} = \epsilon n^\mu n^\nu + \gamma^{ab}\frac{\partial x^\mu}{\partial y^a}\frac{\partial x^\nu}{\partial y^b},
\tag{2.16}
$$

where $\epsilon = n^\mu n_\mu$ equals $+1$ if $\partial\Sigma$ is time-like and -1 if spacelike and y^a are the coordinates on the surface. We also denote

$$
\gamma^{\mu\nu} = \gamma^{ab}\frac{\partial x^\mu}{\partial y^a}\frac{\partial x^\nu}{\partial y^b}.
\tag{2.17}
$$

The requirement that the variation of the metric vanishes on the boundary implies that the tangential derivative $n_\lambda \gamma_{\mu\nu} \partial^\nu \delta g^{\mu\lambda} = 0$ vanishes. The surface term (2.15) then becomes

$$\int_{\partial\Sigma} d^3x \sqrt{|\gamma|} \gamma_{\mu\nu} n_\lambda \partial^\lambda \delta g^{\mu\nu},\tag{2.18}$$

which can be be canceled by introducing a YGH–boundary term

$$S_{YGH} = \int_{\partial\Sigma} d^3x \epsilon \sqrt{|\gamma|} K,\tag{2.19}$$

where $K = \mathring{\nabla}_\mu n^\mu$ is the trace of the extrinsic curvature of the boundary. It is straightforward to verify this by taking the variation of this boundary term (2.19) and see that it yields (2.18) up to a constant. For more details see e.g. [7] and [8, Appendix E]. Thus when taking into account the boundaries, the action that produces the Einstein field equations in the metric formulation is

$$S_{GR} = \int_\Sigma d^4x \sqrt{-g}\mathring{R} + 2\int_{\partial\Sigma} d^3x \epsilon \sqrt{|\gamma|} K.\tag{2.20}$$

This is one method of dealing with the boundary term. In this thesis we are not concerned about the details of these terms; for the physical significance of these terms and further consideration see e.g. [9]. In the next section we will see that in the Palatini formulation no YGH boundary term need to be added.

2.2 Palatini formulation

In the Palatini formulation the metric and the connection are considered as independent degrees of freedom. Thus we have a spacetime manifold with a symmetric Lorentzian metric $g_{\mu\nu}$ and a general connection, for which the metric compatibility does not hold, $\nabla_\sigma g_{\mu\nu} \neq 0$, and the torsion does not necessarily vanish $T^\sigma{}_{\mu\nu} \neq 0$.

The formulation is proceeded by constructing a Lagrangian density, which now is a functional of both the metric and the connection and their derivatives $\mathcal{L} = \mathcal{L}(g, \partial g, \partial\partial g, ..., \Gamma, \partial\Gamma, \partial\partial\Gamma, ...)$. The action is now,

$$S[g_{\mu\nu}, \Gamma^\sigma{}_{\mu\nu}] = \int_\Sigma d^n x \sqrt{-g}\mathcal{L}.\tag{2.21}$$

Next we apply the classical variational principle: first we vary both the metric and the connection separately with the variations vanishing on the boundary $\partial\Sigma$, and then require that the variation of the action vanishes. This yields the equations of motion for both the metric and the connection

$$\frac{\delta\mathcal{L}}{\delta g^{\mu\nu}} = 0, \quad \text{and} \quad \frac{\delta\mathcal{L}}{\delta\Gamma^\sigma{}_{\mu\nu}} = 0.\tag{2.22}$$

The appeal of the Palatini formulation is that in the case of Einstein–Hilbert action the dynamical equation for the connection imposes the connection to be

Levi–Civita and thus we recover the pseudo-Riemannian space with Einstein equations, with no prior assumptions. Furthermore there is no need to introduce YGH–boundary terms to the action. There are, however, some unappealing properties in this formulation when considering the most general connection and when introducing matter into the Lagrangian. We will discuss these problems after we review the variation of the Einstein–Hilbert action and compare the procedure to the metric case.

2.2.1 Einstein–Hilbert action in the Palatini formulation

We consider the case of a completely general connection. We denote the Riemann tensor constructed purely from the connection (and tensors derived from it) with $R^{\rho}{}_{\mu\sigma\nu}$ to differentiate from the Levi–Civita counterpart denoted with $\mathring{R}^{\rho}{}_{\mu\sigma\nu}$. With a completely general connection the Riemann tensor does not have all the symmetry properties as with the Levi–Civita connection. The only property that remains is the antisymmetry of the last two indices, which is apparent from the definition (0.4). This then implies that there is more than one unique contraction that we can take of the Riemann tensor in contrast to the one unique choice in the Levi–Civita case: the Ricci tensor. Thus it is interesting to see whether the Ricci scalar is unique or not. There are in fact three different non vanishing contractions of the Riemann tensor. The Ricci tensor is defined as

$$R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu}, \tag{2.23}$$

which in general is not symmetric. We define the co-Ricci tensor as

$$\hat{R}^{\mu}{}_{\nu} = g^{\lambda\sigma} R^{\mu}{}_{\sigma\nu\lambda}, \tag{2.24}$$

which in general does not have any symmetry properties. Finally we define the antisymmetric Ricci tensor

$$R'_{\mu\nu} = R^{\lambda}{}_{\lambda\mu\nu}, \tag{2.25}$$

which is sometimes also called the homothetic curvature and it is antisymmetric from the definition of the Riemann tensor. Since the metric tensor is symmetric the trace of $g^{\mu\nu} R'_{\mu\nu} = 0$ vanishes, furthermore

$$\hat{R}^{\mu}{}_{\mu} = g^{\lambda\sigma} R^{\mu}{}_{\sigma\mu\lambda} = g^{\lambda\sigma} R_{\lambda\sigma} = R. \tag{2.26}$$

Thus the Ricci scalar still remains unique. These different Ricci tensors come into play in the higher-order Gravity in section 5, but now we can say that the Einstein–Hilbert action is still uniquely determined[1].

The action is

$$S = \int d^4x \sqrt{-g}\, g^{\mu\nu} R_{\mu\nu}(\Gamma). \tag{2.27}$$

---

[1]There exist other candidates other than the Ricci scalar that results in second order field equations, see e.g. [10, 11].

The variation with respect to the metric is rather straightforward, since the variation of the Ricci tensor is zero; so, we have

$$\delta(\sqrt{-g}g^{\mu\nu}R_{\mu\nu}(\Gamma)) = \delta\sqrt{-g}g^{\mu\nu}R_{\mu\nu}(\Gamma) + \delta(g^{\mu\nu})R_{\mu\nu}(\Gamma). \tag{2.28}$$

Using (2.7) the metric variation $\frac{\delta S}{\delta g^{\mu\nu}} = 0$ gives

$$R_{(\mu\nu)} - \frac{1}{2}g_{\mu\nu}R = 0. \tag{2.29}$$

Notice that there were no appearances of any boundary terms.

The variation with respect to the connection is more involved. Since the only object depending on the connection is the Ricci tensor we need to compute its variation $\delta R_{\mu\nu}$. For this we make use of the general Palatini identity

$$\delta R^{\rho}{}_{\mu\lambda\nu} = \nabla_{\lambda}\delta\Gamma^{\rho}_{\nu\mu} - \nabla_{\nu}\delta\Gamma^{\rho}_{\lambda\mu} + T^{\sigma}{}_{\lambda\nu}\,\delta\Gamma^{\rho}_{\sigma\mu}, \tag{2.30}$$

see Appendix A. Taking the contraction of the first and third index gives the variation of $\delta R_{\mu\nu}$. So the variation of the action so far is

$$\delta S = \int d^4x\sqrt{-g}g^{\mu\nu}\left(\nabla_{\lambda}\delta\Gamma^{\lambda}_{\nu\mu} - \nabla_{\nu}\delta\Gamma^{\lambda}_{\lambda\mu} + T^{\sigma}{}_{\lambda\nu}\,\delta\Gamma^{\lambda}_{\sigma\mu}\right). \tag{2.31}$$

The first and the second term inside the brackets need to be simplified further. We will focus on the first term in (2.31), since steps for the second term are the same. With the chain rule we have

$$\sqrt{-g}g^{\mu\nu}\nabla_{\lambda}\delta\Gamma^{\lambda}_{\nu\mu} = \nabla_{\lambda}\left(\sqrt{-g}g^{\mu\nu}\delta\Gamma^{\lambda}_{\nu\mu}\right) - \nabla_{\lambda}\left(\sqrt{-g}g^{\mu\nu}\right)\delta\Gamma^{\lambda}_{\nu\mu}. \tag{2.32}$$

The first term on the right side of the equation resembles a surface term. We cannot, however, turn this into a surface term with the Stokes' theorem (2.13), since it is only valid for the Levi–Civita connection. To deal with this, we use a change of variables and write our general connection as

$$\Gamma^{\sigma}_{\mu\nu} = \mathring{\Gamma}^{\sigma}_{\mu\nu} + C^{\sigma}{}_{\mu\nu}, \tag{2.33}$$

where $\mathring{\Gamma}^{\sigma}_{\mu\nu}$ is the Levi–Civita connection and $C^{\sigma}{}_{\mu\nu}$ is the difference from it and hence a tensor. In addition, the second term on the right hand side of equation (2.32) is a covariant derivative acting on a tensor density of weight $+1$, which we will evaluate in order to get the desired form for the equation. For this we need the covariant derivative of a scalar density, which for a general scalar density weight ω is

$$\nabla_{\mu}\rho = \partial_{\mu}\rho - \omega\Gamma^{\sigma}_{\mu\sigma}\rho \tag{2.34}$$

see Appendix A or [12, ch. III]. Note that it is important which of the contractions is taken from the connection coefficients in (2.34) and this depends on the convention for the covariant derivative (0.2). Expanding (2.32) further

$$\sqrt{-g}g^{\mu\nu}\nabla_{\lambda}\delta\Gamma^{\lambda}_{\nu\mu} = \sqrt{-g}\nabla_{\lambda}(g^{\mu\nu}\delta\Gamma^{\lambda}_{\mu\nu}) + g^{\mu\nu}(\nabla_{\lambda}\sqrt{-g})\delta\Gamma^{\lambda}_{\mu\nu} - \nabla_{\lambda}\left(\sqrt{-g}g^{\mu\nu}\right)\delta\Gamma^{\lambda}_{\nu\mu}. \tag{2.35}$$

Now using the change of variables (2.33) and writing out the derivative in the second term we have

$$
\begin{aligned}
\sqrt{-g}g^{\mu\nu}\nabla_\lambda\delta\Gamma^\lambda_{\nu\mu} =& \sqrt{-g}\mathring{\nabla}_\lambda(g^{\mu\nu}\delta\Gamma^\lambda_{\mu\nu}) + \sqrt{-g}C^\lambda{}_{\lambda\sigma}g^{\mu\nu}\delta\Gamma^\sigma_{\mu\nu} \\
& + g^{\mu\nu}\left[\mathring{\nabla}_\lambda\sqrt{-g} - C^\sigma{}_{\lambda\sigma}\right]\delta\Gamma^\lambda_{\mu\nu} - \nabla_\lambda\left(\sqrt{-g}g^{\mu\nu}\right)\delta\Gamma^\lambda_{\nu\mu}.
\end{aligned}
\tag{2.36}
$$

The first term in this expression can now be turned into a surface term using Stokes' theorem (2.13) and it will vanish. The terms containing the $C^\sigma{}_{\mu\nu}$ tensors can be combined and written using the torsion as $C^\lambda{}_{\lambda\sigma} - C^\lambda{}_{\sigma\lambda} = T^\lambda{}_{\lambda\sigma}$ and finally we have $\mathring{\nabla}_\lambda\sqrt{-g} = 0$. With these (2.36) simplifies to

$$
\sqrt{-g}g^{\mu\nu}\nabla_\lambda\delta\Gamma^\lambda_{\nu\mu} = \left[\sqrt{-g}g^{\mu\nu}T^\lambda{}_{\lambda\sigma} - \nabla_\sigma(\sqrt{-g}g^{\mu\nu})\right]\delta\Gamma^\sigma_{\mu\nu},
\tag{2.37}
$$

up to the vanishing surface terms. The second term in (2.31) goes similarly giving finally the equations of motion for the connection

$$
\begin{aligned}
&\nabla_\lambda(\sqrt{-g}g^{\nu\lambda})\delta^\mu_\sigma - \nabla_\sigma(\sqrt{-g}g^{\mu\nu}) + \\
&\sqrt{-g}\left[g^{\mu\nu}T^\lambda{}_{\lambda\sigma} - g^{\nu\lambda}T^\rho{}_{\rho\lambda}\delta^\mu_\sigma + g^{\nu\lambda}T^\mu{}_{\sigma\lambda}\right] = 0.
\end{aligned}
\tag{2.38}
$$

At this point we can consider the case in which a symmetric connection is assumed. The obtained equation (2.38) simplifies considerably since the torsion vanishes and only the symmetric part of the upper indices contributes, we thus have

$$
\nabla_\lambda(\sqrt{-g}g^{\lambda(\nu})\delta^{\mu)}_\sigma - \nabla_\sigma(\sqrt{-g}g^{\mu\nu}) = 0.
\tag{2.39}
$$

Taking the trace of this with indices μ and σ we end up with

$$
\nabla_\sigma(\sqrt{-g}g^{\mu\nu}) = 0,
\tag{2.40}
$$

which is equivalent to the metric compatibility condition. Thus if we initially assume a symmetric connection the connection dynamically becomes the Levi–Civita connection. Then the equations of motion for the metric reduces to the Einstein field equations, without needing to introduce additional YGH-boundary terms to deal with the boundaries.

There are several good reasons to impose restrictions to the general connection. For example, the equivalence principle in GR states that the connection coefficients can be locally set to zero; however this is only possible with a completely symmetric connection. Moreover, having a space with torsion means that infinitesimal parallelograms do not close. Additionally, autoparallel curves and the extremal curves of the metric[2] are equivalent if and only if the torsion tensor is totally antisymmetric $T_{[\sigma\mu\nu]} = T_{\sigma\mu\nu}$ and the connection metric compatible [13]. Let us now inspect the most general case.

---

[2]Autoparallel curves are curves where vectors are parallel transported with respect to themselves, i.e. straight lines described by the connection. Extremal curves of the metric are curves that tell the shortest or longest path between two points.

In the general case it can be shown that the most general solution to (2.38) is given by [14]

$$\Gamma^{\sigma}_{\mu\nu} = \mathring{\Gamma}^{\sigma}_{\mu\nu} + \delta^{\sigma}_{\nu}V_{\mu}, \tag{2.41}$$

where V_{μ} is an arbitrary vector field. This says that the general solution for the connection differs from the Levi–Civita connection by a projective transformation, which acts as

$$\Gamma^{\sigma}_{\mu\nu} \to \Gamma^{\sigma}_{\mu\nu} + \delta^{\sigma}_{\nu}V_{\mu}. \tag{2.42}$$

It seems that to get the standard Einstein field equations one has to make an additional requirement of setting this arbitrary vector field to zero which is equivalent to saying that the trace of the torsion vanishes $T^{\lambda}_{\mu\lambda} = 0$.[3] There are different types of methods to impose this requirement e.g. where the constrain is introduced in terms of Lagrange multipliers [15]. However, it is easy to see that the Einstein–Hilbert action is in fact invariant under the projective transformation in the case of a symmetric metric. This can be seen by applying the transformation for the Ricci tensor which straightforwardly yields

$$R_{\mu\nu} \to R_{\mu\nu} + 2\partial_{[\mu}V_{\nu]}. \tag{2.43}$$

The projective transformation has the effect of adding an antisymmetric part to the Ricci tensor which vanishes when we take the contraction with symmetric metric, thus the action is invariant under projective transformations. This implies that at least the equations of motion are identical with the general solution (2.41) and the Levi–Civita connection. However, geodesics described by the general solution could still be different from Levi–Civita geodesics. It can be shown that the geodesics of the general solution are pre-geodesics [4] of the Levi–Civita ones, and so the geodesics coincide due to reparametrization invariance. The geodesics of the affine connection are solutions of the equation

$$\dot{x}^{\lambda}\nabla_{\lambda}\dot{x}^{\mu} = 0, \tag{2.44}$$

where $x^{\mu}(\tau)$ is a curve and the dot denotes a derivative with respect to the curve parameter τ. Using the solution (2.41) this can be written as

$$\dot{x}^{\lambda}\mathring{\nabla}_{\lambda}\dot{x}^{\mu} + V_{\lambda}\dot{x}^{\lambda}\dot{x}^{\mu} = 0. \tag{2.45}$$

The pre-geodesics are given by extremizing the arc-length functional

$$s(\lambda) = \int_{0}^{\lambda} \sqrt{g_{\mu\nu}\dot{x}^{\mu}\dot{x}^{\nu}}d\lambda, \tag{2.46}$$

which has the extremum

$$\dot{x}^{\lambda}\nabla_{\lambda}\dot{x}^{\mu} = (\ddot{s}/\dot{s})\dot{x}^{\mu}. \tag{2.47}$$

---

[3]This is obvious if one substitutes the solution (2.41) to the definition of torsion.

[4]Pre-geodesic of a connection is a smooth curve on a manifold, which has a reparametrization that is a geodesic of the connection; i.e. the curves describe the same trajectory with different parametrizations.

Now we can see that with the reparametrization

$$\frac{dx^\mu(\lambda)}{d\lambda} = \frac{dx^\mu(\tau)}{d\tau}\frac{d\tau}{d\lambda} \quad \text{with} \quad \frac{d\tau}{d\lambda} = \exp\left(-\int_0^\lambda V_\mu \frac{dx^\mu}{d\lambda} d\lambda\right) \qquad (2.48)$$

the geodesics of the general connection and the Levi–Civita one coincide. In these two cases, however, parallel transport of a vector will differ by a homothetic transformation (the resulting vectors differ by a scaling factor which depends on the parametrization). It has been argued that the difference between these transports is unobservable and the arbitrary vector field in the general solution is thus also unobservable [16]. These considerations rely on the assumption that the matter part of the Lagrangian is independent of the connection. This is not necessarily the case if one considers fermionic matter. This brings us to consider matter parts of the Lagrangian in the Palatini formulation.

2.2.2 Matter in the Palatini formulation

Under the assumption that the matter Lagrangian depends only on the metric, the matter fields and their derivatives $\mathcal{L}_M = \mathcal{L}_M(g, \Psi_i)$ the equations of motion for the connection do not change and we recover the Einstein field equations with a source term similarly to (2.6). This is straightforward if the matter Lagrangian is naturally independent of the connection, most notably in the case of a canonical scalar field. For a scalar field ϕ we have $\nabla_\mu\phi = \partial_\mu\phi$, and thus it is naturally not coupled to the connection. In general the matter Lagrangian depends on the metric and the covariant derivative. Thus assuming that the matter Lagrangian is independent of the general connection amounts to using the metric compatible covariant derivative in the matter part of the Lagrangian. This leads to the general connection being an auxillary field not related to the geometrical structure of the spacetime, since the geodesics are now defined through the Levi–Civita connection a priori. Under this assumption the theory ends up being equivalent to a metric theory of spacetime [17].

In the literature the Palatini formulation is often defined by assuming that the matter part is independent of the connection and the general case where this is allowed $\mathcal{L}_M = \mathcal{L}_M(g, \Gamma, \Psi_i)$ is called the metric-affine formulation. In the case where the matter is coupled to the connection, which comes naturally when considering fermionic fields, the situation is more complicated. Since now the equations of motion for the connection will also have a source term usually in the literature called the hypermomentum

$$\Delta_\sigma{}^{\mu\nu} = \frac{-2}{\sqrt{-g}}\frac{\delta S_M}{\delta\Gamma^\sigma_{\mu\nu}}. \qquad (2.49)$$

This formulation also has its problems. In the case of Einstein–Hilbert action taking the trace of the equations of motion (2.38), with respect to the indices ν and σ, the left side of the equation vanishes identically. Thus if we have a source term this

leads to a constraint for the hypermomentum $\Delta_\sigma{}^{\mu\sigma} = 0$. This means that if we have a matter Lagrangian that does not satisfy this constraint we get inconsistent field equations. This problem can be traced back to the fact that the Einstein–Hilbert action is invariant under projective transformations, but in general the matter action is not, see [17, Sec. 4.2] and references therein. To get consistent field equations with general matter Lagrangians one is required to: impose some restrictions on the connections with e.g. the Lagrange multiplier method [15], consider a higher-order action in the curvature or consider a non-symmetric metric i.e. Einstein-Strauss theories [18, chapter XII]. However, we consider only scalar field matter and in this case the Palatini and metric-affine formulations are equivalent since the scalar field Lagrangian is naturally independent of the connection, (see e.g. [11, 13, 17, 19] for more details of metric-affine theories).

We have shown that Palatini formulation in the case of the Einstein–Hilbert action, makes the connection dynamically the Levi–Civita connection and the Einstein field equations with a source term are obtained, if matter is not coupled to the connection. To be precise, we cannot state that the metric and Palatini formulations are completely equivalent in the case of Einstein–Hilbert action, since they differ by a surface term and additionally the solution of the general connection includes an extra vector field degree of freedom. It seems to be only a coincidence that the formulations produce the same dynamics with the Einstein–Hilbert action and in general the formulations produce different equations of motion. We will see this explicitly in chapters 4 and 5 where we consider modified gravity sectors. Before that we review the case where scalar field drives inflation.

Chapter 3

Inflationary paradigm

GR enabled us to start analyzing the evolution of the universe and together with the SM of particle physics it gives rise to the hot Big Bang model of the universe. The existence of the cosmic microwave background (CMB) and the abundances of light elements are nicely explained by the hot Big Bang model. However, there are problems that the standard hot Big Bang is unable to address.

Inflation can explain many features left unaddressed by the hot Big Bang model. These include: why the universe is so homogeneous and isotropic (the horizon problem) and why it is spatially flat. Most notably, it can explain the origin of the primordial perturbations that seed the structure formation of the universe.

In this chapter we review cosmic inflation keeping emphasis on the generation of primordial perturbations, for more complete introduction see e.g. [20, part IV][21].

3.1 Friedmann-Lemaître-Robertson-Walker model

Inflation quickly makes the universe spatially homogeneous and isotropic to a good approximation. Imposing these two properties implies that we can foliate our spacetime into spacelike slices that are maximally symmetric, decomposing spacetime into a time and a spacelike three dimensional manifold $\mathbb{R} \times \Sigma_t$. The spatial slices are equipped with an induced maximally symmetric metric. There exist a projection tensor $h_{\mu\nu}$ that projects vectors to the spatial slices, which satisfies

$$h_{\mu\nu}u^\nu = 0, \tag{3.1}$$

where u^μ is the normal vector of the hypersurface Σ_t, which can be thought as the time direction and is normalized by $g_{\mu\nu}u^\mu u^\nu = -1$. With these the full metric can be decomposed into a useful form

$$g_{\mu\nu} = h_{\mu\nu} - u_\mu u_\nu, \quad g^{\mu\nu} = h^{\mu\nu} - u^\mu u^\nu. \tag{3.2}$$

The above costraints yields the Friedmann-Lemaître-Robertson-Walker (FLRW) metric [1, ch. 8.2]

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + sin^2\theta d\phi^2)\right], \tag{3.3}$$

where t is the time coordinate, $a(t)$ the scale factor describing the expansion of space, $\{r, \theta, \phi\}$ are the spherical spatial coordinates and k describes the curvature of space. Since inflation drives the universe to be spatially flat, we can further simplify and

set $k = 0$. This enables us to write the FLRW metric in a simple form in Cartesian coordinates

$$ds^2 = -dt^2 + a^2(t) \left[dx^2 + dy^2 + dz^2 \right]. \tag{3.4}$$

The symmetries of the FLRW model restrict the matter content of the universe to be described by an ideal fluid, for which the stress-energy tensor can be written as

$$T_{\mu\nu} = p g_{\mu\nu} + (\rho + p) u_\mu u_\nu, \tag{3.5}$$

where $u_\mu = (-1, 0, 0, 0)$ is the four-velocity of the fluid, ρ and p are the energy density and pressure of the fluid, respectively. With (3.4) and (3.5) plugged into the Einstein equation (2.6) gives the Friedmann equations

$$H^2 = \frac{\rho}{3}, \quad \frac{\ddot{a}}{a} = -\frac{1}{6}(\rho + 3p), \tag{3.6}$$

where $H = \dot{a}/a$ is the Hubble parameter and the dot denotes the time derivative. Thus for accelerated expansion $\ddot{a} > 0$ we need $\rho + 3p < 0$. Next we will see how a scalar field can drive inflation.

