Constrained Dynamics and Higher Derivative Systems in Modified Gravity

Tai-jun Chen

Department of Applied Mathematics and Theoretical Physics

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

Girton College May 2015

This thesis is dedicated to my beloved wife, Yu-ning Lee, who has supported me for the past twelve years. Her hours of work in supporting our family enabled my research and contemplation.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text. This dissertation contains fewer than 65,000 words, including the appendices, bibliography, footnotes, tables and equations, and has fewer than 150 figures.

Tai-jun Chen
May 2015

Acknowledgements

I wish to thank my supervisor, Eugene A. Lim, for his guidance and support throughout my PhD. I am grateful for his advice not only on research but also on how to improve my professionalism. When my research appeared too solitary, he encouraged me to move outside my comfort zone and helped me to build academic connections with his collaborators, students and friends. Eugene also gave me advice when I made decisions, when I was unsure how to react professionally and when I felt depressed. I am deeply indebted to him.

I also wish to thank my advisor, Paul Shellard, and the other faculty members of the cosmology group, including Daniel Baumann, Anne Davis, John Barrow, and Xingang Chen, for their advice and professionalism, which inspired me and made me a better researcher. I especially wish to thank Daniel for his help with my fourth term interview. I am deeply grateful to my collaborators, Eugene A. Lim, Andrew Tolley and Matteo Fasiello, for their interesting input and discussions about this dissertation.

For stimulating discussions about physics I wish to thank: Mustafa A. Amin, C. Armendariz-Picon, Sean Carroll, Pisin Chen, Jiunn-Wei Chen, Richard Easther, Pau Figueras, Kurt Hinterbichler, Pei-ming Ho, Wayne Hu, Kei-ichi Maeda, Shinji Mukohyama, Ignacy Sawicki, Thomas Sotiriou, Alexander Vikman, and Helvi Witek.

Thanks also go to my friends at DAMTP: Valentin Assassi, Chris Lau, Jeremy Sakiestein, Marcel Schmittfull, Simon Su, and Kei Yamamoto. I especially wish to thank Chris and Simon for the useful, interesting discussions on physics, politics, and many other issues.

I also wish to thank my friends, who enriched my life in Cambridge, especially Yu-ning, Borkae, Eddy, Jason, Candy, Jhen-Si, Nai-Chieh, Alva, Lily, Gareth, Chiachi, and Jinco.

Finally, I am deeply indebted to my family for their great support, encouragement and love throughout my studies, which gave me the time to pursue my research.

Abstract

In this thesis, higher derivative theories and constrained dynamics are investigated in detail. In the first part of the thesis, we discuss how the Ostrogradski instability emerges in non-degenerate higher derivative theories in the context of a one-dimensional point particle where the position of the particle is a function only dependent on time. We show that the instabilities can only be removed by the addition of constraints if the original theory's phase space is reduced. We then generalize this formalism to the most general higher derivative gravity theory where the action is not only linearly dependent on the Ricci scalar but also the quadratic curvature invariants in four-dimensional spacetime. We find that the instabilities can be removed by the judicious addition of constraints at the quadratic level of metric fluctuations around Minkowski and de Sitter backgrounds while the dimensionality of the original phase space is reduced. The constrained higher derivative gravity theory is ghost free as well as preserves the renormalization properties of higher derivative gravity, at the price of giving up the Lorentz invariance. In the second part of the thesis, we study the spherically symmetric static solution of a class of two scalar-field theory, where one of them is a Lagrange multiplier enforcing a constraint relating the value of the other scalar field to the norm of its derivative. We find the spherically symmetric static solution of the theory with an exponential potential. However, when we investigate the stability issue of the solution, the perturbation with the odd type symmetry is stable, while the even modes always contain one ghostlike degree of freedom.

Contents

Chapter 1

Introduction

Since Newton decided to write the second law of motion as $F = m\ddot{q}$, all theories of fundamental physics are based on equations of motion with, at most, second order time derivatives. In the language of Lagrangian formalism, this means that there is no more than second order time derivative terms in the Lagrangian which cannot be removed with integration by parts. The reason why the higher derivative theories are not adopted is because of the accompanying instability, which can be summarized by Ostrogradski's theorem: "Any *non-degenerate*[1] theory whose dynamical variable is higher than second order in time derivative there exist linearly unstable degrees of freedom." [1]

If we consider a theory based on an equation of motion in fourth order time derivatives,[2] since we need four initial conditions to solve the equation, the phase space in the Hamiltonian formalism is four-dimensional. Ostrogradski's theorem states that the extra degree of freedom is a linearly unstable one and the Hamiltonian of the theory is not bounded below along the direction of the extra dimension of the phase space. The theorem is very powerful and can be easily applied to all non-degenerate higher derivative theories.

Even though the powerful Ostrogradski's theorem exists, higher derivative theories have been studied in an attempt to modify the fundamental theories in order to render them compatible with the phenomenology or invent theories with better characteristics. An old example of higher derivative theory is the Abraham–Lorentz equation of motion, where the energy of the charged particle during acceleration is dependent on the time derivative of its acceleration. The equation is later stated in the context of generalized classical electrodynamics and again derived by Fokker, Feymann, and Wheeler [2].

[1]Non degeneracy is a technical term states that the highest time derivative of the variable can be written as a function of canonical coordinates and momenta, we will discuss it later.

[2]In order to have an integer number of degrees of freedom, the number of initial conditions of an equation of motion must be even. We will illustrate this in