3.2 Slow-Roll inflation

Adding a canonical scalar field to (2.4), the action reads

$$S = \int d^4 x \sqrt{-g} \left[\frac{1}{2} R - \frac{1}{2} g^{\mu\nu} \partial_\mu \varphi \partial_\nu \varphi - V(\varphi) \right]. \tag{3.7}$$

Assuming that the scalar field is homogeneous and isotropic, the Friedmann equations turn out to be

$$H^2 = \frac{1}{3} \left[\frac{1}{2} \dot{\varphi}^2 + V(\varphi) \right], \quad \frac{\ddot{a}}{a} = -\frac{1}{3} \left[\dot{\varphi}^2 - V(\varphi) \right], \tag{3.8}$$

and the equations of motion for the scalar field become

$$\ddot{\varphi} + 3H\dot{\varphi} = -V_{,\varphi}(\varphi). \tag{3.9}$$

Thus the condition for inflation is satisfied when $\dot{\varphi}^2 < V$. In slow-roll (SR) approximation we assume $\dot{\varphi}^2 \ll V(\varphi)$. Additionally we assume that the derivative of the field $\dot{\varphi}$ does not change quickly in one Hubble time. This is needed for a sufficiently long period of inflation to occur. The slow-roll conditions are

$$\dot{\varphi}^2 \ll V(\varphi), \tag{3.10}$$

$$|\ddot{\varphi}| \ll |3H\dot{\varphi}|. \tag{3.11}$$

With these conditions satisfied the Friedmann equations (3.8) and (3.9) give the slow-roll equations

$$H^2 = \frac{1}{3} V(\varphi), \quad 3H\dot{\varphi} = -V_{,\varphi}(\varphi). \tag{3.12}$$

Another useful way of writing the requirement for inflation $\ddot{a} > 0$ is

$$\epsilon_H = -\frac{\dot{H}}{H^2} < 1. \tag{3.13}$$

With the slow-roll equations this is approximately

$$\epsilon_H \approx \epsilon_V = \frac{1}{2}\left(\frac{V_{,\varphi}}{V}\right) \overset{SR}{\ll} 1, \tag{3.14}$$

where ϵ_V is the first slow-roll parameter, which describes the quality of the slow-roll approximation. The second slow-roll parameter is

$$|\eta_V| = \left|\frac{V_{,\varphi\varphi}}{V}\right| \overset{SR}{\ll} 1, \tag{3.15}$$

describing the change of the first slow-roll parameter. It can be shown that $\epsilon_V \ll 1$ implies the first slow-roll condition (3.10) and $|\eta_V| \ll 1$ implies (3.11). Smallness of the second and higher-order parameters ($\zeta_V = V_{,\varphi\varphi\varphi}V_{,\varphi}/V^2$ etc.) are needed for long period of inflation. It also turs out that slow-roll is an attractor, meaning that wide range of initial conditions rapidly tends to the slow-roll region, see e.g. [20, sec 18.6.2].

The amount of inflation is described by the number of e-folds N, which can be written as

$$N(t) = \ln\frac{a(t_{end})}{a(t)} = \int_t^{t_{end}} H dt = \int_{\varphi_{end}}^{\varphi} \frac{H}{\dot{\varphi}}d\varphi \overset{SR}{\approx} \int_{\varphi_{end}}^{\varphi} \frac{d\varphi}{\sqrt{2\epsilon_V}}, \tag{3.16}$$

where the subscript 'end' denotes the end of inflation. The end of slow-roll, and to a good approximation the end of inflation, is defined as the point where the first or the second slow-roll parameter becomes of the order of unity, which ever occurs first. Thus the field value at the end of inflation can be solved either from $\epsilon(\varphi_{end}) = 1$ or $|\eta(\varphi_{end})| = 1$. This then allows us to solve the field in terms of N from (3.16). This is useful when evaluating the inflationary observables, which we shall discuss next.

3.3 Perturbations and inflationary observables

Inflation gives us a way of predicting the statistical properties of the primordial perturbations. The perturbations are generated by quantum vacuum fluctuations, which get amplified and stretched to large scales during inflation.

The full calculation, of the statistical properties of the perturbations, uses quantum field theory in curved spacetime. The fields are expanded around the FLRW solution as $\phi \to \phi + \delta\phi$, $g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu}$. The important fact is that perturbations in the metric and Palatini formulations coincide in the case of the Einstein–Hilbert action. For a review of this computation see e.g. [21, Part III]. We will only give a high level overview of the main points and present the results.

Since the perturbations are gauge dependent, care has to be taken over fixing a gauge or constructing gauge invariant variables before proceeding with quantization of the perturbations. For gauge transformation in perturbation theory of gravity see e.g. [22, ch. 5]. Then canonical quantization procedure can be applied: promoting the fields into operators and imposing canonical commutation relations. Due to the curved background, complications arise when constructing the vacuum state, see e.g. [23, ch. 3]. Approximating the background to be close to the static de Sitter space, which is the case during inflation, the vacuum turns out to be the Bunch-Davies vacuum, see e.g. [23, ch. 5]. The perturbations are then assumed to be initially in this adiabatic Bunch-Davies vacuum state. A crucial fact that enables inflationary theories to make predictions about the form of the perturbations is that initially adiabatic perturbations are conserved after Hubble crossing in slow-roll; by Hubble crossing we mean the instance when the wave numbers of the Fourier modes of the perturbations become larger than the Hubble parameter $k = aH$, see e.g. [22, ch. 5.4][24]. This is also the point where the quantum fluctuations get transformed to look like a classical field, see e.g. [25], which allows us to treat the perturbations as classical. In first order perturbation theory scalar, vector and tensor degrees of freedom do not mix; this is sometimes called the decomposition theorem, see e.g. [22, ch. 5]. This allows us to separately calculate the spectrum for scalar and tensor perturbations. It turns out that vector perturbations quickly decay away and can be ignored here.

The gauge invariant quantity describing the scalar perturbations is the so called Sasaki-Mukhanov variable ν, which in the spatially flat gauge is $\nu = a\delta\phi$. The Fourier mode functions ν_k of this variable obey the equations of motion [26]

$$\nu_k'' + \left(k^2 - \frac{z''}{z} \right) \nu_k = 0, \tag{3.17}$$

where $z \equiv a\dot{\phi}/H$, the prime denotes a derivative with respect to the conformal time $d\tau = dt/a$ and k is the comoving wave number of the Fourier mode. This is solved in the slow-roll approximation with the Bunch-Davies vacuum state as the initial condition. These solutions are used to compute the correlation functions that describe the statistical properties of the perturbations. To the first order the perturbations are Gaussian and completely described by the power spectrum (two point correlation function)

$$\mathcal{P}_{\delta\phi} = \frac{1}{a^2}\mathcal{P}_\nu = \frac{k^3}{2\pi^2 a^2}|u_k|^2 = \left(\frac{H}{2\pi} \right)^2 \Bigg|_{k=aH} . \tag{3.18}$$

The comoving curvature perturbation $\mathcal{R}$ turns out to be an useful quantity that can be related to the spectrum measured from the CMB [1]. It is proportional to the

---

[1]Sometimes in the literature another useful quantity is used, the curvature perturbation on uniform-density hypersurfaces ζ. It can be shown that these quantities are equal outside the horizon and during slow-roll.

Sasaki-Mukhanov variable $\mathcal{R} = \nu/z$ [21]. Thus the power spectrum for the curvature perturbation is

$$\mathcal{P}_{\mathcal{R}} = \frac{1}{z^2}\mathcal{P}_\nu = \frac{1}{2\epsilon}\left(\frac{H}{2\pi}\right)^2\Bigg|_{k=aH}, \tag{3.19}$$

where ϵ is the first slow-roll parameter. It is customary to parametrize the spectrum in a nearly scale invariant form

$$\mathcal{P}_{\mathcal{R}} = A_s\left(\frac{k}{k_*}\right)^{n_s-1}, \tag{3.20}$$

where k_* is some reference scale, A_s is the amplitude and n_s is the spectral tilt, which describes the scale dependence of the spectrum. Thus when $n_s = 1$ the spectrum is scale invariant. The amplitude in the slow-roll aproximation can be written as

$$A_s = \frac{1}{24\pi^2}\frac{V}{\epsilon_V} \tag{3.21}$$

and to first order the spectral tilt becomes

$$n_s - 1 \equiv \frac{d\ln\mathcal{P}_{\mathcal{R}}(k)}{d\ln k} \overset{SR}{\approx} -6\epsilon_V + 2\eta_V. \tag{3.22}$$

Inflation also predicts the form of tensor perturbations. Similar calculation as for the scalar perturbations gives the spectrum for the tensor perturbations

$$\mathcal{P}_t = 2\left(\frac{2}{a}\right)^2\mathcal{P}_\nu = 8\left(\frac{H}{2\pi}\right)^2\Bigg|_{k=aH}. \tag{3.23}$$

This can also parametrized in a nearly scale invariant form as

$$\mathcal{P}_t = A_t\left(\frac{k}{k_*}\right)^{n_t}. \tag{3.24}$$

The amplitude in the slow-roll approximation reads

$$A_t = \frac{3}{2\pi^2}V. \tag{3.25}$$

The spectral tilt for this to the first order becomes

$$n_t \equiv \frac{d\ln\mathcal{P}_t}{d\ln k} \overset{SR}{\approx} -2\epsilon_V. \tag{3.26}$$

The amplitude of tensor spectrum is not amplified by the slow-roll parameters, in contrast to the spectrum of the scalar perturbations, and is thus harder to observe. The important quantity that the current observations can set bounds on is the tensor-to-scalar ratio, which in the slow-roll becomes

$$r \equiv \frac{\mathcal{P}_{\mathcal{R}}}{\mathcal{P}_t} \overset{SR}{\approx} 16\epsilon_V. \tag{3.27}$$

We have seen that when a single scalar field drives inflation inflationary predictions can be computed from the inflaton potential alone.

Chapter 4

Higgs inflation

The Standard Model of particle physics (SM) contains one scalar field: the Higgs boson. For this reason it is tempting to identify the Higgs field as the inflaton. However, the Higgs self-coupling ($\lambda \simeq 0.129$) is too large to generate an amplitude for the primordial scalar perturbations that is consistent with current observations. At tree level the calculated amplitude is too large. This remains the case when taking into account quantum corrections to the potential [27, 28, 29].

Adding a non-minimal coupling between the Higgs and gravity can generate a spectrum compatible with current observations. The non-minimal coupling term is well motivated. It is the only new dimension four operator when the Einstein–Hilbert and SM actions are combined. Furthermore, even if the term is absent in the action at classical level this kind of non-minimal coupling term is generated when a scalar field is quantized in a curved background; more specifically when renormalizing the energy-momentum tensor [23, ch. 3, 6].

The feature that only one new parameter is introduced makes this one of the most interesting models to study. Furthermore, the couplings between the Higgs and the rest of the SM are known experimentally, which is advantageous since the reheating period after inflation can be calculated in detail. This is important since the length of the reheating period affects the inflationary observables. In contrast, many models introduce new unknown couplings between the inflaton and the SM.

In Higgs inflation the SM Higgs boson non-minimally coupled to gravity is identified as the inflaton field. The action reads

$$S = \int d^4x\sqrt{-g}\left[(\tfrac{1}{2}M^2 + \xi\mathcal{H}^\dagger\mathcal{H})R - (D_\mu\mathcal{H})^\dagger(D^\mu\mathcal{H}) + m_h\mathcal{H}^\dagger\mathcal{H} - \lambda(\mathcal{H}^\dagger\mathcal{H})^2 + \mathcal{L}_{SM}\right],$$
(4.1)

where M is a mass parameter, ξ is positive dimensionless non-minimal coupling parameter to be fixed by observations, $\mathcal{H}$ is the Higgs doublet, D_μ is the gauge covariant derivative and $\mathcal{L}_{SM}$ is the rest of the standard model. During inflation the rest of the SM acts as spectator fields and do not affect the evolution of the universe. We can write the Higgs doublet in the unitary gauge as $\mathcal{H} = 1/\sqrt{2}\left(\begin{smallmatrix}0\\v+h\end{smallmatrix}\right)$, where $v = 246$ GeV is the vacuum expectation value (vev) of the Higgs field h. We can thus write the relevant part of the action as

$$S = \int d^4x\sqrt{-g}\left[\frac{1}{2}(M^2 + \xi h^2)R - \frac{1}{2}g^{\mu\nu}\partial_\mu h\partial_\nu h - V(h)\right],$$
(4.2)

where $V(h)$ is the symmetry breaking potential

$$V(h) = \frac{\lambda}{4}(h^2 - v^2)^2.$$
(4.3)

The non-minimal coupling to gravity changes the strength of gravitational interaction by effectively changing Newton's gravitational constant, or equivalently the square of Planck mass

$$M_p^2 = M^2 + \xi h^2. \tag{4.4}$$

The Planck mass we measure today in the Jordan frame (the definition of the Jordan frame is discussed in the next section) would then be $M_p^2 = M^2 + \xi v^2$. With a non-minimal coupling in the range $1 \ll \xi \ll M^2/v^2$ we can approximate $M \simeq M_p$, which is what we will do from now on. Additionally we again set M_p to unity. Differences occur only in the large field regime $h \gg M_p/\xi$, which turns out to be the inflationary regime.

The non-minimal coupling makes the computation of inflationary observables more involved. However, we can make use of conformal transformations to remove the non-minimal coupling. Then determining the spectrum of the perturbations will follow the standard computation outlined in section 3.3. Let us next discuss these transformations more closely.

4.1 Conformal transformations

Conformal transformation is a local rescaling of the metric

$$g_{\mu\nu} \to \tilde{g}_{\mu\nu} = \Omega^2(x) g_{\mu\nu}, \tag{4.5}$$

where $\Omega^2(x)$ is a positive smooth function called the conformal factor. This is thus a map between two pseudo-Riemannian spaces. The causal structure is unchanged. Space(Time)like vectors remain space(time)like and null vectors remain as null vectors. Thus null geodesics are left invariant under conformal transformations and light cones are unchanged. Angles between vectors are also unchanged. What conformal transformations do change is the geometry. For example the time-like geodesics generically differ.

To be able to perform this transformation to the action (4.2) we need to compute how the relevant quantities transform. First of all for the inverse metric we have $\tilde{g}_{\lambda\nu}\tilde{g}^{\mu\lambda} = \delta_\nu^\mu$ and thus $\tilde{g}^{\mu\nu} = \Omega^{-2}(x)g^{\mu\nu}$. The metric determinant transforms as

$$\det(\tilde{g}_{\mu\nu}) = \det(\Omega^2 g_{\mu\nu}) = \Omega^{2d} \det(g_{\mu\nu}), \tag{4.6}$$

where d is the number of spacetime dimensions. So we have altogether

$$\begin{aligned} g_{\mu\nu} &= \Omega^{-2}\tilde{g}_{\mu\nu} \\ g^{\mu\nu} &= \Omega^2 \tilde{g}^{\mu\nu} \\ \sqrt{-g} &= \Omega^{-d}\sqrt{-\tilde{g}}. \end{aligned} \tag{4.7}$$

This is all that is needed for the Palatini formulation. For the metric formulation we also need to know how the Ricci scalar transforms under the conformal transformations. First transforming the Levi–Civita connection we find

$$\mathring{\Gamma}^\sigma_{\mu\nu} = \tilde{\mathring{\Gamma}}^\sigma_{\mu\nu} - \left[2\delta^\sigma_{(\nu} \ln \Omega_{,\mu)} - \tilde{g}_{\mu\nu}\tilde{g}^{\sigma\lambda} \ln \Omega_{,\lambda} \right]. \tag{4.8}$$

19

We can plug this into the definition of the Riemann tensor (0.4) and see how it transforms

$$
\begin{aligned}
\mathring{R}^{\sigma}{}_{\mu\lambda\nu} =& \tilde{\mathring{R}}^{\sigma}{}_{\mu\lambda\nu} + 2\tilde{g}_{\mu[\nu}\tilde{g}^{\sigma\rho}\tilde{\nabla}_{\lambda]}\tilde{\nabla}_{\rho}\ln\Omega + 2\delta^{\sigma}_{[\lambda}\tilde{\nabla}_{\nu]}\tilde{\nabla}_{\mu}\ln\Omega + 2\delta^{\sigma}_{[\lambda}\ln\Omega_{,\nu]}\ln\Omega_{,\mu} \\
& 2\tilde{g}_{\mu[\lambda}\delta^{\sigma}_{\nu]}\tilde{g}^{\rho\tau}\ln\Omega_{,\rho}\ln\Omega_{,\tau} + \tilde{g}_{\mu[\nu}\tilde{g}^{\rho\sigma}\ln\Omega_{,\lambda]}\ln\Omega_{,\rho},
\end{aligned}
\tag{4.9}
$$

where quantities with a tilde are constructed from the metric $\tilde{g}_{\mu\nu}$. Taking the contraction we get how the Ricci tensor transforms

$$
\begin{aligned}
\mathring{R}_{\mu\nu} =& \tilde{\mathring{R}}_{\mu\nu} + \tilde{g}_{\mu\nu}\tilde{\Box}\ln\Omega + (d-2)\tilde{\nabla}_{\nu}\tilde{\nabla}_{\mu}\ln\Omega \\
& + (d-2)\ln\Omega_{,\mu}\ln\Omega_{,\nu} + (2-d)\tilde{g}_{\mu\nu}\tilde{g}^{\rho\tau}\ln\Omega_{,\rho}\ln\Omega_{,\tau}.
\end{aligned}
\tag{4.10}
$$

Contracting with the metric $g^{\mu\nu}$ we get how the Ricci scalar transforms

$$
\mathring{R} = \Omega^2\left[\tilde{\mathring{R}} + 2(d-1)\tilde{\Box}\ln\Omega + (d-1)(2-d)\tilde{g}^{\rho\tau}\ln\Omega_{,\rho}\ln\Omega_{,\tau}\right].
\tag{4.11}
$$

It is also straightforward to derive how higher-order curvature terms (discussed later in chapter 5) transform, but with the higher-order terms we focus on the Palatini formulation. Considering higher-order terms in metric formulation is out of the scope of this thesis. In contrast in the Palatini formulation the situation is simple, since the Riemann tensor does not depend on the metric and is thus unchanged. All the different invariants that are second order in curvature, that we will encounter in chapter 5, are invariant under conformal transformations (in 4 dimensions). For example

$$
\begin{aligned}
\sqrt{-g}g^{\mu\lambda}g^{\nu\sigma}R_{\mu\nu}R_{\lambda\sigma} &= \left[-\Omega^{-2\times4}\tilde{g}\right]^{1/2}\Omega^4\tilde{g}^{\mu\lambda}\tilde{g}^{\nu\sigma}R_{\mu\nu}R_{\lambda\sigma} \\
&= \sqrt{-\tilde{g}}\tilde{g}^{\mu\lambda}\tilde{g}^{\nu\sigma}R_{\mu\nu}R_{\lambda\sigma}.
\end{aligned}
\tag{4.12}
$$

The usefulness of these transformations is due to the fact that two different-looking actions related by a conformal transformation are physically equivalent at the classical level[1]. By performing a conformal transformation we thus find a different representation for the theory. Customarily in the literature these different representations that are related by conformal transformations are called different frames. There are two specific important frames that have their own names. The so called Jordan frame, which is characterized by the fact that there is a non-minimal coupling between a scalar field and the curvature term. The other important frame is the so called Einstein frame, where matter is minimally coupled to the curvature. The physical equivalence of these two frames is clear from the fact that the transformation is required to be invertible and thus it gives a one-to-one correspondence between these two frames. The transformation can also be thought simply as a local change of units [30]. When comparing quantities between two frames attention has to be paid to appropriately account for the conformal factor.

The inflationary observables, both the curvature and tensor perturbations, can be shown to be invariant under conformal transformations [31, 32]. The inflationary

---

[1]When quantizing the theory this point is less clear.

observables can thus be calculated in any frame. In the case of Higgs inflation inflationary observables have also been computed in both frames and the results agree, see [33] and references therein.

Transforming the action (4.2) to the Einstein frame is different in the metric and Palatini formulation due to the distinct metric dependence of the actions. Let's now perform this transformation in both cases.

4.1.1 Einstein frame in metric formulation

Let's first consider the metric formulation, where $R(\Gamma) = \mathring{R}(g)$ in (4.2). From the action (4.2) we see that we can remove the non-minimal coupling by choosing the conformal factor to be

$$\Omega^2(x) = 1 + \xi h^2. \tag{4.13}$$

With the previously obtained transformations the action (4.2) becomes

$$
\begin{aligned}
S = \int d^4x \sqrt{-\tilde{g}}\,\Omega^{-4} \Bigg\{ &\frac{1}{2}\Omega^2\Omega^2 \left[\tilde{\mathring{R}} + 3\tilde{\Box}\ln\Omega^2 - \frac{3}{2}\tilde{g}^{\rho\lambda}\ln\Omega^2_{,\rho}\ln\Omega^2_{,\lambda} \right] \\
&- \frac{1}{2}\Omega^2\tilde{g}^{\mu\nu}\partial_\mu h\partial_\nu h - V(h) \Bigg\},
\end{aligned}
\tag{4.14}
$$

where the second term inside the square brackets is a total derivative. This term can be converted to a surface term using the Stokes' theorem (2.13) and thus vanishes. Writing

$$\ln\Omega^2_{,\lambda} = \frac{(\Omega^2)_{,\lambda}}{\Omega^2} = \frac{\Omega^2_{,h}}{\Omega^2}\partial_\lambda h, \tag{4.15}$$

the action simplifies to

$$S = \int d^4x \sqrt{-\tilde{g}} \left\{ \frac{1}{2}\tilde{\mathring{R}} - \frac{1}{2}\frac{\frac{3}{2}(\Omega^2)^2_{,h} + \Omega^2}{\Omega^4}\tilde{g}^{\mu\nu}\partial_\mu h\partial_\nu h - \Omega^{-4}V(h) \right\}. \tag{4.16}$$

We see that we have been able to remove the non-minimal coupling between gravity and the scalar field. The cost for this is a modified matter sector with a non-canonical kinetic term. The kinetic term can be brought into a canonical form by a field redefinition

$$\frac{d\chi}{dh} = \sqrt{\frac{\frac{3}{2}(\Omega^2)^2_{,h} + \Omega^2}{\Omega^4}}. \tag{4.17}$$

With this the action becomes

$$S = \int d^4x \sqrt{-\tilde{g}} \left\{ \frac{1}{2}\tilde{\mathring{R}} - \frac{1}{2}\tilde{g}^{\mu\nu}\partial_\mu\chi\partial_\nu\chi - U(\chi) \right\}, \tag{4.18}$$

where we defined the modified potential

$$U(\chi) \equiv \frac{V(h(\chi))}{\Omega^4(h(\chi))}. \tag{4.19}$$

Note that the h field has to be solved in terms of the new field from the redefinition (4.21), we will come to this later.

4.1.2 Einstein frame in the Palatini formulation

In the Palatini formulation this is considerably easier since the Ricci tensor is unchanged. We can again get rid of the non-minimal coupling by the same conformal factor (4.13). The action (4.2) becomes

$$S = \int d^4x \sqrt{-\widetilde{g}} \left[\frac{1}{2} \widetilde{g}^{\mu\nu} R_{\mu\nu} - \frac{1}{2} \frac{1}{\Omega^2} \widetilde{g}^{\mu\nu} \partial_\mu h \partial_\nu h - \frac{V(h)}{\Omega^4} \right]. \tag{4.20}$$

The non-canonical kinetic term can be again brought in to canonical form by a field redefinition, which now simply reads

$$\frac{d\chi}{dh} = \sqrt{\frac{1}{\Omega^2}}. \tag{4.21}$$

So the action can be written

$$S = \int d^4x \sqrt{-\widetilde{g}} \left\{ \frac{1}{2} \widetilde{R} - \frac{1}{2} \widetilde{g}^{\mu\nu} \partial_\mu \chi \partial_\nu \chi - U(\chi) \right\}, \tag{4.22}$$

with

$$U(\chi) \equiv \frac{V(h(\chi))}{\Omega^4(h(\chi))}. \tag{4.23}$$

We see that the metric and Palatini formulations differ by the form of the potential coming from the fact that the field redefinitions differ. We will see this difference more explicitly in the following section.

4.2 Tree-level inflationary observables

The tree-level inflationary observables are now straightforward to compute in the Einstein frame. As we saw in chapter 3, these can be computed from the potential, which are different between the metric and Palatini formulations. To get the potential we need to solve the new field variable χ from the field redefinition. Let's again start with the metric formulation.

4.2.1 Metric formulation

It is possible to solve χ from the field redefinition (4.21) analytically [33]. However, the form of the solution is cumbersome and not particularly illuminating. We can solve it approximately by assuming that $\xi \gg 1$, which we shall later see is the case in Higgs inflation. Additionally in the inflationary regime the field value is large

$h \gg 1/\xi$. The field redefinition (4.21) is then approximately [2]

$$\frac{d\chi}{dh} = \sqrt{\frac{1 + (1 + 6\xi)\xi h^2}{(1 + \xi h^2)^2}}$$
$$\simeq \frac{\sqrt{6}\xi h}{1 + \xi h^2}. \tag{4.24}$$

Which when solved gives

$$h^2 = \frac{1}{\xi}\left(e^{\frac{2}{\sqrt{6}}\chi} - 1\right). \tag{4.25}$$

Solving for χ gives

$$\chi = \sqrt{\frac{3}{2}}\ln\left(1 + \xi h^2\right) = \sqrt{\frac{3}{2}}\ln\Omega(h)^2. \tag{4.26}$$

We also see that in the small field regime $h \ll 1/\xi$ we simply have $h \simeq \chi$. So the field redefinition (4.21) is approximately solved in different asymptotic regions as

$$\chi \simeq \begin{cases} h & \text{, when } h \ll 1/\xi \\ \sqrt{\frac{3}{2}}\ln\Omega(h)^2 & \text{, when } h \gg 1/\xi \end{cases} \tag{4.27}$$

In the large field region the Higgs vev is much smaller than the field value and we can approximate the Jordan frame potential (4.3) as

$$V(h) \simeq \frac{\lambda}{4}h^4. \tag{4.28}$$

With these, in the large field regime, the potential (4.19) becomes

$$U(\chi) = \frac{\lambda}{4\xi^2}\left(1 - e^{-\frac{2}{\sqrt{6}}\chi}\right)^2. \tag{4.29}$$

From this we see that the potential is asymptotically flat: $U(\chi) \to \lambda/(4\xi^2)$ when $\chi \to \infty$. Thus there is a plateau in the large field regime where slow-roll inflation can occur. This also implies that the action has an asymptotic shift symmetry, i.e. it is symmetric under the shift of the inflaton field ($\chi \to \chi + $ constant) in the limit $\chi \to \infty$.[3] This approximate symmetry turns out to be important when considering quantum corrections [35].

We can now calculate the inflationary observables in the slow-roll approach from the potential. The first slow-roll parameters are

$$\epsilon_U = \frac{1}{2}\left(\frac{U_{,\chi}}{U}\right)^2 = \frac{4}{3}\left(e^{\frac{2}{\sqrt{6}}\chi} - 1\right)^{-2} = \frac{4}{3\xi^2 h^4} \tag{4.30}$$

$$\eta_U = \frac{U_{,\chi\chi}}{U} = \frac{4}{3}\left(2 - e^{\frac{2}{\sqrt{6}}\chi}\right)\left(e^{\frac{2}{\sqrt{6}}\chi} - 1\right)^{-2} = \frac{4}{3\xi^2 h^4}\left(1 - \xi h^2\right). \tag{4.31}$$

---

[2]When this was first done in the literature [34] the approximation $h \gg 1/\sqrt{\xi}$ was made which results in a different form for the potential; the overall approximate form is different but the asymptotic behavior is the same. The first slow-roll parameter turns out to be the same but second and higher order slow-roll parameters are different.

[3]In the Jordan frame this asymptotic symmetry manifests itself as a asymptotic scale symmetry: $x^\mu \to \alpha x^\mu$ and $h(x) \to \alpha^{-1}h(\alpha x)$ where α is a constant.

We can get the field value at the end of inflation defined by the usual condition, which in this case is $\epsilon(\chi_{end}) = 1$. Equating (4.30) to unity and solving for χ gives the field value at the end of inflation

$$\chi_{end} = \sqrt{\frac{3}{2}} \ln\left(1 + \sqrt{\frac{3}{4}}\right). \tag{4.32}$$

As described in the section 3.3 the observables are to be evaluated at a horizon crossing. We can do this by solving the field in terms of the number of e-folds $\chi(N_*)$, where N_* is the number of e-folds when some reference scale k_* crosses the horizon $k_* = a_* H_*$. From (3.16) and (4.30) we have

$$N_* = \int_{\chi_{end}}^{\chi_*} \frac{d\chi}{\sqrt{2\epsilon}} = \int_{\chi_{end}}^{\chi_*} \frac{U}{U_{,\chi}} d\chi = \int_{\chi_{end}}^{\chi_*} \frac{1}{2}\sqrt{\frac{3}{2}} \left(e^{\frac{2}{\sqrt{6}}\chi} - 1\right) d\chi, \tag{4.33}$$

giving us

$$N_* = \frac{3}{4}\left[\exp\left(\frac{2}{\sqrt{6}}\chi_*\right) - \exp\left(\frac{2}{\sqrt{6}}\chi_{end}\right)\right] - \frac{\sqrt{6}}{4}\left[\chi_* - \chi_{end}\right]. \tag{4.34}$$

Plugging in the field value at the end of inflation (4.32) we get

$$N_* = \frac{1}{4}\left[3\exp\left(\frac{2}{\sqrt{6}}\chi_*\right) - \sqrt{6}\chi_*\right] - \frac{3}{4}\left[\ln\left(1 + \sqrt{\frac{3}{4}}\right) - \left(1 + \sqrt{\frac{3}{4}}\right)\right], \tag{4.35}$$

where the numerical factor is of the order of unity. Assuming that the exponential term dominates we finally get

$$N_* \simeq \frac{3}{4}\exp\left(\frac{2}{\sqrt{6}}\chi_*\right). \tag{4.36}$$

It is also possible solve χ_* in terms of N_* and χ_{end} analytically. The approximated form is easier to work with, which gives

$$\chi_*(N_*) = \sqrt{\frac{3}{2}} \ln\left(\frac{4}{3}N_*\right). \tag{4.37}$$

With this we can write the inflationary observables in terms of the number of e-folds. The slow-roll parameters are

$$\epsilon(N_*) = \frac{12}{(4N_* - 3)^2} \simeq \frac{3}{4N_*^2} \tag{4.38}$$

$$\eta(N_*) = 8\frac{3 - 2N_*}{(4N_* - 3)^2} \simeq -\frac{1}{N_*}, \tag{4.39}$$

where we approximated further by assuming $4N_* \gg 1$. The amplitude of the scalar perturbations is

$$A_s = \frac{1}{24\pi^2}\frac{\lambda(4N_* - 3)^4}{768\xi^2 N_*^2} \simeq \frac{1}{24\pi^2}\frac{\lambda N_*^2}{3\xi^2}. \tag{4.40}$$

The spectral tilt of scalar perturbations is

$$n_s = 1 - \frac{8(3 + 4N_*)}{(4N_* - 3)^2} \simeq 1 - \frac{2}{N_*}. \tag{4.41}$$

The spectral tilt of tensor perturbations is

$$n_t = \frac{-24}{(4N_* - 3)^2} \simeq -\frac{3}{2N_*^2}. \tag{4.42}$$

The tensor-to-scalar ratio is

$$r = \frac{192}{(4N_* - 3)^2} \simeq \frac{12}{N_*^2}. \tag{4.43}$$

Analytical expressions, with out the approximations made here, for these observables can be found in [33]. The precise number of e-folds depends on the heating period after inflation, which advantageously can be calculated in detail since the SM parameters are experimentally known. These computations have been carried out in the literature [36, 37, 38]. We use the simplifying assumption that the reheating period was instant, which gives us an estimate of $N_* \simeq 55$ when the usual reference scale $k_* = 0.05 \quad \text{Mpc}^{-1}$ passes the horizon. We can now find an estimate for the non-minimal coupling by matching the amplitude (4.40) with the observed value [39]

$$\ln\left(10^{10} A_s\right) \simeq 3.094 \pm 0.034. \tag{4.44}$$

This gives a relation between the non-minimal coupling and the Higgs self-coupling. With the previous estimate for the number of e-folds we have

$$\xi \simeq 800 N_* \sqrt{\lambda} \simeq 44000 \sqrt{\lambda} \tag{4.45}$$

We see that the non-minimal coupling is required to be quite large since $\lambda \simeq 0.129$, but still $\xi \ll M_p^2/v^2 \sim 10^{32}$ as we previously required. It is also smaller than the upper bound $\xi < 2.6 \times 10^{15}$ set by considering observations about the Higgs at the LHC [40].

We can now also get numerical values for the important observables

$$n_s \simeq 0.964, \quad r \simeq 0.00396, \tag{4.46}$$

which are in good agreement with the current observed values [39].

However, this model is not without problems. The non-minimal coupling to gravity makes this model non-renormalizable in the Jordan frame. Written in terms of canonically normalized variables $g_{\mu\nu} = \eta_{\mu\nu} + M_p^{-1} h_{\mu\nu}$ in the weak field regime $\phi^2 R$ results in a dimension 5 term and is thus non-renormalizable. In the Einstein frame the non-renormalizability comes from the nonlinear form of the interactions in the potential. Non-renormalizability implies that there is some cutoff scale where perturbation theory breaks down, and in the effective field theory point of view one would have to add higher dimension terms in the scalar field to the action suppressed

by the cutoff. These higher-order terms could spoil the inflationary plateau. Hence, it is important to estimate this cutoff in order to figure out whether the theory is consistent in the inflationary region or does it need to be modified. One way of estimating this cutoff is to see where tree-level unitarity is violated, see e.g. [41, ch. 24]. Estimating the cutoff by expanding only the metric around a background leads to a cutoff which is right around the inflationary scale, and thus seems to lead to inconsistencies [42]. However, during inflation both metric and the scalar are naturally expanded around a background. In [35] the cutoff was estimated by expanding both the metric and the scalar field around a background, in both Einstein and Jordan frames. This results in a field dependent cutoff, which is parametrically much larger than the Hubble scale during inflation and is equal to the Planck scale at small field values. These considerations suggest that our tree-level analysis is consistent, however this is still an open question.

4.2.2 Palatini formulation

In the Palatini formulation the analysis follows the same steps. The field redefinition is now simpler

$$\frac{d\chi}{dh} = \frac{1}{\sqrt{\Omega^2}} = \frac{1}{\sqrt{1 + \xi h^2}}, \tag{4.47}$$

which is easily integrated to give

$$\chi(h) = \frac{1}{\sqrt{\xi}} \sinh^{-1}\left(\sqrt{\xi}h\right), \tag{4.48}$$

$$h(\chi) = \frac{1}{\sqrt{\xi}} \sinh\left(\sqrt{\xi}\chi\right). \tag{4.49}$$

Again the Higgs vev is negligible in the large field regime and we take the Jordan frame potential to be (4.28). Plugging (4.48) into the potential (4.23) gives

$$U(\chi) = \frac{\lambda}{4\xi^2} \tanh^4\left(\sqrt{\xi}\chi\right), \tag{4.50}$$

which is asymptotically flat: $\chi \to \infty$, $U(\chi) \to \lambda/(4\xi^2)$ and has a plateau in the large field regime where slow-roll can occur. The slow-roll parameters can be approximated in the large field regime $\sqrt{\xi}\chi \gg 1$ as

$$\epsilon \simeq 128\xi \exp\left(-4\sqrt{\xi}\chi\right), \tag{4.51}$$

$$\eta \simeq -32\xi \exp\left(-2\sqrt{\xi}\chi\right), \tag{4.52}$$

The second slow-roll parameter becomes of order of unity before the first parameter. Thus we get the field value at the end of inflation from $|\eta(\chi_{end})| = 1$. This gives

$$\chi_{end} = \frac{1}{2\sqrt{\xi}} \ln\left(32\xi\right). \tag{4.53}$$

Next we need to solve the field in terms of e-folds. Starting with

$$N_* = \int_{\chi_{end}}^{\chi_*} \frac{U}{U_{,\chi}} d\chi = \int_{\chi_{end}}^{\chi_*} \frac{1}{8\sqrt{\xi}} \sinh\left(2\sqrt{\xi}\chi\right) d\chi$$
$$= \frac{1}{16\xi} \cosh\left(2\sqrt{\xi}\chi\right)\bigg|_{\chi_{end}}^{\chi_*}, \tag{4.54}$$

which gives

$$N_* = \frac{1}{32\xi}\left[\exp\left(2\sqrt{\xi}\chi_*\right) - \exp\left(2\sqrt{\xi}\chi_{end}\right)\right]. \tag{4.55}$$

Plugging in the field value at the end of inflation (4.53) into the previous expression

$$N_* = -1 + \frac{1}{32\xi}\exp\left(2\sqrt{\xi}\chi_*\right), \tag{4.56}$$

then with the assumption that $N_* \gg 1$ we get[4]

$$N_* \simeq \frac{1}{32\xi}\exp\left(2\sqrt{\xi}\chi\right). \tag{4.57}$$

Allowing us to solve the field in terms of the e-folds

$$\chi(N_*) = \frac{1}{2\sqrt{\xi}}\ln\left(32\xi N_*\right). \tag{4.58}$$

The slow-roll parameters in terms of e-folds become

$$\epsilon(N_*) \simeq \frac{1}{8N_*^2\xi} \tag{4.59}$$

$$\eta(N_*) \simeq -\frac{1}{N_*}. \tag{4.60}$$

Which allows us to write the observables in terms of the e-folds

$$A_s \simeq \frac{1}{24\pi^2}\frac{2\lambda N_*^2}{\xi}, \tag{4.61}$$

$$n_s \simeq 1 - \frac{2}{N_*}, \tag{4.62}$$

$$n_t \simeq -\frac{1}{4N_*^2\xi}, \tag{4.63}$$

$$r \simeq \frac{2}{N_*^2\xi}, \tag{4.64}$$

where we again approximated $N_* \gg 1$. The reheating stage after inflation is also different in the Palatini case and it has been studied in [44]. There it was found

---

[4]In the ref. [43] the assumption $\chi_* \gg \chi_{end}$ is made, which in fact result in the same approximate form (4.57) and thus does not change the results that we obtain. However, when looking at (4.53) and (4.58) we see that χ_*, χ_{end} are of the same order. In the reference [43] the end of inflation is defined to be at $\epsilon = 1$, which leads χ_* to be about twice the value of χ_{end}.

that the reheating stage is almost instant, which increases the required number of e-folds. The number of e-folds is estimated to be $N_* \simeq 50$. The number of e-folds depends on the value of the non-minimal coupling ξ. With the range $\xi = 10^6 \dots 10^9$ we have $N_* \simeq 51 \dots 50$. We can now match the observed amplitude of the scalar perturbations (4.44). This now requires a much larger non-minimal coupling

$$\xi \simeq 3.8 \times 10^6 N_*^2 \lambda \simeq 3.9 \times 10^9 \lambda, \tag{4.65}$$

which is many orders of magnitude larger than in the metric formulation. However, it is still much smaller than what we required $\xi \ll M_p^2/v^2 \sim 10^{32}$ and smaller than the LHC upper bound $\xi < 2.6 \times 10^{15}$.

We see that the tilt of the scalar spectrum has the same form to the first order as in the metric case, but with the smaller estimated number of e-folds has a slightly smaller value. The main difference comes in the tensor-to-scalar ratio, which is suppressed by the non-minimal coupling. Future planned experiments could thus rule out Palatini-Higgs inflation if significant tensor-to-scalar ratio is observed. With our estimates the important observables have the values

$$n_s \simeq 0.960, \quad r \simeq 2.1 \times 10^{-13} \lambda^{-1}, \tag{4.66}$$

which are still compatible with observations.

The unitarity violation is less of a problem in Palatini formulation. A straightforward estimation of the cutoff turns out to be parametrically larger than the scale during inflation, for details see [45] and references therein.

Now that we have reviewed the Higgs inflation in Palatini formulation we are ready to consider higher-order curvature terms and Higgs inflation. We shall begin by reviewing some basic properties of higher-order gravity.

Chapter 5

Higher-order gravity

GR has been well tested on different scales, since its formulation, and it has been able to predict new phenomena. Despite the unquestioned success of the theory there are good reasons to consider modified theories of gravity. From a theoretical standpoint the lack of theory of quantum gravity suggests that this is not the full story. Moreover, the non-renormalizability of GR (see e.g.[41, Ch. 22.4]), further suggest that GR is a effective field theory of some underlying more complete theory.

There are many ways to start modifying GR (for a exhaustive survey in view of cosmology see e.g. [46]). Here we are interested in higher-order gravity theories, where terms that are higher-order in the curvature are added to the action. Existence of these terms can be motivated in several ways: considering relativistic QFT in classical curved background generates terms to the action which are quadratic in the curvature tensor [47][23]. Additionally, Weinberg and Deser suggested that adding quadratic curvature terms to the action makes the theory renormalizable[1], which was later proven [49]. Also these kind of terms are generated in string theories as low energy effective actions at classical level, see [50, 51][52, Ch. 7].

However, adding these kinds of terms introduces new kinds of problems. In the case of metric formulation the equations of motion contain higher-order derivative terms of the form, which by the Ostrogradsky theorem [53, Ch 2.5], lead to instabilities. By instabilities we mean that there exists solutions for which the energies are not bounded from below and can thus have arbitrary negative energies[2]. These negative energy modes are often called ghosts in the literature. These lead into serious difficulties when trying to interpret the theory.

In the Palatini formulation appearance of ghosts is more unclear. Naively one might think that since in Palatini formulation the resulting equations of motion are only second order, and thus there are no Ostrogradsky instabilities, there would not be any ghosts. However, by analyzing the degrees of freedom of some quadratic curvature terms it has been shown that ghost modes can still appear [55]. The question, what is the most general form of the Lagrangian where ghosts do not appear, requires further research.

In this chapter we will consider the addition of quadratic curvature invariants in both the metric and Palatini formulations and derive the equations of motion for these terms.

---

[1] Renormalization is out of the scope of this thesis. For an introduction see e.g. [41, Part III][48, Part II]

[2] It has been suggested that with higher-order terms all the unstable solutions are unphysical and all the physical solutions remain stable in semiclassical limit [54].

5.1 $f(R)$ gravity

To illustrate the differences between the metric and Palatini formulations arising from adding higher-orders curvature terms to the action, we first consider adding higher powers of the Ricci scalar. We can do this in general by considering a general function of the Ricci scalar $f(R)$. This is called $f(R)$ gravity, for a review see e.g. [19]. It is well known that in both the metric and Palatini formulations these kind of theories are equivalent to GR coupled to a scalar field known as scalar-tensor theories[3] [56]. It is also known that with the condition $f''(R) \geq 0$ (prime denotes the derivative with respect to the Ricci scalar R) these theories are ghost free [19, Ch. V, and references therein].

Let's see how the differences between the metric and Palatini formulation appear.

5.1.1 In the metric formulation

The action under consideration is

$$S = \int d^4x \sqrt{-g} f(\mathring{R}). \tag{5.1}$$

The steps for taking the variation of this action are very similar as those we skimmed over in section 2.1.1. To deal with the boundaries without imposing extra constraints YGH-term needs to be added similarly as in section 2.1.2 [4]. The differences occur when performing partial integrations, since this now introduces second order derivative terms of $f'(\mathring{R})$. The resulting equations of motion are

$$f'(\mathring{R})\mathring{R}_{\mu\nu} - \frac{1}{2}f(\mathring{R})g_{\mu\nu} - \mathring{\nabla}_\mu\mathring{\nabla}_\nu f'(\mathring{R}) + g_{\mu\nu}\mathring{\Box} f'(\mathring{R}) = 0, \tag{5.2}$$

where $\mathring{\Box} \equiv \mathring{\nabla}_\mu\mathring{\nabla}^\mu$ is the d'Alembertian operator. These are fourth order partial differential equations of the metric, if $f'(\mathring{R})$ is dynamical; and reduce to second order equations, if $f'(\mathring{R})$ is a constant.

The difference between the metric and Palatini formulation is more easily seen when moving to the Einstein frame. If we assume that $f''(\mathring{R}) \neq 0$ we can introduce a new scalar field ϕ and write the action as

$$S = \int d^4x \sqrt{-g} \left[f(\phi) + f'(\phi)(\mathring{R} - \phi) \right]. \tag{5.3}$$

We see that the variation with respect to the field ϕ leads to a constraint $\mathring{R} = \phi$ and thus we recover equivalent equations of motion to the original action. Now with a

---

[3]Note that this equivalence does not apply in metric-affine formulation.

[4]The exact form of the term differs from the one in the section 2.1.2. The term needed is

$$S_{YGH} = 2\int_{\partial\Sigma} d^3x \epsilon \sqrt{\gamma} f'(\mathring{R})\mathring{\nabla}_\mu n^\mu,$$

for a detailed derivation see [7].

simple change of variables $\Phi(\phi) = f'(\phi)$, and with the requirement that the inverse $\phi(\Phi)$ exists, we can write $V(\Phi) = \phi(\Phi) - f(\phi(\Phi))$, which brings the action to the form

$$S = \int d^4x\sqrt{-g}\left[\Phi\mathring{R} - V(\Phi)\right].\tag{5.4}$$

There now appears a non-minimal coupling between the scalar and gravity. We can now remove this non-minimal coupling by performing a conformal transformation, as we have done before. With

$$g_{\mu\nu} \rightarrow \Phi g_{\mu\nu},\tag{5.5}$$

and using relations from section 4.1, the action can be written in Einstein frame

$$S = \int d^4x\sqrt{-g}\left[\mathring{R} - \frac{3}{2\Phi^2}\partial^\mu\Phi\partial_\mu\Phi - \Phi^{-2}V(\Phi)\right],\tag{5.6}$$

where a surface term was dropped. Thus we see that in the metric formulation there appears a new gravitational degree of freedom.

5.1.2   In the Palatini formulation

Next we will see how the equations of motion turn out in the Palatini formulation. Variation of the action (5.1) with respect to the metric gives

$$f'(R)R_{(\mu\nu)} - \frac{1}{2}f(R)g_{\mu\nu} = 0.\tag{5.7}$$

Variation of the action with respect to the connection follows the same steps as we did in section 2.2.1. Thus the equations of motion for the connection are

$$\begin{aligned}&\nabla_\lambda(\sqrt{-g}f'(R)g^{\nu\lambda})\delta^\mu_\sigma - \nabla_\lambda(\sqrt{-g}f'(R)g^{\mu\nu})\\&+ \sqrt{-g}f'(R)\left[g^{\mu\nu}T^\lambda_{\;\lambda\sigma} - g^{\nu\lambda}T^\rho_{\;\rho\lambda}\delta^\mu_\sigma + g^{\nu\lambda}T^\mu_{\;\sigma\lambda}\right] = 0.\end{aligned}\tag{5.8}$$

Let us now move to the Einstein frame. Often in the literature the torsion is assumed to vanish; this is not, however, necessary for finding the Einstein frame. We start by noticing that the equations of motion for the connection (5.8) with the change of variables $q^{\mu\nu} = f'(R)g^{\mu\nu}$ are the same as for the Einstein–Hilbert action. We know the general solution for this equation, namely (2.41). Thus (5.8) has the general solution

$$\Gamma^\sigma_{\mu\nu} = \mathring{\Gamma}^\sigma_{\mu\nu}(q) + \delta^\sigma_\nu V_\mu,\tag{5.9}$$

where $\mathring{\Gamma}^\sigma_{\mu\nu}(q)$ is the Levi–Civita connection for the metric $q^{\mu\nu}$ and V_μ is an arbitrary vector field. The action is invariant under the projective transformations (2.42), and thus the arbitrary vector field does not come into the action. We can thus say that $R_{\mu\nu}$ is equivalent to the Ricci scalar with Levi–Civita connection of the metric $q^{\mu\nu}$, $R_{\mu\nu} = \mathring{R}_{\mu\nu}(q)$. So the action becomes

$$S = \int d^4x\sqrt{-g}\left[\Phi g^{\mu\nu}\mathring{R}_{\mu\nu}(q) - V(\Phi)\right].\tag{5.10}$$

We can now perform a conformal transform to get rid of the $q^{\mu\nu}$ dependence, since $g^{\mu\nu} = \Phi^{-1}q^{\mu\nu}$. Using the relations from section 4.1 and performing the conformal transformation the action can be written as

$$S = \int d^4x\sqrt{-g}\left[\Phi\mathring{R} + \frac{3}{2\Phi}\partial_\mu\Phi\partial^\mu\Phi - V(\Phi)\right], \tag{5.11}$$

where a surface term was ignored. There is still a non-minimal coupling present, so we do a another conformal transformation to the Einstein frame. With (5.5) the action becomes

$$S = \int d^4x\sqrt{-g}\left[\mathring{R} + \frac{3}{2\Phi^2}\partial_\mu\Phi\partial^\mu\Phi - \frac{3}{2\Phi^2}\partial_\mu\Phi\partial^\mu\Phi - \Phi^{-2}V(\Phi)\right]. \tag{5.12}$$

The kinetic term vanishes and the action becomes

$$S = \int d^4x\sqrt{-g}\left[\mathring{R} - \Phi^{-2}V(\Phi)\right] + S_m\left[\Phi g_{\mu\nu}, \Psi_i\right], \tag{5.13}$$

where we wrote explicitly the matter part of the action to illustrate that in the Palatini formulation there is no new degrees of freedom, since the scalar Φ does not have a kinetic term, but the relation between existing degrees of freedom is altered.

From the Einstein frame action (5.6) in the metric formulation we observe there is one extra degree of freedom. In contrast from the Einstein frame action (5.13) in the Palatini formalism we see that there are no new degrees of freedom[5].

5.2 Quadratic gravity in metric formulation

Let us next investigate the other possible quadratic curvature terms, starting with the metric formulation.

Due to the high number of symmetries of the Riemann tensor, in the metric formulation, there are only a few independent terms that can be written down, namely: $\mathring{R}^2$, $\mathring{R}_{\mu\nu}\mathring{R}^{\mu\nu}$ and $\mathring{R}_{\mu\nu\sigma\lambda}\mathring{R}^{\mu\nu\sigma\lambda}$ (there is one more possible term $\epsilon^{\mu\nu\sigma\rho}\mathring{R}^{\alpha\beta}{}_{\mu\nu}\mathring{R}_{\alpha\beta\sigma\rho}$, but this turns out to be a total derivative).

Equations of motion for the $\mathring{R}^2$ term can be easily read from the $f(\mathring{R})$ gravity equations of motion (5.2). Derivation of the equations of motion for the last two terms is straightforward but somewhat lengthy. We give some needed tricks and just state the results. Keep in mind that when raising and lowering the indices of the variation: $\delta g_{\mu\nu} = -g_{\mu\alpha}g_{\nu\beta}\delta g^{\alpha\beta}$. We make use of the Bianchi identities $\mathring{R}_{\mu\nu[\sigma\lambda;\rho]} = 0$ (more specifically the contracted Bianchi identity) and the definition of how the commutator of two covariant derivatives $[\mathring{\nabla}_\mu, \mathring{\nabla}_\nu]$ act on a general tensor (see e.g. [1, p. 123]). With these in mind and ignoring surface terms the equations of motion

---

[5]Comparing the Jordan frame actions, the metric case (5.4) is of the form of Brans-Dicke action with $\omega_0 = 0$ and the Palatini case $\omega_0 = -3/2$

turn out to be[6]

$$\frac{\delta(\mathring{R}^2)}{\sqrt{-g}\delta g^{\mu\nu}} = -\frac{1}{2}\mathring{R}^2 g_{\mu\nu} + 2\mathring{R}\mathring{R}_{\mu\nu} - 2\mathring{\nabla}_\nu\mathring{\nabla}_\mu\mathring{R} + 2g_{\mu\nu}\mathring{\Box}\mathring{R} = 0, \tag{5.14}$$

$$\frac{\delta(\mathring{R}_{\alpha\beta}\mathring{R}^{\alpha\beta})}{\sqrt{-g}\delta g^{\mu\nu}} = -\frac{1}{2}g_{\mu\nu}\mathring{R}_{\alpha\beta}\mathring{R}^{\alpha\beta} + \mathring{\Box}\mathring{R}_{\mu\nu} + \frac{1}{2}g_{\mu\nu}\mathring{\Box}\mathring{R}$$
$$- \mathring{\nabla}_\nu\mathring{\nabla}_\mu\mathring{R} + 2\mathring{R}^{\lambda\sigma}R_{\mu\lambda\nu\sigma} = 0, \tag{5.15}$$

$$\frac{\delta(\mathring{R}_{\alpha\beta\sigma\lambda}\mathring{R}^{\alpha\beta\sigma\lambda})}{\sqrt{-g}\delta g^{\mu\nu}} = -\frac{1}{2}g_{\mu\nu}\mathring{R}_{\alpha\beta\sigma\lambda}\mathring{R}^{\alpha\beta\sigma\lambda} + 2\mathring{R}_\mu{}^{\alpha\beta\sigma}\mathring{R}_{\nu\alpha\beta\sigma} + 4\mathring{\Box}\mathring{R}_{\mu\nu}$$
$$- 2\mathring{\nabla}_\nu\mathring{\nabla}_\mu\mathring{R} + 4\mathring{R}^{\lambda\sigma}\mathring{R}_{\lambda\mu\sigma\nu} - 4\mathring{R}_{\mu\lambda}\mathring{R}_\nu{}^\lambda = 0. \tag{5.16}$$

These are again seen to be fourth order differential equations. Notice that the sum of these terms with the right coefficients leads to the higher-order derivative terms to cancel out. Namely the combination

$$\mathcal{L}_{GB} = \mathring{R}^2 - 4\mathring{R}_{\mu\nu}R^{\mu\nu} + \mathring{R}_{\mu\nu\sigma\lambda}\mathring{R}^{\mu\nu\sigma\lambda}, \tag{5.17}$$

called the Gauss-Bonnet term. The variation of this term can be written in a compact form using the Weyl tensor (0.8). This is identically zero and it is called the Bach-Lanczos identity [57]

$$\frac{\delta\mathcal{L}_{GB}}{\delta g^{\mu\nu}} = W_\mu{}^{\alpha\beta\sigma}W_{\nu\alpha\beta\sigma} - \frac{1}{4}g_{\mu\nu}W_{\alpha\beta\sigma\lambda}W^{\alpha\beta\sigma\lambda} = 0, \tag{5.18}$$

for a simple proof see e.g. [58, Appendix]. The combination (5.17) does not contribute to the equations of motion in 4 dimensions. The term (5.17) is a topological invariant (Euler characteristic) of the manifold. The generalized Gauss-Bonnet theorem in 4 dimensions states that

$$\frac{1}{32\pi^2}\int_M d^4x\sqrt{-g}\left[\mathring{R}^2 - 4\mathring{R}_{\mu\nu}\mathring{R}^{\mu\nu} + \mathring{R}_{\mu\nu\sigma\lambda}\mathring{R}^{\mu\nu\sigma\lambda}\right] = \chi(M), \tag{5.19}$$

where $\chi(M)$ is a topological invariant of the manifold M. This implies that only two of the three different higher-order terms are independent. We can solve out one of the three terms (choosing the most complicated one including the Riemann tensor) and write the general action quadratic in the curvature as (ignoring the constant term)

$$S = \int d^4x\sqrt{-g}\left(\alpha\mathring{R}_{\mu\nu}\mathring{R}^{\mu\nu} + \beta\mathring{R}^2 + \gamma\mathring{R}\right), \tag{5.20}$$

where α, β and γ are constants. This is exactly the form of the Lagrangian that has been proven to be renormalizable in all orders of perturbation theory [49]. This, however, comes with the cost of massive spin-2 ghost [59][60]. Next we will consider Lagrangians that do not lead to ghosts in the metric formulation and see that this requirement restricts the form of the Lagrangian considerably.

---

[6]This derivation gives rise to boundary terms that can be canceled by adding YHG-terms to the action as done before; these terms are ignored here.

5.2.1   Lovelock gravity

Lovelock constructed the most general Lagrangian in metric formulation, that leads to second order field equations, and do not contain ghosts, in arbitrary number of spacetime dimension [61, 62]. The restrictions that Lovelock imposed were: the general Einstein tensor can only depend on the metric and its first two derivatives and it should be divergence free. It turns out that in four spacetime dimensions the only such tensors are the metric and the Einstein tensor of GR. Thus the Lovelock action is the Einstein–Hilbert action; this is also called the Lovelock's theorem [61, Theorem 5]. The Lovelock Lagrangian in d dimensions is

$$\mathcal{L}^d_{Lovelock} = \sum_{n=0}^{n<d/2} c_n \mathcal{L}_n, \tag{5.21}$$

where c_n are constants with dimensionality $2n - d$ and

$$\mathcal{L}_n = \frac{1}{2^n} \delta^{\alpha_1 \beta_1 \ldots \alpha_n \beta_n}_{\mu_1 \nu_1 \ldots \mu_n \nu_n} \prod_{m=1}^{n} \mathring{R}^{\mu_m \nu_m}{}_{\alpha_m \beta_m}, \tag{5.22}$$

where $\delta^{\ldots}_{\ldots}$ is the generalized Kronecker delta (0.10). Notice that if $n > d/2$ then $\mathcal{L}_n$ will be zero, since there are more indices in the Kronecker delta than there are dimensions. The first few terms of $\mathcal{L}_n$ are

$$\mathcal{L}_0 = \Lambda, \tag{5.23a}$$

$$\mathcal{L}_1 = R, \tag{5.23b}$$

$$\mathcal{L}_2 = \mathring{R}^2 - 4\mathring{R}_{\mu\nu} R^{\mu\nu} + \mathring{R}_{\mu\nu\sigma\lambda} \mathring{R}^{\mu\nu\sigma\lambda}. \tag{5.23c}$$

The term with $n = d/2$ turns out to be a topological invariant also in general and this justifies why it is not included in the sum (5.21).

It turns out that for Lovelock gravity every solution in the metric formulation is also a solution in the Palatini formulation, but not the other way around. Not every solution of the Palatini formulation is a solution of the metric formulation. Thus in the case of Lovelock gravity metric formulation is contained in the Palatini formulation. There are also other forms of higher-order curvature Lagrangians that have this relation between these two formulations [63].

5.3   Quadratic gravity in the Palatini formulation

Let us now consider adding quadratic curvature invariants to the action in the Palatini formulation. In contrast to the metric case there are more different invariants that we can construct out of two Riemann tensors. We can think of the most general

Lagrangian second order in the Riemann tensor, that does not violate parity (i.e. ignoring terms with the Levi–Civita symbol) as [7]

$$S = \int d^4x \sqrt{-g} \left(\text{all contractions of } R^{\sigma}{}_{\rho\mu\nu} R^{\alpha}{}_{\beta\tau\delta} \right). \tag{5.24}$$

There are 16 possible independent contractions due to the Riemann tensor having only one symmetry. These contractions can be written in terms of the different possible Ricci tensors found in (2.23)–(2.25). With these the action can be written as

$$
\begin{aligned}
S = \int d^4x \sqrt{-g} \Big[&\alpha R^2 + \beta_1 R_{\mu\nu} R^{\mu\nu} + \beta_2 R_{\mu\nu} R^{\nu\mu} + \beta_3 R_{\mu\nu} \hat{R}^{\mu\nu} + \beta_4 R_{\mu\nu} \hat{R}^{\nu\mu} + \\
&\beta_5 \hat{R}_{\mu\nu} \hat{R}^{\mu\nu} + \beta_6 \hat{R}_{\mu\nu} \hat{R}^{\nu\mu} + \beta_7 \hat{R}_{\mu\nu} R'^{\mu\nu} + \beta_8 R'_{\mu\nu} R'^{\mu\nu} + \beta_9 R_{\mu\nu} R'^{\mu\nu} + \\
&\gamma_1 R_{\mu\nu\sigma\lambda} R^{\mu\nu\sigma\lambda} + \gamma_2 R_{\mu\nu\sigma\lambda} R^{\mu\sigma\nu\lambda} + \gamma_3 R_{\mu\nu\sigma\lambda} R^{\nu\mu\sigma\lambda} + \gamma_4 R_{\mu\nu\sigma\lambda} R^{\nu\sigma\mu\lambda} + \\
&\gamma_5 R_{\mu\nu\sigma\lambda} R^{\sigma\nu\mu\lambda} + \gamma_6 R_{\mu\nu\sigma\lambda} R^{\sigma\lambda\mu\nu} \Big].
\end{aligned}
\tag{5.25}
$$

Derivation of the equations of motion for these terms follows the similar procedure as that we did for the Einstein–Hilbert action in the section 2.2.1. We will derive the equations of motion for one of these terms for clarity and present the rest of the equations of motion for each term in the appendix B.

We will consider the term $R_{\mu\nu} R^{\mu\nu}$. Starting with the metric variation

$$\delta \int d^4x \sqrt{-g} R_{\mu\nu} R^{\mu\nu} = \int d^4x \left[R_{\mu\nu} R^{\mu\nu} \delta\sqrt{-g} + \sqrt{-g} \delta(R_{\mu\nu} R^{\mu\nu}) \right]. \tag{5.26}$$

Taking the variation of $R_{\mu\nu} R^{\mu\nu}$ gives

$$
\begin{aligned}
\delta \left(R_{\mu\nu} R^{\mu\nu} \right) &= \delta \left(g^{\tau\mu} g^{\lambda\nu} R^{\sigma}{}_{\mu\sigma\nu} R^{\rho}{}_{\tau\rho\lambda} \right) \\
&= (\delta g^{\tau\mu}) g^{\lambda\nu} R_{\mu\nu} R_{\tau\lambda} + g^{\tau\mu} (\delta g^{\lambda\nu}) R_{\mu\nu} R_{\tau\lambda} \\
&= \left(R_{\mu}{}^{\lambda} R_{\nu\lambda} + R^{\lambda}{}_{\mu} R_{\lambda\nu} \right) \delta g^{\mu\nu}
\end{aligned}
\tag{5.27}
$$

With the previous identity and (2.7) the equation of motion turn out to be

$$R_{\mu}{}^{\lambda} R_{\nu\lambda} + R^{\lambda}{}_{\mu} R_{\lambda\nu} - \frac{1}{2} g_{\mu\nu} R_{\alpha\beta} R^{\alpha\beta} = 0. \tag{5.28}$$

Next let's look at the variation with respect to the connection. First of all we have

$$
\begin{aligned}
\int d^4x \sqrt{-g} \delta(R_{\mu\nu} R^{\mu\nu}) &= \int d^4x \sqrt{-g} g^{\tau\mu} g^{\lambda\nu} \left[R_{\tau\lambda} \delta R_{\mu\nu} + R_{\mu\nu} \delta R_{\tau\lambda} \right] \\
&= \int d^4x \sqrt{-g} 2 R^{\mu\nu} \delta R_{\mu\nu}.
\end{aligned}
\tag{5.29}
$$

Then using the Palatini identity (2.30) this becomes

$$\int d^4x \sqrt{-g} 2 R^{\mu\nu} \left[\nabla_{\sigma} \delta\Gamma^{\sigma}_{\nu\mu} - \nabla_{\nu} \delta\Gamma^{\sigma}_{\sigma\mu} + T^{\sigma}{}_{\lambda\nu} \delta\Gamma^{\lambda}_{\sigma\mu} \right]. \tag{5.30}$$

---

[7]There might be some redundancy in these terms, if there exists some equivalent more general version of the Gauss-Bonnet theorem in metric affine space.

Like previously the first and the second term inside the brackets have to be simplified further. This computation is very similar to the one we performed in section 2.2.1. First we expand out the first term in (5.30)

$$
\begin{aligned}
2R^{\mu\nu}\nabla_\sigma\delta\Gamma^\sigma_{\nu\mu} = {}& 2\sqrt{-g}\nabla_\sigma(R^{\mu\nu}\delta\Gamma^\sigma_{\nu\mu}) + 2R^{\mu\nu}\nabla_\sigma\sqrt{-g}\delta\Gamma^\sigma_{\nu\mu} \\
& - 2\nabla_\sigma(\sqrt{-g}R^{\mu\nu})\delta\Gamma^\sigma_{\nu\mu},
\end{aligned}
\tag{5.31}
$$

using the change of variables $\Gamma^\sigma_{\mu\nu} = \mathring{\Gamma}^\sigma_{\mu\nu} + C^\sigma{}_{\mu\nu}$ this becomes

$$
\begin{aligned}
2R^{\mu\nu}\nabla_\sigma\delta\Gamma^\sigma_{\nu\mu} = {}& 2\sqrt{-g}\mathring{\nabla}_\sigma(R^{\mu\nu}\delta\Gamma^\sigma_{\nu\mu}) + 2\sqrt{-g}C^\lambda{}_{\lambda\sigma}(R^{\mu\nu}\delta\Gamma^\sigma_{\nu\mu}) + \\
& 2R^{\mu\nu}\left[\mathring{\nabla}_\nu\sqrt{-g} - C^\lambda{}_{\sigma\lambda}\right]\delta\Gamma^\sigma_{\nu\mu} - 2\nabla_\sigma(\sqrt{-g}R^{\mu\nu})\delta\Gamma^\sigma_{\nu\mu}.
\end{aligned}
\tag{5.32}
$$

The first term in this expression is now a total divergence and can be turned into a surface term by Stokes' theorem (2.13) and it will vanish. The terms containing the $C^\sigma{}_{\mu\nu}$ tensors can be again written in terms of the torsion tensor. With these this simplifies to

$$
2R^{\mu\nu}\nabla_\sigma\delta\Gamma^\sigma_{\nu\mu} = \left[2\sqrt{-g}R^{\mu\nu}T^\lambda{}_{\lambda\sigma} - 2\nabla_\sigma(\sqrt{-g}R^{\mu\nu})\right]\delta\Gamma^\sigma_{\nu\mu},
\tag{5.33}
$$

up to the surface terms. The second term in (5.30) goes similarly and the equations of motion turn out to be

$$
\begin{aligned}
& 2\nabla_\lambda(\sqrt{-g}R^{\mu\lambda})\delta^\nu_\sigma - 2\nabla_\sigma(\sqrt{-g}R^{\mu\nu}) \\
& + 2\sqrt{-g}\left[R^{\mu\nu}T^\lambda{}_{\lambda\sigma} - R^{\mu\rho}T^\lambda{}_{\lambda\rho} + R^{\mu\lambda}T^\nu{}_{\sigma\lambda}\right] = 0.
\end{aligned}
\tag{5.34}
$$

The non-metricity nature of these equations becomes more transparent if we write it in terms of the non-metricity tensor, defined by

$$
Q_{\sigma\mu\nu} \equiv \nabla_\sigma g_{\mu\nu}.
\tag{5.35}
$$

Note that it is symmetric in the last two indices. It is important to also note that $\nabla_\sigma g^{\mu\nu} \neq Q_\sigma{}^{\mu\nu}$. A quick computation will give us the right relation,

$$
\nabla_\sigma g^{\mu\nu} = \nabla_\sigma(g^{\mu\alpha}g^{\nu\beta}g_{\alpha\beta}) = Q_\sigma{}^{\mu\nu} + 2\nabla_\sigma g^{\mu\nu},
\tag{5.36}
$$

and thus

$$
\nabla_\sigma g^{\mu\nu} = -Q_\sigma{}^{\mu\nu}.
\tag{5.37}
$$

With the definition of non-metricity and torsion it is straightforward to verify that a general connection can be written as [8]

$$
\Gamma^\sigma_{\mu\nu} = \mathring{\Gamma}^\sigma_{\mu\nu} + \frac{1}{2}\left[-Q_\mu{}^\sigma{}_\nu + Q^\sigma{}_{\nu\mu} - Q_{\nu\mu}{}^\sigma\right] + \frac{1}{2}\left[T^\sigma{}_{\mu\nu} - T_{\mu\nu}{}^\sigma + T_\nu{}^\sigma{}_\mu\right].
\tag{5.38}
$$

---

[8]Plug in the definition of non-metricity (5.35) into (5.38) and write out the covariant derivative of the metric. With a few steps of algebra the left and right hand side of the equation will match.

Using this we can write $\nabla_\mu \sqrt{-g}$ in terms of the non-metricity, since

$$\nabla_\mu \sqrt{-g} = \mathring{\nabla}_\mu \sqrt{-g} - C^\lambda{}_{\mu\lambda} \sqrt{-g}. \tag{5.39}$$

Taking the trace of (5.38) we get

$$\begin{aligned}
C^\lambda{}_{\mu\lambda} &= \frac{1}{2} \left[-Q_\mu{}^\lambda{}_\lambda + Q^\lambda{}_{\lambda\mu} - Q_{\lambda\mu}{}^\lambda \right] + \frac{1}{2} \left[T^\lambda{}_{\mu\lambda} - T_{\mu\lambda}{}^\lambda + T_\lambda{}^\lambda{}_\mu \right] \\
&= -\frac{1}{2} Q_\mu{}^\lambda{}_\lambda,
\end{aligned} \tag{5.40}$$

where the torsion tensors vanish due to them being antisymmetric in the last two indices and two of the non-metricity terms vanish due to the symmetry of the last two indices. Thus we have

$$\nabla_\mu \sqrt{-g} = \frac{1}{2} Q_\mu{}^\lambda{}_\lambda \sqrt{-g}. \tag{5.41}$$

Using this we can write the equations of motion (5.34) as

$$\begin{aligned}
2R^{\mu\lambda} &\left[\frac{1}{2} Q_\lambda{}^\rho{}_\rho \delta^\nu_\sigma - T^\rho{}_{\rho\lambda} \delta^\nu_\sigma + T^\nu{}_{\sigma\lambda} \right] + 2R^{\mu\nu} \left[-\frac{1}{2} Q_\sigma{}^\rho{}_\rho + T^\lambda{}_{\lambda\sigma} \right] \\
&- Q_\lambda{}^{\mu\lambda} \delta^\nu_\sigma + Q_\sigma{}^{\mu\nu} + 2\delta^\nu_\sigma \nabla_\lambda R^{\mu\lambda} - 2\nabla_\sigma R^{\mu\nu} = 0.
\end{aligned} \tag{5.42}$$

The non-metricity and torsion dependence of the last two terms containing covariant derivatives of the Ricci tensor can also be written out using (5.38). This form for the equation is useful when trying to find solutions in the general case. However, as we can see even for this one quadratic term $R_{\mu\nu} R^{\mu\nu}$ the equations of motion turn out to be very complicated. Furthermore, the fact that some of the extra degrees of freedom coming from the higher-order terms turn out to be ghost degrees of freedom complicate the picture [64, 65, 55]. To be able to say something about higher-order terms and Higgs inflation we will restrict our action such that no new degrees of freedom appear, and thus avoid any inconsistencies. In the next chapter we shall see how this is accomplished.

Chapter 6

Higher-order gravity and Higgs inflation

In this chapter we present a new result. We analyze Higgs inflation with higher-order curvature terms in the Palatini formulation. We have to make several simplifying assumptions to make the calculations analytically tractable. The first simplification that we make is to assume vanishing torsion, i.e. our general connection is assumed to be symmetric $\Gamma^\sigma_{\nu\mu} = \Gamma^\sigma_{\mu\nu}$. We will only add terms constructed from the Ricci tensor and additionally we will consider an action invariant under projective transformations, which does not introduce extra gravitational degrees of freedom, that would complicate the picture. The action that we will consider in the Jordan frame reads

$$
S = \int d^4x \sqrt{-g} \left[\frac{1}{2} \left\{ \left(1 + \xi h^2\right) R + \beta R^2 + \alpha R_{(\mu\nu)} R^{(\mu\nu)} \right\} \right.
$$
$$
\left. - \frac{1}{2} g^{\mu\nu} \partial_\mu h \partial_\nu h - V(h) \right].
\tag{6.1}
$$

Like in the case of pure Higgs inflation one could work with this action and compute the inflationary observables. However we can again simplify our life by transforming our action into an Einstein–Hilbert like action; we will again call this the Einstein frame. However, this is now a more involved task due to the higher-order curvature terms. We will also present an alternative method of computing the slow-roll equations without going into the Einstein frame.

First we will carry out the computation in the Einstein frame. We start by introducing a general procedure how to bring an projective invariant Ricci based theory, like we have, into the Einstein frame. We then use this method to compute the slow-roll equations and parameters with our action (6.1). Next we will derive the slow-roll equations without going to the Einstein frame and see that our results agree. Finally we will show how the inflationary observables are modified by the higher-order curvature terms.

6.1   Finding the Einstein frame representation

Here we will outline the general procedure how to transform an projective invariant Ricci based theory to the Einstein frame. This general procedure was also described in [66]. We will then use this procedure in our special case to obtain our new result.

We write our general action as

$$S = \frac{1}{2} \int d^4x \sqrt{-g} \left[F \left(g_{\mu\nu}, R_{(\mu\nu)}, \psi_i, \partial_\mu \psi_i \right) + \mathcal{L}_m \left(g_{\mu\nu}, \psi_i, \partial_\mu \psi_i \right) \right], \qquad (6.2)$$

where F is a general function of the symmetric part of the Ricci tensor $R_{\mu\nu}$, the metric $g_{\mu\nu}$ and matter fields ψ_i; and $\mathcal{L}_m$ is the matter part of the action. We can perform a Legendre transformation so that the action becomes

$$\begin{aligned} S = \frac{1}{2} \int d^4x \sqrt{-g} \Bigg[& F \left(g_{\mu\nu}, \Sigma_{\mu\nu}, \psi_i, \partial_\mu \psi_i \right) + \frac{\partial F}{\partial \Sigma_{\mu\nu}} \left(R_{(\mu\nu)} - \Sigma_{\mu\nu} \right) \\ & + \mathcal{L}_m \left[g_{\mu\nu}, \psi_i, \partial_\mu \psi_i \right] \Bigg], \end{aligned} \qquad (6.3)$$

where $\Sigma_{\mu\nu}$ is an auxillary field, which is symmetric by definition. Taking the variation of the action (6.3) with respect to $\Sigma_{\mu\nu}$ gives the constraint $\Sigma_{\mu\nu} = R_{(\mu\nu)}$ and thus the actions (6.2) and (6.3) describe the same theory. The Legendre transformed action is now linear in the Ricci tensor and we are one step closer to finding the Einstein frame. We next make a field redefinition

$$\sqrt{-q} q^{\mu\nu} = \sqrt{-g} \frac{\partial F}{\partial \Sigma_{\mu\nu}}, \qquad (6.4)$$

where $q \equiv \det q_{\mu\nu}$ and $q_{\mu\nu}$ is the inverse of $q^{\mu\nu}$, which means that $q^{\mu\lambda} q_{\nu\lambda} = \delta^\mu_\nu$. Note that the new variable is also symmetric $q_{\mu\nu} = q_{\nu\mu}$. From this redefinition we can solve the auxillary field $\Sigma_{\mu\nu}$ in terms of the metric $g_{\mu\nu}$, the matter fields ψ_i and the new variable $q_{\mu\nu}$, i.e. $\Sigma_{\mu\nu} = \Sigma_{\mu\nu}(g_{\mu\nu}, q_{\mu\nu}, \psi_i)$. This enables us to write the action as

$$\begin{aligned} S = \frac{1}{2} \int d^4x \Bigg\{ & \sqrt{-q} q^{\mu\nu} R_{\mu\nu} \\ & - \sqrt{-g} \left[\frac{\partial F}{\partial \Sigma_{\mu\nu}} \Sigma_{\mu\nu}(q_{\mu\nu}, g_{\mu\nu}, \psi_i) - F(q_{\mu\nu}, g_{\mu\nu}, \psi_i) + \mathcal{L}_m(g_{\mu\nu}, \psi_i) \right] \Bigg\}. \end{aligned}$$
$$(6.5)$$

From this action it is apparent that the metric $g_{\mu\nu}$ does not have a kinetic term and thus its equations of motion are algebraic equations. We can thus solve the original metric $g_{\mu\nu}$ in terms of the new field $q_{\mu\nu}$ and the matter fields ψ_i, that is $g_{\mu\nu} = g_{\mu\nu}(q_{\mu\nu}, \psi_i)$. Plugging this back into the (6.5) gives us an action that depends only on the independent connection $\Gamma^\sigma_{\mu\nu}$, the new field $q_{\mu\nu}$ and the matter fields ψ_i

$$S = \frac{1}{2} \int d^4x \left\{ \sqrt{-q} q^{\mu\nu} R_{\mu\nu} - \sqrt{-g} \left[\frac{\partial F}{\partial \Sigma_{\mu\nu}} \Sigma_{\mu\nu}(q_{\mu\nu}, \psi_i) - \frac{1}{2} F(q_{\mu\nu}, \psi_i) - \mathcal{L}_m(q, \psi) \right] \right\},$$
$$(6.6)$$

where $\sqrt{-g}$ is also solved in terms of $q_{\mu\nu}$ and ψ_i. Thus we have managed to transform our action into the Einstein frame where the gravitational sector is the Einstein–Hilbert action for the metric $q_{\mu\nu}$.

In general, if we do not assume a symmetric Ricci tensor the new metric $q_{\mu\nu}$ is not symmetric. Even if a symmetric connection is assumed the Ricci tensor can

have an antisymmetric part: $R_{[\mu\nu]} = -\nabla_{[\nu}Q_{\mu]}$. and thus from the field redefinition (6.4) it is clear that also $q^{\mu\nu}$ has an antisymmetric part. The non-symmetric part will then contain new gravitational degrees of freedom and in general these can lead to instabilities in the theory [67][68]. In fact in the non-symmetric case the action can be seen to be equivalent to a non-symmetric gravity theory, which has been shown to contain instabilities. However, it is possible to construct a non-projective invariant theory in Ricci based gravity without instabilities by requiring that the torsion will vanish by means of Lagrange multipliers. Then the resulting new gravitational degrees of freedom from the antisymmetric part will result in one massive vector field [55]. In fact, in a cosmological setting with the FLRW universe the Ricci tensor is restricted to be symmetric since the only available tensors are the projection tensor for the spacelike slices $h_{\mu\nu}$ and the timelike vector u_μ. Thus the background evolution will be the same with a general Ricci tensor, but the perturbations will include these extra degrees of freedom. Considering these extra degrees of freedom are out of the scope of this thesis.

If the action is projective invariant there are no new gravitational degrees of freedom and $q^{\mu\nu}$ is symmetric. The action (6.6) is then the Einstein–Hilbert action for the metric $q_{\mu\nu}$ with modified matter sector. The connection is thus the Levi–Civita connection for the metric $q_{\mu\nu}$.

In what follows we shall carry out this procedure in the case of Higgs inflation with higher-order curvature terms.

6.2 Einstein frame

The general procedure works for non-minimally coupled actions. However, we will simplify our action (6.1) by removing the non-minimal coupling with a conformal transformation like previously

$$g^{\mu\nu} \to \Omega^2(x)g^{\mu\nu}, \quad \Omega^2 = 1 + \xi h^2. \tag{6.7}$$

Since the higher-order terms are invariant under conformal transformations in the Palatini formulation, as we noted in section 4.1, only the matter sector is modified. The action becomes

$$S = \int d^4x \sqrt{-g} \left[\frac{1}{2} \left\{ R + \beta R^2 + \alpha R_{\mu\nu}R^{\mu\nu} \right\} - \frac{1}{2}\Omega^{-2}g^{\mu\nu}\partial_\mu h \partial_\nu h - \Omega^{-4}V(h) \right]. \tag{6.8}$$

The kinetic term is brought to the canonical form with the same field redefinition as before (4.21). And the potential $U(\chi)$ is defined as (4.23). With these the action simplifies to

$$S = \int d^4x \sqrt{-g} \left[\frac{1}{2} \left\{ R + \beta R^2 + \alpha R_{\mu\nu}R^{\mu\nu} \right\} - \frac{1}{2}g^{\mu\nu}\partial_\mu\chi\partial_\nu\chi - U(\chi) \right]. \tag{6.9}$$

We will call the action in this form the higher-order gravity frame. We will also later derive the slow-roll equations in this frame to compare with the Einstein frame

results. First, let us carry out the procedure outlined above and transform (6.9) into the Einstein frame.

With our action (6.9) the function F is

$$F(g_{\mu\nu}, R_{\mu\nu}) = R + \beta R^2 + \alpha R_{\mu\nu}R^{\mu\nu}, \tag{6.10}$$

and the matter sector is

$$\mathcal{L}_m(g_{\mu\nu}, \chi) = -\frac{1}{2}g^{\mu\nu}\partial_\mu\chi\partial_\nu\chi - U(\chi). \tag{6.11}$$

To bring our action explicitly to the form of (6.6) we start by solving the auxillary field $\Sigma_{\mu\nu}$ in terms of $q^{\mu\nu}$ and $g^{\mu\nu}$. We will write out all the metrics explicitly when performing contractions to see more easily the metric dependence and to make it clear which metric is used when making the contractions. Writing out (6.4) with (6.10) we have

$$\sqrt{-q}q^{\mu\nu} = \sqrt{-g}\left[(1 + 2\beta g^{\sigma\lambda}\Sigma_{\sigma\lambda})g^{\mu\nu} + 2\alpha g^{\sigma\mu}g^{\lambda\nu}\Sigma_{\sigma\lambda}\right]. \tag{6.12}$$

Taking the trace of this we have

$$\sqrt{-q}q^{\mu\nu}g_{\mu\nu} = \sqrt{-g}\left[4(1 + 2\beta g^{\sigma\lambda}\Sigma_{\sigma\lambda}) + 2\alpha g^{\sigma\lambda}\Sigma_{\sigma\lambda}\right], \tag{6.13}$$

from which we can solve

$$g^{\sigma\lambda}\Sigma_{\sigma\lambda} = \frac{\sqrt{-q}q^{\mu\nu}g_{\mu\nu}}{\sqrt{-g}(8\beta + 2\alpha)} - \frac{4}{8\beta + 2\alpha}. \tag{6.14}$$

Substituting this back into (6.12) and solving for $\Sigma_{\mu\nu}$ we get

$$\Sigma_{\mu\nu} = \frac{1}{2\alpha}\frac{\sqrt{-q}}{\sqrt{-g}}q^{\sigma\lambda}g_{\sigma\mu}g_{\lambda\nu} - \frac{1}{2\alpha + 8\beta}\left[1 + \frac{\beta\sqrt{-q}}{\alpha\sqrt{-g}}q^{\sigma\lambda}g_{\sigma\lambda}\right]g_{\mu\nu}. \tag{6.15}$$

Next we need to substitute this back to the Legendre transformed action (6.3) and derive the equations of motion for the original metric $g_{\mu\nu}$. Computing the variation is straightforward but lengthy, see appendix C.1 for a few more details, here we will just present the result

$$
\begin{aligned}
\frac{1}{\sqrt{-g}}\frac{\delta S}{\delta g^{\mu\nu}} = &-\frac{1}{4(\alpha + 4\beta)}\frac{\sqrt{-q}}{\sqrt{-g}}q^{\sigma\lambda}g_{\sigma\mu}g_{\lambda\nu} \\
&+ \frac{1}{4\alpha}\frac{q}{g}\left[q^{\sigma\lambda}q^{\rho\delta}g_{\lambda\delta}g_{\rho\nu}g_{\sigma\mu} - \frac{\beta}{(\alpha + 4\beta)}q^{\delta\rho}g_{\delta\rho}q^{\sigma\lambda}g_{\sigma\mu}g_{\lambda\nu}\right] \\
&+ \frac{1}{2}g_{\mu\nu}\left[\frac{1}{\alpha + 4\beta}\left(\frac{\beta}{8\alpha}\frac{q}{g}q^{\lambda\sigma}g_{\lambda\sigma}q^{\rho\delta}g_{\rho\delta}\right) - \frac{q}{g}\frac{1}{8\alpha}q^{\lambda\sigma}q^{\delta\rho}g_{\lambda\delta}g_{\sigma\rho}\right] \\
&- \frac{1}{2}\partial_\mu\chi\partial_\nu\chi + \frac{1}{2}g_{\mu\nu}\left[U(\chi) + \frac{1}{2}g^{\lambda\sigma}\partial_\lambda\chi\partial_\sigma\chi\right] = 0.
\end{aligned}
\tag{6.16}
$$

We now need to solve the original metric $g_{\mu\nu}$ in terms of the new metric $q_{\mu\nu}$ and the scalar field χ. This turns out to be the most difficult part. We will solve this by introducing the following ansatz

$$g_{\mu\nu} = \gamma_1(\chi, X_q)q_{\mu\nu} + \gamma_2(\chi, X_q)\partial_\mu\chi\partial_\nu\chi, \tag{6.17}$$

where $X_q = q^{\mu\nu}\partial_\mu\chi\partial_\nu\chi$. This is in the form of a disformal transformation. We are thus effectively finding the disformal transformation which brings the action to the Einstein frame. Let us briefly discuss these kind of metric transformations.

6.2.1 Disformal transformations

Another useful type of metric transformations that is able to bring more complicated actions into the Einstein frame, is disformal transformations. These are of the form

$$q_{\mu\nu} = \Gamma_1(\chi, X_g)g_{\mu\nu} + \Gamma_2(\chi, X_g)\partial_\mu\chi\partial_\nu\chi, \tag{6.18}$$

where $X_g = g^{\mu\nu}\partial_\mu\chi\partial_\nu\chi$ and Γ_1, Γ_2 are called the conformal and disformal factor respectively. We write the inverse of this transformation as

$$g_{\mu\nu} = \gamma_1(\chi, X_q)q_{\mu\nu} + \gamma_2(\chi, X_q)\partial_\mu\chi\partial_\nu\chi. \tag{6.19}$$

Notice that both of the factors are allowed to depend on the kinetic term of the scalar field and thus these factors depend on the metric $g_{\mu\nu}$ and in the case of the inverse transformation they depend on the new metric $q_{\mu\nu}$. The factors between the transformation and its inverse have the following relations

$$\Gamma_1(\chi, X_g) = \frac{1}{\gamma_1(\chi, X_q(X_g))}, \quad \Gamma_2(\chi, X_g) = -\frac{\gamma_2(\chi, X_q(X_g))}{\gamma_1(\chi, X_q(X_g))}, \tag{6.20}$$

where the X_q has to be solved in terms of X_g. We will get a more explicit form for $X_q(X_g)$ later. The factors Γ_1, Γ_2 are restricted by requiring that the new metric describes a well defined spacetime. For that we require the following four conditions:

1. The transformation should be invertible, so that there exists an inverse metric $q^{\mu\nu}$.

2. The transformation does not change the Lorentzian signature of the metric.

3. Causal trajectories have to remain causal.

4. The transformation has to be invertible.

Sometimes in the literature also another condition is listed, which is that the volume element has to be non-zero. This is however redundant since this is already guaranteed by the first condition listed. These requirements assure that the transformation is a map between two pseudo-Riemannian spaces and the two representations are physically equivalent, for investigations of this see e.g. [69, 70].

It is easier to compute the constraints in a cosmological setting and since we are interested in analyzing inflation this is enough for our purposes. Let's look at how these requirements restrict the two factors. Looking at the invariant line element

$$d\hat{s}^2 = \Gamma_1 g_{\mu\nu}dx^\mu dx^\nu + \Gamma_2\partial_\mu\chi\partial_\nu\chi dx^\mu dx^\nu = \Gamma_1 ds^2 + \Gamma_2(\partial_\mu\chi dx^\mu)^2, \tag{6.21}$$

we see that the light-cones get modified and are either stretched or squeezed depending on the sing of Γ_2. To ensure that causal trajectories, $ds^2 < 0$, are also causal with the transformed metric we must have $d\hat{s}^2 < 0$, i.e. $\Gamma_2 < 0$ everywhere.

To maintain the Lorentzian signature of the transformed metric, i.e. $q_{00} < 0$ with $q_{0i} = 0$, we must have

$$q_{00} = \Gamma_1 g_{00} + \Gamma_2 \partial_0 \chi \partial_0 \chi < 0. \tag{6.22}$$

Multiplying both sides with g^{00}, which is negative, we get

$$\Gamma_1 + \Gamma_2 g^{00} \partial_0 \chi \partial_0 \chi > 0. \tag{6.23}$$

In a FLRW model $\partial_\mu \chi$ is time-like; thus, taking a frame where $\partial_\mu \chi = (\partial_0 \chi, 0, 0, 0)$ we can write $X_g = g^{00} \partial_0 \chi \partial_0 \chi$. So we have the frame independent requirement

$$\Gamma_1 + \Gamma_2 X_g > 0. \tag{6.24}$$

The inverse of the transformed metric $q_{\mu\nu}$ can be obtained from $q_{\lambda\nu} q^{\mu\lambda} = \delta^\mu_\nu$. We introduce an ansatz for the inverse metric

$$q^{\mu\nu} = \Gamma_3(\chi, X_g) g^{\mu\nu} + \Gamma_4(\chi, X_g) g^{\mu\lambda} g^{\nu\sigma} \partial_\lambda \chi \partial_\sigma \chi \tag{6.25}$$

Substituting this and (6.18) into the identity $\delta^\mu_\nu = q^{\mu\lambda} q_{\nu\lambda}$ we get

$$\begin{aligned}
\delta^\mu_\nu &= \left(\Gamma_3 g^{\mu\lambda} + \Gamma_4 g^{\mu\rho} q^{\lambda\tau} \partial_\rho \chi \partial_\tau \chi \right) \left(\Gamma_1 g_{\mu\nu} + \Gamma_2 \partial_\mu \chi \partial_\nu \chi \right) \\
&= \Gamma_1 \Gamma_3 \delta^\mu_\nu + \left(\Gamma_2 \Gamma_4 X_g + \Gamma_1 \Gamma_4 + \Gamma_2 \Gamma_3 \right) g^{\mu\lambda} \partial_\lambda \chi \partial_\nu \chi.
\end{aligned} \tag{6.26}$$

Requiring for the coefficients to match on the left and right side of the equation we get

$$\begin{cases} \Gamma_1 \Gamma_3 = 1 \\ \Gamma_2 \Gamma_4 X_g + \Gamma_1 \Gamma_4 + \Gamma_2 \Gamma_3 = 0. \end{cases} \tag{6.27}$$

Solving for Γ_3 and Γ_4 we get

$$\begin{cases} \Gamma_3 = \dfrac{1}{\Gamma_1} \\ \Gamma_4 = \dfrac{-\Gamma_2}{\Gamma_1(\Gamma_1 + X_g \Gamma_2)}. \end{cases} \tag{6.28}$$

The inverse metric is thus

$$q^{\mu\nu} = \frac{1}{\Gamma_1} g^{\mu\nu} - \frac{\Gamma_2}{\Gamma_1(\Gamma_1 + X_g \Gamma_2)} g^{\lambda\mu} g^{\sigma\nu} \partial_\lambda \chi \partial_\sigma \chi, \tag{6.29}$$

For this to be non-singular we must require $\Gamma_1 > 0$ and $\Gamma_1 + X_q \Gamma_2 > 0$ everywhere. These two conditions are enough to assure the existence of the inverse metric and that the Lorentz signature is unchanged. We can now relate X_q with X_g by multiplying (6.29) with $\partial_\mu \chi \partial_\nu \chi$ which yields

$$X_q = \frac{X_g}{\Gamma_1 + X_g \Gamma_2}. \tag{6.30}$$

The existence of the inverse transformation (6.29) is guaranteed by the condition of non vanishing Jacobian determinant, which yields the condition [71]

$$\Gamma_1 \left(\Gamma_1 + \frac{1}{2} X_g \frac{\partial \Gamma_1}{\partial X_g} + \frac{1}{2} X_g^2 \frac{\partial \Gamma_2}{\partial X_g} \right) \neq 0. \tag{6.31}$$

This constraint also guarantees that the equation (6.30) is solvable. Next let's look at how the determinant of the metric transforms. We begin by taking (6.18) and contracting both sides with $g^{\nu\sigma}$ and taking the determinant

$$\det(q_{\mu\nu} g^{\nu\sigma}) = \det \left[\Gamma_1 \delta_\mu^\sigma + \Gamma_2 g^{\nu\sigma} \partial_\mu \chi \partial_\nu \chi \right]. \tag{6.32}$$

Thinking in terms of matrices the left-hand side is a determinant of the product of two square matrices which is equal to the product of their determinants. Additionally $\det(M^{-1}) = (\det M)^{-1}$, so we have

$$q g^{-1} = \det \left[\Gamma_1 \delta_\sigma^\mu + \Gamma_2 g^{\nu\sigma} \partial_\mu \chi \partial_\nu \chi \right]. \tag{6.33}$$

For the right hand side we write the determinant in terms of the completely anti-symmetric Levi–Civita symbol $\epsilon^{\mu\nu\sigma\lambda}$ as

$$\begin{aligned}
q g^{-1} = \epsilon^{\mu\nu\sigma\lambda} \big[& \left(\Gamma_1 \delta_\mu^0 + \Gamma_2 \partial_\mu \chi \partial_\rho \chi q^{0\rho} \right) \\
& \times \left(\Gamma_1 \delta_\nu^1 + \Gamma_2 \partial_\nu \chi \partial_\rho \chi g^{1\rho} \right) \\
& \times \left(\Gamma_1 \delta_\sigma^2 + \Gamma_2 \partial_\sigma \chi \partial_\rho \chi g^{2\rho} \right) \\
& \times \left(\Gamma_1 \delta_\lambda^3 + \Gamma_2 \partial_\lambda \chi \partial_\rho \chi g^{3\rho} \right) \big].
\end{aligned} \tag{6.34}$$

Since the left hand side is a ratio of two scalar densities with the same weight it is a scalar and so we can pick any convenient frame to evaluate this expression. Thus, again picking a frame where $\partial_\mu \chi = (\partial_0 \chi, 0, 0, 0)$ makes all except few terms vanish giving us

$$\begin{aligned}
q g^{-1} &= \Gamma_1^4 \epsilon^{\mu\nu\sigma\lambda} \delta_\mu^0 \delta_\nu^1 \delta_\sigma^2 \delta_\lambda^3 + \Gamma_1^3 \Gamma_2 \partial_\mu \chi \partial_\rho \chi g^{0\rho} \delta_\nu^1 \delta_\sigma^2 \delta_\lambda^3 \epsilon^{\mu\nu\sigma\lambda} \\
&= \Gamma_1^4 + \Gamma_1^3 \Gamma_2 \partial_0 \chi \partial_0 \chi g^{00} \\
&= \Gamma_1^3 \left(\Gamma_1 + X_g \Gamma_2 \right).
\end{aligned} \tag{6.35}$$

Finally the determinant of $q_{\mu\nu}$ in terms of $g_{\mu\nu}$ and χ becomes

$$q = g \Gamma_1^3 \left(\Gamma_1 + X_g \Gamma_2 \right). \tag{6.36}$$

Thus, the volume element transforms as

$$d^4 x \sqrt{-q} = d^4 x \sqrt{-g} \sqrt{\Gamma_1^3 (\Gamma_1 + X_g \Gamma_2)}, \tag{6.37}$$

which is already guaranteed to be non-zero by the previously found constraints. The invariance of inflationary observables under disformal transformations has been investigated in the literature [69, 72, 73, 74, 75].

For our computation we need to perform the inverse transformation (6.19) and thus we need relations for the inverse metric $g^{\mu\nu}$ and the metric determinant in

terms of the factors γ_1, γ_2. We see that the computations can be performed with the exactly same steps as we did above and so we have

$$g^{\mu\nu} = \frac{1}{\gamma_1} g^{\mu\nu} - \frac{\gamma_2}{\gamma_1(\gamma_1 + X_q\gamma_2)} q^{\lambda\mu} q^{\sigma\nu} \partial_\lambda\chi\partial_\sigma\chi, \tag{6.38}$$

for the inverse metric $g^{\mu\nu}$ and

$$g = q\gamma_1^3 \left(\gamma_1 + X_q\gamma_2\right), \tag{6.39}$$

for the metric determinant. Similarly X_g can be written in terms of X_q as

$$X_g = \frac{X_q}{\gamma_1 + X_q\gamma_2}. \tag{6.40}$$

Since we will be solving only the form of the inverse transformation it is easier to check the four conditions that we require in terms of the factors γ_1 and γ_2. Using the relations between the transformation and its inverse (6.20) and how the kinetic terms are related (6.40) the requirements for the existence of the inverse metric become

$$\frac{1}{\gamma_1} > 0 \iff \gamma_1 > 0, \quad \frac{1}{\gamma_1 + X_q\gamma_2} > 0 \iff \gamma_1 + X_q\gamma_2 > 0. \tag{6.41}$$

The requirement for causal trajectories to remain causal becomes

$$-\frac{\gamma_2}{\gamma_1} < 0 \iff \gamma_2 > 0. \tag{6.42}$$

Finally the invertibility condition becomes

$$\gamma_1 \left(\gamma_1 + \frac{1}{2}X_q\frac{\partial\gamma_1}{\partial X_q} + \frac{1}{2}X_q^2\frac{\partial\gamma_2}{\partial X_q}\right) \neq 0. \tag{6.43}$$

Now we have all the necessary tools to carry out the computation.

6.2.2 Solving the metric

In the previous section we found the inverse (6.38) for the ansatz of the metric $g_{\mu\nu}$ (6.17), and additionally the metric determinant g in terms of the new metric $q_{\mu\nu}$ and the scalar χ (6.39).

Now we can substitute the ansatz (6.17),(6.29) and the determinant (6.39) to the constraint equation for the metric $g_{\mu\nu}$ (6.16). This will give us algebraic equations for the coefficients γ_1 and γ_2. After doing the substitutions and requiring that the coefficients of $q_{\mu\nu}$ and $\partial_\mu\chi\partial_\nu\chi$ vanish identically we get the system of equations:

$$\frac{1}{16\alpha(\alpha + 4\beta)\gamma_1^2(\gamma_1 + X_q\gamma_2)} \Bigg[-4\alpha\gamma_1^4 + 2X_q(\alpha + 2\beta)\gamma_1\gamma_2$$
$$+ X_q^2(\alpha + 3\beta)\gamma_2^2 - 4X_q\alpha\gamma_1^3(\alpha + 4\beta + \gamma_2) + 4\alpha\gamma_1\sqrt{\gamma_1^3(\gamma_1 + X_q\gamma_2)} \Bigg] \tag{6.44}$$
$$-\frac{1}{2}\gamma_1 U(\chi) = 0,$$

from the coefficient of $q_{\mu\nu}$ and

$$
\frac{1}{16\alpha(\alpha+4\beta)\gamma_1^3(\gamma_1+X_q\gamma_2)} \left[-4\alpha\gamma_1^4(2\alpha+8\beta-\gamma_2)+8(\alpha+2\beta)\gamma_1^2\gamma_2 \right.
$$
$$
-4X_q\alpha\gamma_1^3(\alpha+4\beta-\gamma_2)\gamma_2 + 2X_q(5\alpha+12\beta)\gamma_1\gamma_2^2 + 3X_q^2(\alpha+3\beta)\gamma_2^3 \tag{6.45}
$$
$$
\left. -\left(8\alpha\gamma_1\gamma_2 + 4X_q\alpha\gamma_2^2\right)\sqrt{\gamma_1^3(\gamma_1+X_q\gamma_2)} \right] + \frac{1}{2}\gamma_2 U(\chi) = 0,
$$

from the coefficient of $\partial_\mu\chi\partial_\nu\chi$. We can now check what happens in the vacuum limit. Taking $X_q \to 0$,$U(\chi) \to 0$ and $\chi \to 0$ the equation (6.44) simplifies to $\gamma_1 - 1 = 0$, the second equation is nonexistent $(0 = 0)$ since there is no kinetic term for the scalar. Substituting $\gamma_1 = 1$ to the modified scalar sector in the action (6.6) gives zero. Thus, we obtain the Einstein equations in vacuum.

In general it is hard to find solutions to this set of algebraic equations. However, in slow-roll approximation the square of the time derivative of the field is small $\dot\chi^2/H^2 \ll 1$. In FLRW universe we have $X_q = -\dot\chi^2$, thus we can treat X_q as a small parameter and find approximate solutions of the form

$$
\gamma_1(\chi, X_q) = a_0 + a_1 X_q + a_2 X_q^2 + \ldots \quad ,
$$
$$
\gamma_2(\chi, X_q) = b_0 + b_1 X_q + b_2 X_q^2 + \ldots \tag{6.46}
$$

We can then find solutions of this form by substituting (6.46) into the equations (6.44) and (6.45) and again expanding in terms of X_q and requiring that each of the coefficients vanish identically. This gives us systems of equations from which we can solve the coefficients $\{a_i, b_i\}$. In the zeroth order we get the system of equations

$$
\frac{a_0\left[1 + 2\alpha U(\chi) + 8\beta U(\chi)\right] - 1}{\alpha + 4\beta} = 0
$$
$$
\frac{-2\alpha(\alpha+4\beta)a_0^2 + (4\beta + \alpha(2+(-2+a_0)a_0))b_0}{4\alpha(\alpha+4\beta)a_0^2} + \frac{1}{2}b_0 U(\chi) = 0, \tag{6.47}
$$

which are easily solved giving

$$
a_0 = \frac{1}{1 + (2\alpha + 8\beta)U(\chi)}
$$
$$
b_0 = \frac{2\alpha}{\left[1 + (2\alpha + 8\beta)U(\chi)\right]\left[1 + (4\alpha + 8\beta)U(\chi)\right]}. \tag{6.48}
$$

These expressions quickly become very cumbersome to write down and are not particularly illuminating. So we will not present any more terms here; a few more coefficients can be found from the appendix C.2. It can be now verified that the requirements that we required for the transformation to be physical are met. To this first order the requirements reduce to $a_0 > 0, b_0 > 0$ and $a_0 + b_0 X_q > 0$.

Now substituting the solution of the form (6.46) with the solved coefficients $\{a_i, b_i\}$ to the action (6.6) and writing this again in series of X_q we get the modified

matter sector

$$\mathcal{L}_m = -\frac{U}{1 + (2\alpha + 8\beta)U} - \frac{\frac{1}{2}q^{\mu\nu}\partial_\mu\chi\partial_\nu\chi}{1 + (2\alpha + 8\beta)U}$$
$$+ \frac{[\alpha^2 U + \beta(1 + 8\beta U) + \alpha(1 + 6\beta U)]}{2\left[1 + (2\alpha + 8\beta)U\right]\left[1 + (4\alpha + 8\beta)U\right]}(q^{\mu\nu}\partial_\mu\chi\partial_\nu\chi)^2 \qquad (6.49)$$
$$+ \ldots$$

As before we can bring the kinetic term to a canonical form by another field redefinition

$$\frac{d\chi}{d\phi} = \sqrt{1 + (2\alpha + 8\beta)U}. \qquad (6.50)$$

The potential of the new field ϕ reads

$$\tilde{U}(\phi) = \frac{U(\chi(\phi))}{1 + (2\alpha + 8\beta)U(\chi(\phi))}. \qquad (6.51)$$

This brings the matter Lagrangian to the form

$$\mathcal{L}_m = -\tilde{U} - \frac{1}{2}q^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$$
$$+ \frac{[\alpha^2 U + \beta(1 + 8\beta U) + \alpha(1 + 6\beta U)]\left[1 + (2\alpha + 8\beta)U\right]}{2\left[1 + (4\alpha + 8\beta)U\right]}(q^{\mu\nu}\partial_\mu\phi\partial_\nu\phi)^2 \qquad (6.52)$$
$$+ \ldots$$

For later convenience we will write this as

$$\mathcal{L}_m = -\tilde{U} - \frac{1}{2}q^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$$
$$+ c_2(U(\phi))(q^{\mu\nu}\partial_\mu\phi\partial_\nu\phi)^2 + \cdots + c_n(U(\phi))(q^{\mu\nu}\partial_\mu\phi\partial_\nu\phi)^n + \ldots \qquad (6.53)$$

At this point we can also note that when taking the limit $\alpha \to 0$ we obtain the result of [76], where non-minimally coupled scalar with the addition of only R^2 term was analyzed. This is, however, with the assumption that all the higher than second order kinetic terms vanish when the limit is taken. This is expected to be the case, but should still be explicitly proven. Easier way of seeing this is to take the limit $\alpha \to 0$ from the start. This simplifies the computations and with some effort the same result is obtained.

When the $\alpha \to 0$ limit is taken we have to require $\beta > 0$ to have the correct sign for the kinetic terms in (6.52) and avoid instabilities. We see that the correct signs for the kinetic terms are preserved in the general case if both $\alpha, \beta > 0$. There might exist more of the parameter space where this is the case, but this needs further investigation. Thus here we require that $\alpha, \beta > 0$.

We are now ready to derive the equations of motion, which now is an easy task since the gravitational sector is nothing but the Einstein–Hilbert action.

6.2.3 Equations of motion

We have managed to bring our action to the form

$$S = \int d^4x \sqrt{-q} \left[\frac{1}{2} q^{\mu\nu} R_{\mu\nu} - \tilde{U} - \frac{1}{2} X_q + c_2(U(\phi)) X_q^2 + \dots \right], \tag{6.54}$$

where $c_n(\phi)$ are functions of the field defined in (6.53). Thus as previously shown the connection is solved to be the Levi–Civita connection of the metric $q^{\mu\nu}$. The equations of motion for the metric thus are

$$G_{\mu\nu}^{(q)} \equiv R_{\mu\nu}^{(q)} - \frac{1}{2} R^{(q)} q_{\mu\nu} =$$
$$- q_{\mu\nu} \left[\tilde{U} + \frac{1}{2} \left(X_q + c_2 X_q^2 + \dots \right) \right] \tag{6.55}$$
$$+ \partial_\mu \phi \partial_\nu \phi \left[1 + c_2 X_q + \dots \right) \right],$$

and for the scalar ϕ we have

$$\Box \phi (1 + 4c_2 X_q + \dots)$$
$$+ q^{\mu\nu} \partial_\nu \phi \partial_\mu \phi \left[c_2 X_q + c_4 X_q^3 + \dots \right] + \left[c_2'(\chi) X_q^2 + \dots \right] = \tilde{U}_{,\phi}. \tag{6.56}$$

With the assumption that the kinetic term for the field is sufficiently small such that we can ignore all terms of order higher than one, we get the Einstein equations with a scalar field. With the FLRW solution with zero spatial curvature the temporal component of the metric field equations becomes

$$3H^2 = \tilde{U} + \frac{1}{2} \dot{\phi}^2, \tag{6.57}$$

and the equations of motion for the scalar field become

$$\ddot{\phi} + 3H\dot{\phi} = -\tilde{U}_{,\phi} \tag{6.58}$$

With the slow-roll conditions $|\ddot{\phi}| \ll |3H\dot{\phi}|$ and $\dot{\phi}^2 \ll \tilde{U}$ the equations reduce to the familiar slow-roll equations with the modified potential $\tilde{U}$

$$3H^2 = \tilde{U}$$
$$3H\dot{\phi} = -\tilde{U}', \tag{6.59}$$

The first few slow-roll parameter are thus

$$\tilde{\epsilon} = \frac{1}{2} \left(\frac{\tilde{U}'}{\tilde{U}} \right)^2$$
$$\tilde{\eta} = \frac{\tilde{U}''}{\tilde{U}} \tag{6.60}$$
$$\tilde{\zeta} = \frac{\tilde{U}' \tilde{U}'''}{\tilde{U}^2}.$$

Now it is clear that $\dot{\phi}^2 = 2/3\tilde{U}\epsilon$ and with (6.50) we get that $\dot{\chi}^2 = 2/3U\epsilon$ thus validating our assumption that $\dot{\chi}^2$ is small in slow-roll. What is left to determine is when does this approximation break down and when the higher-order terms in X_q should be taken in to account.

6.3   Higher-order gravity frame representation

Now we return to the second method of arriving to these slow-roll equations. We derive the equations of motion for the action (6.9) and show that by performing the disformal transformation they are equal to the Einstein frame equations. We derived the equations of motion for the higher-order terms in the most general case in the chapter 5.3. In the case of the action (6.1) and vanishing torsion these simplify considerably. For the metric we obtain

$$\left(\frac{1}{2} + \beta R\right) R_{\mu\nu} + \alpha R_\mu{}^\lambda R_{\nu\lambda} - \frac{1}{2}\partial_\mu\chi\partial_\nu\chi$$
$$- \frac{1}{2}g_{\mu\nu}\left[\frac{1}{2}(1 + \beta R)R + \frac{\alpha}{2}R_{\sigma\lambda}R^{\sigma\lambda} + \mathcal{L}_m\right] = 0, \tag{6.61}$$

where $\mathcal{L}_m$ is defined in (6.11) as before. For the connection we get

$$\nabla_\sigma\left\{\sqrt{-g}\left[(1 + 2\beta R)\,g^{\mu\nu} + 2\alpha R^{\mu\nu}\right]\right\} = 0. \tag{6.62}$$

We see that this is a metric compatibility condition for a metric defined with

$$\sqrt{-q}q^{\mu\nu} = \sqrt{-g}\left[(1 + 2\beta R)\,g^{\mu\nu} + 2\alpha R^{\mu\nu}\right]. \tag{6.63}$$

Thus, the connection is the Levi–Civita connection for the metric $q_{\mu\nu}$. We need to solve $q_{\mu\nu}$ in terms of $g_{\mu\nu}$ and the scalar field χ. This is accomplished by solving R and $R_{\mu\nu}$. We can start by solving R, which is easily solved by taking the trace of (6.61). This gives

$$R = g^{\mu\nu}\partial_\mu\chi\partial_\nu\chi + 4U(\chi). \tag{6.64}$$

Solving $R_{\mu\nu}$ is more complicated. In a FLRW universe the Ricci tensor must be of the form

$$R_{\mu\nu} = R_1(g_{\lambda\sigma}, \chi)g_{\mu\nu} + R_2(g_{\lambda\sigma}, \chi)u_\mu u_\nu. \tag{6.65}$$

We can plug in this ansatz into the the equations of motion for the metric (6.61) and obtain a system of equations for R_1 and R_2 ($\partial_\mu\chi = \dot{\chi}u_\mu$):

$$\left(\frac{1}{2} + \beta R\right) R_1 + \alpha R_1^2 - \frac{1}{4}\left(R + \beta R^2 + 2\mathcal{L}_m\right)$$
$$- \frac{\alpha}{4}\left(4R_1^2 + R_2^2 - 2R_1 R_2\right) = 0, \tag{6.66}$$

$$\left(\frac{1}{2} + \beta R\right) R_2 - \alpha R_2^2 + 2\alpha R_1 R_2 - \frac{1}{2}\dot{\chi}^2 = 0, \tag{6.67}$$

This system of equations is solved by

$$R_1^\pm = \frac{1 + 12\alpha\mathcal{L}_m + 2R(3\alpha + 2\beta)(1 + \beta R) - 3\alpha\dot{\chi}^2 \pm A}{8(\alpha + 2\alpha\beta R)}$$

$$\tag{6.68}$$

$$R_2^\pm = \frac{1 + 4\alpha\mathcal{L}_m + 2R(\alpha + 2\beta)(1 + \beta R) - \alpha\dot{\chi}^2 \pm A}{2\alpha(1 + 2\beta R)},$$

with

$$A \equiv \sqrt{-4\alpha(1 + 2\beta R)^2 \dot{\chi}^2 + [1 + 4\alpha\mathcal{L}_m + 2R(\alpha + 2\beta)(1 + \beta R) - \alpha\dot{\chi}^2]^2}. \quad (6.69)$$

We can now solve the metric $q^{\mu\nu}$ from (6.63) giving

$$
\begin{aligned}
q^{\mu\nu} &= \frac{(1 + 2\beta R + 2\alpha R_1)g^{\mu\nu} + 2\alpha R_2 u^\mu u^\nu}{\sqrt{(1 + 2\beta R + 2\alpha R_1)^3 \left[(1 + 2\beta R + 2\alpha R_1) - 2\alpha R_2\right]}} \\
&\equiv \frac{1}{q_1}g^{\mu\nu} + q_3 u^\mu u^\nu
\end{aligned}
\quad (6.70)
$$

where we have used the previously found identity (6.39) to write the determinant of $q_{\mu\nu}$ in terms of $g_{\mu\nu}$ and χ. We can use the previously obtained identity for the inverse metric (6.29) and write the inverse, for later convenience, as

$$q_{\mu\nu} \equiv q_1 g_{\mu\nu} + q_2 u_\mu u_\nu, \quad q_2 = \frac{q_3}{\frac{1}{q_1}\left(\frac{q_3}{q_1} - \frac{1}{q_1}\right)} \quad (6.71)$$

The solution (6.68) with the minus signs will correspond to the solution found with the earlier method and it is in fact the inverse of the transformation found before. The second solution (6.68) with the plus signs turns out to be non-invertible, at least in the inflationary region, and thus was not found by the earlier method. Additionally, the resulting disformal transformation does not preserve the Lorentzian signature of the metric $q_{\mu\nu}$ in the regime where $\dot{\chi}^2 \ll U(\chi)$. With the positive signed solution the constraint for preserving the Lorentzian signature (6.24), to zeroth order in $\dot{\chi}^2$, simplifies to

$$-\left[1 + 8(\alpha + \beta)U(\chi)\right] > 0, \quad (6.72)$$

which is not satisfied. In contrast with the negatively signed solution the constraint to zeroth order becomes

$$1 + (2\alpha + 8\beta)U(\chi) > 0, \quad (6.73)$$

which is satisfied. The positively signed solution (6.68) is thus dismissed as unphysical.

Now let us continue deriving the equations of motion. We can start by computing the connection explicitly by plugging in the solution of $q_{\mu\nu}$ into the definition of Levi–Civita connection. After a straightforward simplification, with the abbreviations (6.70),(6.71) and making use of (3.2), we obtain

$$
\begin{aligned}
\mathring{\Gamma}^\sigma_{\mu\nu}(q) &= \mathring{\Gamma}^\sigma_{\mu\nu}(g) + \left[\frac{1}{2}\dot{q}_1\left(q_3 - \frac{1}{q_1}\right) + q_3 q_1 H\right]u^\sigma h_{\mu\nu} \\
&\quad + \frac{1}{2}\left[\frac{1}{q_1}(\dot{q}_2 - \dot{q}_1) + q_3(\dot{q}_1 - \dot{q}_2)\right]u^\sigma u_\mu u_\nu + \frac{1}{2}\frac{\dot{q}_1}{q_1}h^\sigma_{\ (\mu}u_{\nu)} \\
&\equiv \mathring{\Gamma}^\sigma_{\mu\nu}(g) + L^\sigma_{\mu\nu},
\end{aligned}
\quad (6.74)
$$

where $H = \dot{a}/a$ is the Hubble parameter and $h_{\mu\nu}$ is the metric of the spacelike slices. We also used the fact that $u_\lambda \mathring{\Gamma}^\lambda_{\mu\nu}(g) = H h_{\mu\nu}$ in FLRW universe. Now we can plug this into the definition of the Riemann tensor and see how it is modified, this gives

$$
\begin{aligned}
R^{(q)}_{\mu\nu} &= R^{(g)}_{\mu\nu} + \mathring{\nabla}^{(g)}_\lambda L^\lambda_{\mu\nu} - \mathring{\nabla}^{(g)}_\nu L^\lambda_{\mu\lambda} + L^\lambda_{\lambda\sigma} L^\sigma_{\mu\nu} - L^\lambda_{\nu\sigma} L^\sigma_{\lambda\mu} \\
&\equiv R^{(g)}_{\mu\nu} + L_1 g_{\mu\nu} + L_2 u_\mu u_\nu.
\end{aligned}
\tag{6.75}
$$

Now instead of plugging this into the full equations of motion for the metric we can plug this into (6.65) and we get

$$
R^{(g)}_{\mu\nu} = (R_1 - L_1)g_{\mu\nu} + (R_2 - L_2)u_\mu u_\nu,
\tag{6.76}
$$

where L_1 and L_2 have to be explicitly computed from (6.74) and (6.75). This is a straightforward but tedious task and so we shall not do it explicitly here. When L_1, L_2 are computed using the solutions (6.68) for R_1^-, R_2^- and plugged into (6.76) we get the Friedmann equations. The full equation is cumbersome to write down and we will not need it for our purposes, thus we will again expand in powers of $\dot{\chi}$ and do not write down terms $\mathcal{O}(\dot{\chi}^2)$. For the negatively signed solution $\{R_1^-, R_2^-\}$ we get

$$
3H^2 = U - \frac{12H\dot{\chi}U'\left[\alpha + 3\beta + 3(\alpha^2 + 6\alpha\beta + 8\beta^2)\right]}{\left[1 + (4\alpha + 8\beta)U\right]\left[1 + (2\alpha + 8\beta)U\right]} + \mathcal{O}(\dot{\chi}^2).
\tag{6.77}
$$

The equations of motion for the scalar field are not modified and are simply

$$
\ddot{\chi} + 3H\dot{\chi} = -U_{,\chi}.
\tag{6.78}
$$

These are the equations of motion in the higher-order gravity frame. To see that these are indeed equivalent to the Einstein frame counterparts, that we obtained earlier, we need to know how the Hubble parameter and the time derivatives change under the disformal transformation.

These transformation rules can be seen by looking at the invariant line element. With $g_{\mu\nu} = \gamma_1 q_{\mu\nu} + \gamma_2 u_\mu u_\nu$:

$$
\begin{aligned}
ds^2 = g_{\mu\nu}dx^\mu dx^\nu &= \gamma_2 u_\mu u_\nu dx^\mu dx^\nu + \gamma_1 q_{\mu\nu} dx^\mu dx^\nu \\
&= -(\gamma_1 - \gamma_2)dt^2 + \gamma_1 a^2(t)\delta_{ij}dx^i dx^j,
\end{aligned}
\tag{6.79}
$$

where we used a frame $u_\mu = (-1, 0, 0, 0)$ and $q_{\mu\nu} = \text{diag}(-1, a^2, a^2, a^2)$. From this we see that the scale factor and time coordinate transform as

$$
\begin{aligned}
dt &\to \sqrt{\gamma_1 - \gamma_2}\, dt \\
a(t) &\to \sqrt{\gamma_1}\, a(t).
\end{aligned}
\tag{6.80}
$$

Thus the Hubble parameter transforms like

$$
H \to \frac{1}{\sqrt{\gamma_1 - \gamma_2}}\left[H + \frac{1}{2}\frac{\dot{\gamma}_1}{\gamma_1}\right].
\tag{6.81}
$$

Applying these rules to the scalar field equations of motion (with our previous definitions $\gamma_1 = 1/q_1$ and $\gamma_2 = -q_2/(q_1 - q_2)$) and keeping only first order terms in $\dot{\chi}$ we get

$$[1 + (\alpha + 4\beta)U] (\ddot{\chi} + 3H\dot{\chi}) = -U_{,\chi}. \tag{6.82}$$

Now making the field redefinition (6.50) we get the same result as before (6.58). The equation (6.77) is also straightforward. Plugging in the transformations and keeping only linear terms we get

$$
\begin{aligned}
3H^2 \left[1 + (2\alpha + 8\beta)U\right] &- \frac{12H\dot{\chi}U'\left[\alpha + 3\beta + 3(\alpha^2 + 6\alpha\beta + 8\beta^2)\right]}{1 + (4\alpha + 8\beta)U} \\
&= U - \frac{12H\dot{\chi}U'\left[\alpha + 3\beta + 3(\alpha^2 + 6\alpha\beta + 8\beta^2)\right]}{1 + (4\alpha + 8\beta)U} + \mathcal{O}(\dot{\chi}^2),
\end{aligned}
\tag{6.83}
$$

which simplifies to the previous result (6.57). This was just to show another way of obtaining the slow-roll equations in the Einstein frame. We will not pursue the analysis in the higher-order gravity frame further.

With the slow-roll equations of motions in the Einstein frame we are now ready to analyze how the inflationary behavior is modified due to the higher-order terms.

6.4 Changes to inflationary observables

We can now see how the inflationary observables are modified by the higher-order curvature terms. It is easier to write the slow-roll parameters (6.60) with the field variable χ. Using (6.50) and (6.51), simple manipulations yield

$$
\begin{aligned}
\tilde{\epsilon}_U &= \frac{1}{2}\left(\frac{\tilde{U}'}{\tilde{U}}\right)^2 = \frac{\epsilon_U}{1 + (2\alpha + 8\beta)U(\chi)} \\
\tilde{\eta}_U &= \frac{\tilde{U}''}{\tilde{U}} = \eta - 3\frac{(2\alpha + 8\beta)U(\chi)}{1 + (2\alpha + 8\beta)U(\chi)}\epsilon,
\end{aligned}
\tag{6.84}
$$

where $\epsilon_U = 1/2(U_{,\chi}/U)^2$, $\eta_U = U_{,\chi\chi}/U$ are the slow-roll parameters that we obtained previously without the higher-order curvature terms in chapter 4. The field value at the end of inflation is again obtained from $|\tilde{\eta}(\chi_{end})| = 1$, which in the large field limit $\sqrt{\xi}\chi \gg 1$ results in the same value as obtained earlier

$$\chi_{end} = \frac{1}{2\sqrt{\xi}} \ln\left(32\xi\right). \tag{6.85}$$

We also find that in the slow-roll approximation the formula for the number of e-folds is unchanged

$$N_* \simeq \int_{\phi_{end}}^{\phi_*} \frac{d\phi}{\sqrt{2\tilde{\epsilon}}} = \int_{\chi_{end}}^{\chi_*} \frac{d\chi}{\sqrt{2\epsilon_U}}. \tag{6.86}$$

We thus obtain the same relation between the field and the number of e-folds

$$\chi(N_*) = \frac{1}{2\sqrt{\xi}} \ln\left(32\xi N_*\right). \tag{6.87}$$

This holds assuming that our expansion in terms of the kinetic term of the field is still valid when approaching the end of inflation. This needs further investigations.

Assuming our expansion to be valid we see that the scalar spectrum amplitude is unchanged

$$A_s = \frac{1}{24\pi^2}\frac{\tilde{U}}{\tilde{\epsilon}} = \frac{1}{24\pi^2}\frac{1 + (2\alpha + 8\beta)U}{1 + (2\alpha + 8\beta)U}\frac{U}{\epsilon_U} = \frac{1}{24\pi^2}\frac{U}{\epsilon_U}. \tag{6.88}$$

Thus matching this to the observed amplitude gives the same requirement for the non-minimal coupling (4.65). The tilt of the scalar spectrum is also unchanged to first order in the slow-roll parameters

$$n_s - 1 = -6\tilde{\epsilon} + 2\tilde{\eta} = -6\frac{\epsilon_U}{1 + (2\alpha + 8\beta)U(\chi)} + 2\eta_U - 6\frac{(2\alpha + 8\beta)U(\chi)}{1 + (2\alpha + 8\beta)U(\chi)}\epsilon_U$$
$$= -6\epsilon_U + 2\eta_U. \tag{6.89}$$

What changes is the spectrum of tensor perturbations. The amplitude

$$A_t = \frac{2}{3\pi^2}\frac{U}{1 + (2\alpha + 8\beta)U}, \tag{6.90}$$

tilt of the spectrum

$$n_t = -2\tilde{\epsilon} = \frac{-2\epsilon_U}{1 + (2\alpha + 8\beta)U}, \tag{6.91}$$

and the tensor-to-scalar ratio

$$r = 16\tilde{\epsilon} = \frac{16\epsilon_U}{1 + (2\alpha + 8\beta)U} \tag{6.92}$$

are all suppressed by the same factor. The observables in terms of the number of e-folds (approximating with $\xi \gg 1$) become

$$A_t \simeq \frac{\lambda}{\lambda(2\alpha + 8\beta) + 4\xi^2} \tag{6.93}$$

$$n_t \simeq -\frac{\xi}{2N_*^2\left[\lambda(\alpha + 4\beta) + 2\xi^2\right]} \tag{6.94}$$

$$r \simeq -\frac{4\xi}{N_*^2\left[\lambda(\alpha + 4\beta) + 2\xi^2\right]}. \tag{6.95}$$

Thus we see that both higher-order curvature terms βR^2 and $\alpha R_{(\mu\nu)}R^{(\mu\nu)}$ have the same effect of suppressing the tensor perturbation spectrum but keeping the scalar perturbation spectrum unchanged. Again when we take the limit $\alpha \to 0$ we recover the results of [76].

Higgs inflation and higher-order gravity in the metric formulation has been studied in the literature. In ref. [77] Higgs inflation with the addition of $\mathring{R}^2$ term was analyzed. As we saw in section 5.1 in the metric formulation there is additional degrees of freedom of gravity and the inflationary model becomes a multi field model. It was found out that in the case of metric formulation the pure Higgs inflation, where the inflation is driven completely by the Higgs field, is ruined.

Chapter 7

Conclusions

We have reviewed the metric and Palatini formulations in a comparative manner and discussed their subtle differences in the case of the Einstein–Hilbert action and more obvious differences when considering higher-order theories of gravity.

We reviewed Higgs inflation in both the metric and Palatini formulations and obtained the inflationary observables in terms of the number of e-folds. We computed the numerical values for the observables, and saw that they are compatible with current observational bounds. The tensor-to-scalar ratio is suppressed by the large non-minimal coupling parameter ξ in the Palatini formulation $r \simeq 2.1 \times 10^{-13}\lambda^{-1}$, compared to $r \simeq 0.00396$ in metric formulation. The spectral tilt is also slightly smaller in Palatini formulation due to the different reheating period. These allow us to distinguish Palatini Higgs inflation from the metric Higgs inflation. This could possibly be accomplished by the next generation of planned experiments [78].

We briefly discussed the problems of ghost modes appearing in both metric and Palatini formulations when considering higher-order gravity. This problem requires further research. We derived the equations of motion for all the different parity preserving quadratic curvature invariants in the Palatini formulation.

When we considered adding higher-order curvature terms to the Higgs inflation scenario we simplified our action to contain only the symmetric part of the Ricci tensor. Actions constructed this way do not suffer from ghosts, due to the fact that no new gravitational degrees of freedom are present. This also allows us to transform our action to the Einstein frame. By adding only terms $\beta R^2 + \alpha R_{(\mu\nu)}R^{(\mu\nu)}$ we were able to find the Einstein frame representation in the inflationary region by finding an approximate solution in powers of the kinetic term of the inflaton field, i.e. the small parameter $\dot{\chi}^2/H^2$. This method could possibly be applied to more general Lagrangians than the one we considered here. However, the region of validity of this approximation needs further investigation. Furthermore, the two parameters α and β were assumed to be positive to assure correct signs for the kinetic terms. However, there could be other valid regions for these parameters that give the correct sings, which were not explored in this thesis.

With the assumption that we can extrapolate the approximate solution to the end of inflation the number of e-folds turns out to be unchanged when comparing to the case where no higher-order terms are present. This results to the same relation between the number of e-folds and the inflation field. The scalar amplitude and

tilt of the scalar spectrum are also unchanged. The changes appeared in the tensor perturbation spectrum. The amplitude, spectral tilt and the tensor to scalar ratio are all suppressed by the combination of the couplings of the higher-order terms and the Higgs self coupling: $1 + (2\alpha + 8\beta)U$. We conclude that in this simple case the higher-order curvature terms do not ruin the Palatini Higgs inflation.

Appendix A

Useful identities

A.1   Palatini identity

Start from the definition of the Riemann tensor (0.4) and take the variation with respect to the connection

$$\delta R^\rho{}_{\mu\lambda\nu} = \partial_\lambda \delta\Gamma^\rho_{\nu\mu} + \Gamma^\rho_{\lambda\sigma}\delta\Gamma^\sigma_{\nu\mu} + \delta\Gamma^\rho_{\lambda\sigma}\Gamma^\sigma_{\nu\mu} - (\lambda \leftrightarrow \nu). \tag{A.1}$$

Notice that difference of connections is a tensor, hence $\delta\Gamma^\rho_{\nu\mu}$ is a tensor, so the covariant derivative is well defined

$$\nabla_\lambda(\delta\Gamma^\rho_{\nu\mu}) = \partial_\lambda \delta\Gamma^\rho_{\nu\mu} + \Gamma^\rho_{\lambda\sigma}\delta\Gamma^\sigma_{\nu\mu} - \Gamma^\sigma_{\lambda\nu}\delta\Gamma^\rho_{\sigma\mu} - \Gamma^\sigma_{\lambda\mu}\delta\Gamma^\rho_{\nu\sigma}. \tag{A.2}$$

Now gathering terms

$$\delta R^\rho{}_{\mu\lambda\nu} = \left[\partial_\lambda \delta\Gamma^\rho_{\nu\mu} + \Gamma^\rho_{\lambda\sigma}\delta\Gamma^\sigma_{\nu\mu} - \Gamma^\sigma_{\lambda\mu}\delta\Gamma^\rho_{\nu\sigma}\right] \\ - \left[\partial_\nu \delta\Gamma^\rho_{\lambda\mu} + \Gamma^\rho_{\nu\sigma}\delta\Gamma^\sigma_{\lambda\mu} - \Gamma^\sigma_{\nu\mu}\delta\Gamma^\rho_{\lambda\sigma}\right]. \tag{A.3}$$

Notice the terms inside the brackets almost equal the covariant derivative of the variation of the connection, we can make it equal by adding and subtracting the missing terms

$$\delta R^\rho{}_{\mu\lambda\nu} = \left[\partial_\lambda \delta\Gamma^\rho_{\nu\mu} + \Gamma^\rho_{\lambda\sigma}\delta\Gamma^\sigma_{\nu\mu} - \Gamma^\sigma_{\lambda\mu}\delta\Gamma^\rho_{\nu\sigma} - \Gamma^\sigma_{\lambda\nu}\delta\Gamma^\rho_{\sigma\mu}\right] \\ - \left[\partial_\nu \delta\Gamma^\rho_{\lambda\mu} + \Gamma^\rho_{\nu\sigma}\delta\Gamma^\sigma_{\lambda\mu} - \Gamma^\sigma_{\nu\mu}\delta\Gamma^\rho_{\lambda\sigma} - \Gamma^\sigma_{\nu\lambda}\delta\Gamma^\rho_{\sigma\mu}\right] + T^\sigma{}_{\lambda\nu}\,\delta\Gamma^\rho_{\sigma\mu} \tag{A.4}$$

and we get the result

$$\delta R^\rho{}_{\mu\lambda\nu} = \nabla_\lambda(\delta\Gamma^\rho_{\nu\mu}) - \nabla_\nu(\delta\Gamma^\rho_{\lambda\mu}) + T^\sigma{}_{\lambda\nu}\,\delta\Gamma^\rho_{\sigma\mu}. \tag{A.5}$$

A.2   Tensor densities

A general tensor density $\mathfrak{T}$ of weight ω is an object that transforms as

$$\mathfrak{T}'^{\mu_1...\mu_n}{}_{\nu_1...\nu_n} = \left|\frac{\partial x^\alpha}{\partial x'_\beta}\right|^\omega \frac{\partial x'^{\mu_1}}{\partial x^{\sigma_1}}\cdots\frac{\partial x'^{\mu_n}}{\partial x^{\sigma_n}}\cdots\frac{\partial x^{\tau_1}}{\partial x'^{\nu_1}}\cdots\frac{\partial x^{\tau_n}}{\partial x'^{\nu_n}}\mathfrak{T}^{\sigma_1...\sigma_n}{}_{\tau_1...\tau_n}, \tag{A.6}$$

where || denotes the determinant. Scalar density of weight ω transforms as

$$\rho' = \left|\frac{\partial x^\alpha}{\partial x'_\beta}\right|^\omega \rho. \tag{A.7}$$

A way of looking at a scalar density of weight one is as a single independent component of a covariant antisymmetric tensor $\rho = \rho_{[\mu_1 \dots \mu_n]} = \rho_{\mu_1 \dots \mu_n}$, see e.g. [18, ch. II]. It is easy to see that a single component of this transforms as a scalar density. Now

$$\nabla_\mu(\rho_{\nu_1 \dots \nu_n}) = \partial_\mu \rho_{\nu_1 \dots \nu_n} - \Gamma^\sigma_{\mu\nu_1} \rho_{\sigma\nu_2 \dots \nu_n} - \Gamma^\sigma_{\mu\nu_2} \rho_{\nu_1\sigma \dots n} - \dots - \Gamma^\sigma_{\mu\nu_n} \rho_{\nu_1 \dots \sigma}. \tag{A.8}$$

Since all the non zero components are equal up-to a sign we can focus on one of them

$$\nabla_\mu(\rho_{0 \dots n}) = \partial_\mu \rho_{0 \dots n} - \Gamma^\sigma_{\mu 0} \rho_{\sigma 1 \dots n} - \Gamma^\sigma_{\mu 1} \rho_{0\sigma \dots n} - \dots - \Gamma^\sigma_{\mu n} \rho_{0 \dots \sigma}, \tag{A.9}$$

we see that this equals

$$\nabla_\mu \rho = \partial_\mu \rho - \Gamma^\sigma_{\mu\sigma} \rho. \tag{A.10}$$

This is the result for a scalar density of weight $+1$. From this it is easy to generalize to arbitrary scalar density of weight ω by requiring that the covariant derivative of such a density transforms as a $(0,1)-$tensor density. One finally obtains

$$\nabla_\mu \rho = \partial_\mu \rho - \omega \Gamma^\sigma_{\mu\sigma} \rho, \tag{A.11}$$

for a scalar density ρ of weight ω.

Appendix B

Equations of motion in the Palatini formulation

Here we list the equations of motion for the different curvature invariants in the action (5.25) in Palatini formulation.

- $\mathcal{L} = R^2$

$$2RR_{(\alpha\beta)} - \frac{1}{2}g_{\alpha\beta}R^2 = 0. \tag{B.1}$$

$$2\nabla_\lambda(\sqrt{-g}Rg^{\nu\lambda})\delta^\mu_\sigma - 2\nabla_\sigma(\sqrt{-g}Rg^{\mu\nu})$$
$$+ 2\sqrt{-g}R\left[g^{\mu\nu}T^\lambda_{\ \lambda\nu} - g^{\nu\rho}T^\lambda_{\ \lambda\rho}\delta^\mu_\sigma + g^{\nu\lambda}T^\mu_{\ \sigma\lambda}\right] = 0. \tag{B.2}$$

- $\mathcal{L} = R^{\mu\nu}R_{\mu\nu}$

$$R_\alpha^{\ \lambda}R_{\beta\lambda} + R^\lambda_{\ \alpha}R_{\lambda\beta} - \frac{1}{2}g_{\alpha\beta}R^{\mu\nu}R_{\mu\nu} = 0. \tag{B.3}$$

$$2\nabla_\lambda(\sqrt{-g}R^{\mu\lambda})\delta^\nu_\sigma - 2\nabla_\sigma(\sqrt{-g}R^{\mu\nu})$$
$$+ 2\sqrt{-g}\left[R^{\mu\nu}T^\lambda_{\ \lambda\sigma} - R^{\mu\rho}T^\lambda_{\ \lambda\rho}\delta^\nu_\sigma + R^{\mu\lambda}T^\nu_{\ \sigma\lambda}\right] = 0. \tag{B.4}$$

- $\mathcal{L} = R^{\mu\nu}R_{\nu\mu}$

$$R_\alpha^{\ \lambda}R_{\lambda\beta} + R^\lambda_{\ \alpha}R_{\beta\lambda} - \frac{1}{2}g_{\alpha\beta}R^{\mu\nu}R_{\nu\mu} = 0. \tag{B.5}$$

$$2\nabla_\lambda(\sqrt{-g}R^{\lambda\mu})\delta^\nu_\sigma - 2\nabla_\sigma(\sqrt{-g}R^{\nu\mu})$$
$$+ 2\sqrt{-g}\left[R^{\nu\mu}T^\lambda_{\ \lambda\sigma} - R^{\rho\mu}T^\lambda_{\ \lambda\rho}\delta^\nu_\sigma + R^{\lambda\mu}T^\nu_{\ \sigma\lambda}\right] = 0. \tag{B.6}$$

- $\mathcal{L} = \hat{R}^{\mu\nu}\hat{R}_{\mu\nu}$

$$2R_{\mu(\alpha|\nu|\beta)}\hat{R}^{\mu\nu} + \hat{R}_{\mu\alpha}\hat{R}^\mu_{\ \beta} - \hat{R}_{\alpha\nu}\hat{R}_\beta^{\ \nu} - \frac{1}{2}g_{\alpha\beta}\hat{R}^{\mu\nu}\hat{R}_{\mu\nu} = 0. \tag{B.7}$$

$$2\nabla_\lambda(\sqrt{-g}g^{\mu\lambda}\hat{R}_\sigma^{\ \nu}) - 2\nabla_\lambda(\sqrt{-g}g^{\mu\nu}\hat{R}_\sigma^{\ \lambda})$$
$$+ 2\sqrt{-g}\left[g^{\mu\nu}\hat{R}_\sigma^{\ \lambda}T^\rho_{\ \rho\lambda} - g^{\mu\lambda}\hat{R}_\sigma^{\ \nu}T^\rho_{\ \rho\lambda} + g^{\mu\lambda}\hat{R}_\sigma^{\ \rho}T^\nu_{\ \rho\lambda}\right] = 0. \tag{B.8}$$

- $\mathcal{L} = \hat{R}^{\mu\nu}\hat{R}_{\nu\mu}$

$$2R_{\mu(\alpha|\nu|\beta)}\hat{R}^{\nu\mu} - \frac{1}{2}g_{\alpha\beta}\hat{R}^{\mu\nu}\hat{R}_{\nu\mu} = 0. \tag{B.9}$$

$$2\nabla_\lambda(\sqrt{-g}g^{\mu\lambda}\hat{R}^\nu_{\ \sigma}) - 2\nabla_\lambda(\sqrt{-g}g^{\mu\nu}\hat{R}^\lambda_{\ \sigma})$$
$$+ 2\sqrt{-g}\left[g^{\mu\nu}\hat{R}^\lambda_{\ \sigma}T^\rho_{\ \rho\lambda} - g^{\mu\lambda}\hat{R}^\nu_{\ \sigma}T^\rho_{\ \rho\lambda} + g^{\mu\lambda}\hat{R}^\rho_{\ \sigma}T^\nu_{\ \rho\lambda}\right] = 0. \tag{B.10}$$

- $\mathcal{L} = R_{\mu\nu}\hat{R}^{\mu\nu}$

$$R_{\mu(\alpha|\nu|\beta)}R^{\mu\nu} + \frac{1}{2}R_{\mu\alpha}\hat{R}^{\mu}{}_{\beta} + \frac{1}{2}R_{\mu\beta}\hat{R}^{\mu}{}_{\alpha} - \frac{1}{2}g_{\alpha\beta}R_{\mu\nu}\hat{R}^{\mu\nu} = 0. \tag{B.11}$$

$$\begin{aligned}
&\nabla_\lambda(\sqrt{-g}\hat{R}^{\mu\lambda})\delta^\nu_\sigma - \nabla_\sigma(\sqrt{-g}\hat{R}^{\mu\nu}) \\
&+ \nabla_\lambda(\sqrt{-g}g^{\mu\lambda}R_\sigma{}^\nu) - \nabla_\lambda(\sqrt{-g}g^{\mu\nu}R_\sigma{}^\lambda) \\
&+ 2\sqrt{-g}\left[\hat{R}^{\mu\nu}T^\lambda{}_{\lambda\sigma} - \hat{R}^{\nu\lambda}T^\rho{}_{\rho\lambda}\delta^\mu_\sigma + \hat{R}^{\nu\lambda}T^\mu{}_{\sigma\lambda}\right] \\
&+ 2\sqrt{-g}\left[g^{\mu\nu}R_\sigma{}^\lambda T^\rho{}_{\rho\lambda} - g^{\mu\lambda}R_\sigma{}^\nu T^\rho{}_{\rho\lambda} + g^{\mu\rho}R_\sigma{}^\lambda T^\nu{}_{\lambda\rho}\right] = 0.
\end{aligned} \tag{B.12}$$

- $\mathcal{L} = R_{\mu\nu}\hat{R}^{\nu\mu}$

$$R_{\mu(\alpha|\nu|\beta)}R^{\nu\mu} + \frac{1}{2}R_{\beta\nu}\hat{R}^\nu{}_\alpha + \frac{1}{2}R_{\alpha\nu}\hat{R}^\nu{}_\beta - \frac{1}{2}g_{\alpha\beta}R_{\mu\nu}\hat{R}^{\nu\mu} = 0. \tag{B.13}$$

$$\begin{aligned}
&\nabla_\lambda(\sqrt{-g}\hat{R}^{\lambda\mu})\delta^\nu_\sigma - \nabla_\sigma(\sqrt{-g}\hat{R}^{\nu\mu}) \\
&+ \nabla_\lambda(\sqrt{-g}g^{\mu\lambda}R^\nu{}_\sigma) - \nabla_\lambda(\sqrt{-g}g^{\mu\nu}R^\lambda{}_\sigma) \\
&+ 2\sqrt{-g}\left[\hat{R}^{\nu\mu}T^\lambda{}_{\lambda\sigma} - \hat{R}^{\lambda\nu}T^\rho{}_{\rho\lambda}\delta^\mu_\sigma + \hat{R}^{\lambda\nu}T^\mu{}_{\sigma\lambda}\right] \\
&+ 2\sqrt{-g}\left[g^{\mu\nu}R^\lambda{}_\sigma T^\rho{}_{\rho\lambda} - g^{\mu\lambda}R^\nu{}_\sigma T^\rho{}_{\rho\lambda} + g^{\mu\rho}R^\lambda{}_\sigma T^\nu{}_{\lambda\rho}\right] = 0.
\end{aligned} \tag{B.14}$$

- $\mathcal{L} = R'_{\mu\nu}R'^{\mu\nu}$

$$2R'_{\alpha\mu}R_\beta{}^\mu - \frac{1}{2}g_{\alpha\beta}R'_{\mu\nu}R'^{\mu\nu} = 0. \tag{B.15}$$

$$\left[4\nabla_\lambda(\sqrt{-g}R'^{\nu\lambda}) + 4\sqrt{-g}R'^{\lambda\nu}T^\rho{}_{\rho\lambda} + 2\sqrt{-g}R'^{\lambda\rho}T^\nu{}_{\lambda\rho}\right]\delta^\mu_\sigma = 0. \tag{B.16}$$

- $\mathcal{L} = R_{[\mu\nu]}R'^{\mu\nu}$

$$R_{[\beta\nu]}R'_\alpha{}^\nu + R_{[\alpha\nu]}R'_\beta{}^\nu - \frac{1}{2}g_{\alpha\beta}R_{[\mu\nu]}R'^{\mu\nu} = 0. \tag{B.17}$$

$$\begin{aligned}
&\left[4\nabla_\lambda(\sqrt{-g}R^{[\nu\lambda]}) + 4\sqrt{-g}T^\rho{}_{\rho\lambda}R^{[\lambda\nu]} + 2\sqrt{-g}R^{[\lambda\rho]}T^\nu{}_{\lambda\rho}\right]\delta^\mu_\sigma \\
&+ \frac{1}{2}\left[\nabla_\sigma(\sqrt{-g}R'^{\nu\mu}) - \nabla_\lambda(\sqrt{-g}R'^{\mu\lambda})\delta^\nu_\sigma + \sqrt{-g}R'^{\mu\lambda}T^\rho{}_{\rho\lambda}\delta^\nu_\sigma \right. \\
&\left. - \sqrt{-g}R'^{\nu\mu}T^\rho{}_{\rho\sigma} + \sqrt{-g}T^\nu{}_{\sigma\lambda}R'^{\mu\lambda}\right] = 0.
\end{aligned} \tag{B.18}$$

- $\mathcal{L} = \hat{R}_{[\mu\nu]}R'^{\mu\nu}$

$$R_{\mu(\alpha|\nu|\beta)}R'^{\mu\nu} + \hat{R}_{[\mu\beta]}R'^\mu{}_\alpha - \frac{1}{2}g_{\alpha\beta}\hat{R}_{[\mu\nu]}R'^{\mu\nu} = 0. \tag{B.19}$$

$$\begin{aligned}
&\left[4\nabla_\lambda(\sqrt{-g}\hat{R}^{[\nu\lambda]}) + 4\sqrt{-g}\hat{R}^{[\lambda\nu]}T^\rho{}_{\rho\lambda} + \sqrt{-g}\hat{R}^{[\lambda\rho]}T^\nu{}_{\lambda\rho}\right]\delta^\mu_\sigma \\
&\nabla_\lambda(\sqrt{-g}g^{\mu\lambda}R'_\sigma{}^\nu) - \nabla_\lambda(\sqrt{-g}g^{\mu\nu}R'_\sigma{}^\nu) + \sqrt{-g}\left[g^{\mu\nu}R'_\sigma{}^\lambda T^\rho{}_{\rho\lambda}\right. \\
&\left. - g^{\mu\lambda}R'_\sigma{}^\nu T^\rho{}_{\rho\lambda} + g^{\mu\lambda}R'_\sigma{}^\rho T^\nu{}_{\rho\lambda}\right] = 0.
\end{aligned} \tag{B.20}$$

- $\mathcal{L} = R_{\mu\nu\rho\lambda}R^{\mu\nu\rho\lambda}$

$$R_{\mu\alpha\rho\lambda}R^{\mu}{}_{\beta}{}^{\rho\lambda} + 2R_{\mu\nu\rho\alpha}R^{\mu\nu\rho}{}_{\beta} - R_{\alpha\nu\rho\lambda}R_{\beta}{}^{\nu\rho\lambda} - \frac{1}{2}g_{\alpha\beta}R_{\mu\nu\rho\lambda}R^{\mu\nu\rho\lambda} = 0. \tag{B.21}$$

$$4\nabla_{\lambda}(\sqrt{-g}R_{\sigma}{}^{\mu\nu\lambda}) + 2\sqrt{-g}\left[2R_{\sigma}{}^{\mu\lambda\nu}T^{\rho}{}_{\rho\lambda} + R_{\sigma}{}^{\mu\lambda\rho}T^{\nu}{}_{\lambda\rho}\right] = 0. \tag{B.22}$$

- $\mathcal{L} = R_{\mu\nu\rho\lambda}R^{\nu\mu\rho\lambda}$

$$2R_{\mu\nu\alpha\lambda}R^{\nu\mu}{}_{\beta}{}^{\lambda} - \frac{1}{2}g_{\alpha\beta}R_{\mu\nu\rho\lambda}R^{\nu\mu\rho\lambda} = 0. \tag{B.23}$$

$$4\nabla_{\lambda}(\sqrt{-g}R^{\mu}{}_{\sigma}{}^{\nu\lambda}) + 2\sqrt{-g}\left[2R^{\mu}{}_{\sigma}{}^{\lambda\nu}T^{\rho}{}_{\rho\lambda} + R^{\mu}{}_{\sigma}{}^{\lambda\rho}T^{\nu}{}_{\lambda\rho}\right] = 0. \tag{B.24}$$

- $\mathcal{L} = R_{\mu\nu\rho\lambda}R^{\rho\lambda\mu\nu}$

$$R^{\mu\nu}{}_{\rho\alpha}R^{\rho}{}_{\beta\mu\nu} + R^{\mu\nu}{}_{\rho\beta}R^{\rho}{}_{\alpha\mu\nu} + -\frac{1}{2}g_{\alpha\beta}R_{\mu\nu\rho\lambda}R^{\rho\lambda\mu\nu} = 0. \tag{B.25}$$

$$\begin{aligned}
&2\nabla_{\lambda}(\sqrt{-g}R^{\nu\lambda}{}_{\sigma}{}^{\mu}) - 2\nabla_{\lambda}(\sqrt{-g}R^{\lambda\nu}{}_{\sigma}{}^{\mu}) \\
&+ 2\sqrt{-g}\left[R^{\lambda\nu}{}_{\sigma}{}^{\mu}T^{\rho}{}_{\rho\lambda} - R^{\nu\lambda}{}_{\sigma}{}^{\mu}T^{\rho}{}_{\rho\lambda} + R^{\lambda\rho}{}_{\sigma}{}^{\mu}T^{\nu}{}_{\lambda\rho}\right] = 0.
\end{aligned} \tag{B.26}$$

- $\mathcal{L} = R_{\mu\nu\rho\lambda}R^{\rho\nu\mu\lambda}$

$$R^{\mu}{}_{\beta\nu}{}^{\lambda}R^{\nu}{}_{\alpha\mu\lambda} + R^{\mu\nu}{}_{\rho\beta}R^{\rho}{}_{\nu\mu\alpha} + -\frac{1}{2}g_{\alpha\beta}R_{\mu\nu\rho\lambda}R^{\rho\nu\mu\lambda} = 0. \tag{B.27}$$

$$\begin{aligned}
&2\nabla_{\lambda}(\sqrt{-g}R^{\nu\mu}{}_{\sigma}{}^{\lambda}) - 2\nabla_{\lambda}(\sqrt{-g}R^{\lambda\mu}{}_{\sigma}{}^{\nu}) \\
&+ 2\sqrt{-g}\left[R^{\lambda\mu}{}_{\sigma}{}^{\nu}T^{\rho}{}_{\rho\lambda} - R^{\nu\mu}{}_{\sigma}{}^{\lambda}T^{\rho}{}_{\rho\lambda} + R^{\rho\mu}{}_{\sigma}{}^{\lambda}T^{\nu}{}_{\lambda\rho}\right] = 0.
\end{aligned} \tag{B.28}$$

- $\mathcal{L} = R_{\mu\nu\rho\lambda}R^{\nu\rho\mu\lambda}$

$$R^{\mu}{}_{\nu\beta}{}^{\lambda}R^{\nu}{}_{\alpha\mu\lambda} + R^{\mu}{}_{\nu}{}^{\lambda}{}_{\beta}R^{\nu}{}_{\lambda\mu\alpha} - \frac{1}{2}g_{\alpha\beta}R_{\mu\nu\rho\lambda}R^{\nu\rho\mu\lambda} = 0. \tag{B.29}$$

$$\begin{aligned}
&\nabla_{\lambda}(\sqrt{-g}R^{\nu}{}_{\sigma}{}^{\mu\lambda}) - \nabla_{\lambda}(\sqrt{-g}R^{\lambda}{}_{\sigma}{}^{\mu\nu}) \\
&+ \nabla_{\lambda}(\sqrt{-g}R^{\nu\mu}{}_{\sigma}{}^{\lambda}) - \nabla_{\lambda}(\sqrt{-g}R^{\mu\lambda}{}_{\sigma}{}^{\nu}) \\
&+ \sqrt{-g}\left[R^{\lambda}{}_{\sigma}{}^{\mu\nu}T^{\rho}{}_{\rho\lambda} - R^{\nu}{}_{\sigma}{}^{\mu\lambda}T^{\rho}{}_{\rho\lambda} + R^{\lambda}{}_{\sigma}{}^{\mu\rho}T^{\nu}{}_{\lambda\rho}\right] \\
&+ \sqrt{-g}\left[R^{\mu\lambda}{}_{\sigma}{}^{\nu}T^{\rho}{}_{\rho\lambda} - R^{\nu\mu}{}_{\sigma}{}^{\lambda}T^{\rho}{}_{\rho\lambda} + R^{\lambda\mu}{}_{\sigma}{}^{\rho}T^{\nu}{}_{\lambda\rho}\right] = 0.
\end{aligned} \tag{B.30}$$

- $\mathcal{L} = R_{\mu\nu\rho\lambda}R^{\mu\rho\nu\lambda}$

$$2R_{\mu\nu(\beta|\lambda}R^{\mu}{}_{\alpha)}{}^{\nu\lambda} + R_{\lambda\mu\nu\beta}R^{\lambda\nu\mu}{}_{\alpha} - R_{\beta\mu\nu\lambda}R_{\alpha}{}^{\nu\mu\lambda} - \frac{1}{2}g_{\alpha\beta}R_{\mu\nu\rho\lambda}R^{\mu\rho\nu\lambda} = 0. \tag{B.31}$$

$$\begin{aligned}
&2\nabla_{\lambda}(\sqrt{-g}R_{\sigma}{}^{\nu\mu\lambda}) - 2\nabla_{\lambda}(\sqrt{-g}R_{\sigma}{}^{\lambda\mu\nu}) \\
&+ 2\sqrt{-g}\left[R_{\sigma}{}^{\lambda\mu\nu}T^{\rho}{}_{\rho\lambda} - R_{\sigma}{}^{\nu\mu\lambda}T^{\rho}{}_{\rho\lambda} + R_{\sigma}{}^{\lambda\mu\rho}T^{\nu}{}_{\lambda\rho}\right] = 0.
\end{aligned} \tag{B.32}$$

Appendix C

Details of the series solution

C.1 Variation of transformed action

Here the variation of the Legendre transformed action (6.5), where the auxillary field $\Sigma_{\mu\nu}$ is written in terms of $q^{\mu\nu}$ and $g^{\mu\nu}$ is performed with respect to the metric $g_{\mu\nu}$.

The part of the Lagrangian density depending on $g_{\mu\nu}$ is

$$\sqrt{-g}\mathcal{L}_g \equiv \mathfrak{L}_g = -\sqrt{-g}\left[\frac{\partial F}{\partial \Sigma_{\mu\nu}}\Sigma_{\mu\nu}(q,g) - F(q,g) - \mathcal{L}_m(g,\chi)\right]. \tag{C.1}$$

Written out in terms of $\Sigma_{\mu\nu}$

$$\mathfrak{L}_g = -\sqrt{-g}\left[\beta\Sigma^2 + \alpha\Sigma_{\lambda\sigma}\Sigma^{\lambda\sigma} + U(\chi) + \frac{1}{2}g^{\lambda\sigma}\partial_\lambda\chi\partial_\sigma\chi\right], \tag{C.2}$$

where $\Sigma \equiv g^{\mu\nu}\Sigma_{\mu\nu}$. The variation is

$$
\begin{aligned}
\delta\mathfrak{L}_g = &-\sqrt{-g}\left[2\beta\Sigma\delta\Sigma + 2\alpha g^{\lambda\gamma}\Sigma_{\lambda\sigma}\delta(g^{\rho\sigma}\Sigma_{\gamma\rho}) + \frac{1}{2}\partial_\lambda\chi\partial_\sigma\chi\delta g^{\lambda\sigma}\right] \\
&+ \frac{1}{2}\sqrt{-g}g^{\mu\nu}\mathcal{L}_g\delta g_{\mu\nu}
\end{aligned} \tag{C.3}
$$

First looking at $\delta\Sigma$ separately

$$\delta\Sigma = \delta\left(\frac{-4}{2\alpha+8\beta} + \frac{1}{2\alpha+8\beta}\frac{\sqrt{-q}}{\sqrt{-g}}q^{\sigma\lambda}g_{\sigma\lambda}\right), \tag{C.4}$$

with the identity $\delta\sqrt{-g} = 1/2\sqrt{-g}g^{\mu\nu}\delta g_{\mu\nu}$ we have

$$\delta\Sigma = \frac{1}{2\alpha+8\beta}\frac{\sqrt{-q}}{\sqrt{-g}}\left(q^{\mu\nu} - \frac{1}{2}q^{\sigma\lambda}g_{\sigma\lambda}g^{\mu\nu}\right)\delta g_{\mu\nu}. \tag{C.5}$$

Now separately looking at $\delta\Sigma_{\mu\nu}$

$$\delta\Sigma_{\mu\nu} = \delta\left[\frac{1}{2\alpha}\frac{\sqrt{-q}}{\sqrt{-q}}q^{\sigma\lambda}g_{\sigma\mu}g_{\lambda\nu} - \frac{1}{2\alpha+8\beta}\left(1 + \frac{\beta}{\alpha}\frac{\sqrt{-q}}{\sqrt{-q}}q^{\sigma\lambda}g_{\sigma\lambda}\right)g_{\mu\nu}\right], \tag{C.6}$$

which becomes

$$
\begin{aligned}
\delta\Sigma_{\mu\nu} = &-\frac{\delta g_{\mu\nu}}{2\alpha+8\beta} + \frac{\sqrt{-q}}{\sqrt{-g}}\left[-\frac{1}{4\alpha}q^{\sigma\lambda}g_{\sigma\mu}g_{\lambda\nu}g^{\rho\gamma}\delta g_{\rho\gamma} + \frac{1}{\alpha}q^{\sigma\lambda}g_{\sigma\mu}\delta g_{\lambda\nu}\right. \\
&\left. -\frac{\beta}{\alpha(2\alpha+8\beta)}\left(q^{\lambda\sigma}g_{\lambda\sigma}\delta g_{\mu\nu} - \frac{1}{2}q^{\sigma\lambda}g_{\sigma\lambda}g_{\mu\nu}g^{\rho\gamma}\delta g_{\rho\gamma} + q^{\sigma\lambda}g_{\mu\nu}\delta g_{\sigma\lambda}\right)\right].
\end{aligned} \tag{C.7}
$$

With these out of the way it is just a question of plugging these in to (C.3) and simplifying the expression. In the end we get

$$
\begin{aligned}
\delta\mathfrak{L}_g = \sqrt{-g}\Bigg\{ & -\frac{1}{4(\alpha+4\beta)}\frac{\sqrt{-q}}{\sqrt{-g}}q^{\sigma\lambda}g_{\sigma\mu}g_{\lambda\nu} \\
& +\frac{1}{4\alpha}\frac{q}{g}\left[q^{\sigma\lambda}q^{\rho\delta}g_{\lambda\delta}g_{\rho\nu}g_{\sigma\mu}-\frac{\beta}{(\alpha+4\beta)}q^{\delta\rho}g_{\delta\rho}q^{\sigma\lambda}g_{\sigma\mu}g_{\lambda\nu}\right] \\
& +\frac{1}{2}g_{\mu\nu}\left[\frac{1}{\alpha+4\beta}\left(\frac{\beta}{8\alpha}\frac{q}{g}q^{\lambda\sigma}g_{\lambda\sigma}q^{\rho\delta}g_{\rho\delta}\right)-\frac{q}{g}\frac{1}{8\alpha}q^{\lambda\sigma}q^{\delta\rho}g_{\lambda\delta}g_{\sigma\rho}\right] \\
& -\frac{1}{2}\partial_\mu\chi\partial_\nu\chi+\frac{1}{2}g_{\mu\nu}\left[U(\chi)+\frac{1}{2}g^{\lambda\sigma}\partial_\lambda\chi\partial_\sigma\chi\right]\Bigg\}\delta g^{\mu\nu}.
\end{aligned}
\tag{C.8}
$$

C.2 Series solution of the modified scalar sector

Here a few coefficients for the approximate solution (6.46) of the disformal transformation (6.17) parameters and the scalar sector with a few more terms are presented.

$$
a_0 = \frac{1}{1+(2\alpha+8\beta)U(\chi)}
$$

$$
b_0 = \frac{2\alpha}{[1+(2\alpha+8\beta)U(\chi)]\,[1+(4\alpha+8\beta)U(\chi)]}
$$

$$
a_1 = -\frac{\alpha+2\beta}{1+(4\alpha+8\beta)U}
$$

$$
b_1 = \frac{4\alpha\left[-2(\alpha+\beta)-(3\alpha+4\beta)(3\alpha+8\beta)U-8(\alpha+2\beta)^2(\alpha+4\beta)U^2\right]}{(1+2\alpha U+8\beta U)(1+4\alpha U+8\beta U)^3}
$$

$$
a_2 = \frac{\alpha\left[7\alpha+10\beta+4(3\alpha+5\beta)(3\alpha+8\beta)U+40(\alpha+2\beta)^2(\alpha+4\beta)U^2\right]}{2(1+2\alpha U+8\beta U)(1+4\alpha U+8\beta U)^3}
$$

$$
\begin{aligned}
b_2 = {} & \frac{2\alpha}{(1+2\alpha U+8\beta U)(1+4\alpha U+8\beta U)^5} \\
& \times\Bigg[\,64\alpha^6 U^4+4\beta^2(1+8\beta U)^4+128\alpha^5 U^3(3+8\beta U) \\
& +5\alpha\beta(1+8\beta U)^3(5+16\beta U)+2\alpha^3 U(1+8\beta U)(85+128\beta U(7+11\beta U)) \\
& +4\alpha^4 U^2(111+8\beta U(131+208\beta U))+(\alpha+8\alpha\beta U)^2(21+2\beta U(173+328\beta U))\,\Bigg]
\end{aligned}
\tag{C.9}
$$

62

The scalar sector with few more terms

$$
\begin{aligned}
\mathcal{L}_m = & - \frac{U}{1 + (2\alpha + 8\beta)U} - \frac{\frac{1}{2}X_q}{1 + (2\alpha + 8\beta)U} \\
& + \frac{\alpha^2 U + \beta(1 + 8\beta U) + \alpha(1 + 6\beta U)}{2\left[1 + (2\alpha + 8\beta)U\right]\left[1 + (4\alpha + 8\beta)U\right]}X_q^2 \\
& - \frac{\alpha\left[12\alpha^3 U^2 + 3\beta(1 + 8\beta U)^2 + 2\alpha^2 U(7 + 48\beta U) + 3\alpha(1 + 18\beta U + 80\beta^2 U^2)\right]}{2(1 + 2\alpha U + 8\beta U)(1 + 4\alpha U + 8\beta U)^3}X_q^3 \\
& + \frac{3\alpha}{8(1 + 2\alpha U + 8\beta U)(1 + 4\alpha U + 8\beta U)^5}X_q^4 \left[64\alpha^6 U^4 \dots \right.
\end{aligned}
\tag{C.10}
$$

Bibliography

[1] S. M. Carroll, *Spacetime and geometry: An introduction to general relativity.* 2004.

[2] J. Martin, C. Ringeval and V. Vennin, *Encyclopædia Inflationaris*, *Phys. Dark Univ.* **5-6** (2014) 75–235, [`1303.3787`].

[3] M. Ferraris, M. Francaviglia and C. Reina, *Variational formulation of general relativity from 1915 to 1925 "Palatini's method" discovered by Einstein in 1925*, *General Relativity and Gravitation* **14** (Mar., 1982) 243–254.

[4] G. W. Gibbons and S. W. Hawking, *Action integrals and partition functions in quantum gravity*, *Phys. Rev. D* **15** (May, 1977) 2752–2756.

[5] J. W. York, *Role of conformal three-geometry in the dynamics of gravitation*, *Phys. Rev. Lett.* **28** (Apr, 1972) 1082–1085.

[6] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2011, 10.1017/CBO9780511524646.

[7] A. Guarnizo, L. Castaneda and J. M. Tejeiro, *Boundary Term in Metric f(R) Gravity: Field Equations in the Metric Formalism*, *Gen. Rel. Grav.* **42** (2010) 2713–2728, [`1002.0617`].

[8] R. M. Wald, *General Relativity.* Chicago Univ. Pr., Chicago, USA, 1984, 10.7208/chicago/9780226870373.001.0001.

[9] E. Dyer and K. Hinterbichler, *Boundary Terms, Variational Principles and Higher Derivative Modified Gravity*, *Phys. Rev.* **D79** (2009) 024028, [`0809.4033`].

[10] H. Burton and R. B. Mann, *Palatini variational principle for an extended Einstein-Hilbert action*, *Phys. Rev.* **D57** (1998) 4754–4759, [`gr-qc/9711003`].

[11] V. Vitagliano, T. P. Sotiriou and S. Liberati, *The dynamics of metric-affine gravity*, *Annals Phys.* **326** (2011) 1259–1273, [`1008.0171`].

[12] J. A. Schouten, *Ricci-Calculus.* Springer-Verlag Berlin Heidelberg, 1954, 10.1007/978-3-662-12927-2.

[13] F. W. Hehl, P. von der Heyde, G. D. Kerlick and J. M. Nester, *General relativity with spin and torsion: Foundations and prospects*, *Rev. Mod. Phys.* **48** (Jul, 1976) 393–416.

[14] N. Dadhich and J. M. Pons, *On the equivalence of the Einstein-Hilbert and the Einstein-Palatini formulations of general relativity for an arbitrary connection*, *Gen. Rel. Grav.* **44** (2012) 2337–2352, [`1010.0869`].

[15] F. W. Hehl and G. D. Kerlick, *Metric-affine variational principles in general relativity. I - Riemannian space-time*, *General Relativity and Gravitation* **9** (Aug., 1978) 691–710.

[16] A. N. Bernal, B. Janssen, A. Jimenez-Cano, J. A. Orejuela, M. Sanchez and P. Sanchez-Moreno, *On the (non-)uniqueness of the Levi-Civita solution in the Einstein–Hilbert–Palatini formalism*, *Phys. Lett.* **B768** (2017) 280–287, [`1606.08756`].

[17] T. P. Sotiriou and S. Liberati, *Metric-affine f(R) theories of gravity*, *Annals Phys.* **322** (2007) 935–966, [`gr-qc/0604006`].

[18] E. Schrödinger, *Space-Time Structure*. Cambridge Science Classics. Cambridge University Press, 1985, 10.1017/CBO9780511586446.

[19] T. P. Sotiriou and V. Faraoni, *f(R) Theories Of Gravity*, *Rev. Mod. Phys.* **82** (2010) 451–497, [`0805.1726`].

[20] D. H. Lyth and A. R. Liddle, *The primordial density perturbation: Cosmology, inflation and the origin of structure*. 2009.

[21] D. Baumann, *Inflation*, in *Physics of the large and the small, TASI 09, proceedings of the Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, Colorado, USA, 1-26 June 2009*, pp. 523–686, 2011, `0907.5424`, DOI.

[22] S. Weinberg, *Cosmology*. 2008.

[23] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space*. Cambridge Monographs on Mathematical Physics. Cambridge Univ. Press, Cambridge, UK, 1984, 10.1017/CBO9780511622632.

[24] D. H. Lyth, K. A. Malik and M. Sasaki, *A General proof of the conservation of the curvature perturbation*, *JCAP* **0505** (2005) 004, [`astro-ph/0411220`].

[25] A. Albrecht, P. Ferreira, M. Joyce and T. Prokopec, *Inflation and squeezed quantum states*, *Phys. Rev.* **D50** (1994) 4807–4820, [`astro-ph/9303001`].

[26] V. F. Mukhanov, *Quantum Theory of Gauge Invariant Cosmological Perturbations*, *Sov. Phys. JETP* **67** (1988) 1297–1302.

[27] G. Isidori, V. S. Rychkov, A. Strumia and N. Tetradis, *Gravitational corrections to standard model vacuum decay*, *Phys. Rev.* **D77** (2008) 025034, [`0712.0242`].

[28] Y. Hamada, H. Kawai and K.-y. Oda, *Minimal Higgs inflation*, *PTEP* **2014** (2014) 023B02, [1308.6651].

[29] M. Fairbairn, P. Grothaus and R. Hogan, *The Problem with False Vacuum Higgs Inflation*, *JCAP* **1406** (2014) 039, [1403.7483].

[30] E. E. Flanagan, *The Conformal frame freedom in theories of gravitation*, *Class. Quant. Grav.* **21** (2004) 3817, [gr-qc0403063].

[31] S. Tsujikawa and B. Gumjudpai, *Density perturbations in generalized Einstein scenarios and constraints on nonminimal couplings from the Cosmic Microwave Background*, *Phys. Rev.* **D69** (2004) 123523, [astro-ph/0402185].

[32] M. Postma and M. Volponi, *Equivalence of the Einstein and Jordan frames*, *Phys. Rev.* **D90** (2014) 103516, [1407.6874].

[33] J. Rubio, *Higgs inflation*, *Front. Astron. Space Sci.* **5** (2019) 50, [1807.02376].

[34] F. L. Bezrukov and M. Shaposhnikov, *The Standard Model Higgs boson as the inflaton*, *Phys. Lett.* **B659** (2008) 703–706, [0710.3755].

[35] F. Bezrukov, A. Magnin, M. Shaposhnikov and S. Sibiryakov, *Higgs inflation: consistency and generalisations*, *JHEP* **01** (2011) 016, [1008.5157].

[36] J. Garcia-Bellido, D. G. Figueroa and J. Rubio, *Preheating in the Standard Model with the Higgs-Inflaton coupled to gravity*, *Phys. Rev.* **D79** (2009) 063531, [0812.4624].

[37] J. Repond and J. Rubio, *Combined Preheating on the lattice with applications to Higgs inflation*, *JCAP* **1607** (2016) 043, [1604.08238].

[38] F. Bezrukov, D. Gorbunov and M. Shaposhnikov, *On initial conditions for the Hot Big Bang*, *JCAP* **0906** (2009) 029, [0812.3622].

[39] Planck collaboration, Y. Akrami et al., *Planck 2018 results. X. Constraints on inflation*, 1807.06211.

[40] M. Atkins and X. Calmet, *Bounds on the Nonminimal Coupling of the Higgs Boson to Gravity*, *Phys. Rev. Lett.* **110** (2013) 051301, [1211.0281].

[41] M. D. Schwartz, *Quantum Field Theory and the Standard Model.* Cambridge University Press, 2014.

[42] J. L. F. Barbon and J. R. Espinosa, *On the Naturalness of Higgs Inflation*, *Phys. Rev.* **D79** (2009) 081302, [0903.0355].

[43] F. Bauer and D. A. Demir, *Inflation with Non-Minimal Coupling: Metric versus Palatini Formulations*, *Phys. Lett.* **B665** (2008) 222–226, [0803.2664].

[44] J. Rubio and E. S. Tomberg, *Preheating in Palatini Higgs inflation*, *JCAP* **1904** (2019) 021, [1902.10148].

[45] F. Bauer and D. A. Demir, *Higgs-Palatini Inflation and Unitarity*, *Phys. Lett.* **B698** (2011) 425–429, [1012.2900].

[46] T. Clifton, P. G. Ferreira, A. Padilla and C. Skordis, *Modified Gravity and Cosmology*, *Phys. Rept.* **513** (2012) 1–189, [1106.2476].

[47] R. Utiyama and B. S. DeWitt, *Renormalization of a classical gravitational field interacting with quantized matter fields*, *Journal of Mathematical Physics* **3** (1962) 608–618, [https://doi.org/10.1063/1.1724264].

[48] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory.* Addison-Wesley, Reading, USA, 1995.

[49] K. S. Stelle, *Renormalization of Higher Derivative Quantum Gravity*, *Phys. Rev.* **D16** (1977) 953–969.

[50] D. G. Boulware and S. Deser, *String Generated Gravity Models*, *Phys. Rev. Lett.* **55** (1985) 2656.

[51] S. Deser, *Gravity from strings*, *Physica Scripta* **T15** (jan, 1987) 138–142.

[52] D. Tong, *String Theory*, 0908.0333.

[53] R. P. Woodard, *How Far Are We from the Quantum Theory of Gravity?*, *Rept. Prog. Phys.* **72** (2009) 126002, [0907.4238].

[54] L. Parker and J. Z. Simon, *Einstein equation with quantum corrections reduced to second order*, *Phys. Rev.* **D47** (1993) 1339–1355, [gr-qc/9211002].

[55] J. Beltrán Jiménez and A. Delhom, *Ghosts in metric-affine higher order curvature gravity*, 1901.08988.

[56] T. P. Sotiriou, *f(R) gravity and scalar-tensor theory*, *Class. Quant. Grav.* **23** (2006) 5117–5128, [gr-qc/0604028].

[57] C. Lanczos, *A remarkable property of the riemann-christoffel tensor in four dimensions*, *Annals of Mathematics* **39** (1938) 842–850.

[58] E. Alvarez and S. Gonzalez-Martin, *Weyl Gravity Revisited*, *JCAP* **1702** (2017) 011, [1610.03539].

[59] A. Salvio, *Quadratic Gravity*, *Front.in Phys.* **6** (2018) 77, [1804.09944].

[60] K. S. Stelle, *Classical gravity with higher derivatives*, *General Relativity and Gravitation* **9** (Apr, 1978) 353–371.

[61] D. Lovelock, *The uniqueness of the einstein field equations in a four-dimensional space*, *Archive for Rational Mechanics and Analysis* **33** (Jan, 1969) 54–70.

[62] D. Lovelock, *The einstein tensor and its generalizations*, *Journal of Mathematical Physics* **12** (1971) 498–501, [https://doi.org/10.1063/1.1665613].

[63] M. Borunda, B. Janssen and M. Bastero-Gil, *Palatini versus metric formulation in higher curvature gravity*, *JCAP* **0811** (2008) 008, [0804.4440].

[64] E. Alvarez, J. Anero, S. Gonzalez-Martin and R. Santos-Garcia, *Physical content of Quadratic Gravity*, *Eur. Phys. J.* **C78** (2018) 794, [1802.05922].

[65] L. Alvarez-Gaume, A. Kehagias, C. Kounnas, D. Lüst and A. Riotto, *Aspects of Quadratic Gravity*, *Fortsch. Phys.* **64** (2016) 176–189, [1505.07657].

[66] V. I. Afonso, C. Bejarano, J. Beltran Jimenez, G. J. Olmo and E. Orazi, *The trivial role of torsion in projective invariant theories of gravity with non-minimally coupled matter fields*, *Class. Quant. Grav.* **34** (2017) 235003, [1705.03806].

[67] J. W. Moffat, *New theory of gravitation*, *Phys. Rev. D* **19** (Jun, 1979) 3554–3558.

[68] T. Damour, S. Deser and J. G. McCarthy, *Nonsymmetric gravity theories: Inconsistencies and a cure*, *Phys. Rev.* **D47** (1993) 1541–1556, [gr-qc/9207003].

[69] G. Domènech, A. Naruko and M. Sasaki, *Cosmological disformal invariance*, *JCAP* **1510** (2015) 067, [1505.00174].

[70] K. Takahashi, H. Motohashi, T. Suyama and T. Kobayashi, *General invertible transformation and physical degrees of freedom*, *Phys. Rev.* **D95** (2017) 084053, [1702.01849].

[71] M. Zumalacárregui and J. García-Bellido, *Transforming gravity: from derivative couplings to matter to second-order scalar-tensor theories beyond the Horndeski Lagrangian*, *Phys. Rev.* **D89** (2014) 064046, [1308.4685].

[72] H. Motohashi and J. White, *Disformal invariance of curvature perturbation*, *JCAP* **1602** (2016) 065, [1504.00846].

[73] Y. Watanabe, A. Naruko and M. Sasaki, *Multi-disformal invariance of non-linear primordial perturbations*, *EPL* **111** (2015) 39002, [1504.00672].

[74] S. Tsujikawa, *Disformal invariance of cosmological perturbations in a generalized class of Horndeski theories*, *JCAP* **1504** (2015) 043, [1412.6210].

[75] M. Minamitsuji, *Disformal transformation of cosmological perturbations*, *Phys. Lett.* **B737** (2014) 139–150, [1409.1566].

[76] V.-M. Enckell, K. Enqvist, S. Rasanen and L.-P. Wahlman, *Inflation with R^2 term in the Palatini formalism*, 1810.05536.

[77] V.-M. Enckell, K. Enqvist, S. Rasanen and L.-P. Wahlman, *Higgs-R^2 inflation - full slow-roll study at tree-level*, 1812.08754.

[78] H. Sugai, P. A. R. Ade, Y. Akiba, D. Alonso, K. Arnold, J. Aumont et al., *Updated design of the cmb polarization experiment satellite litebird*, 2020.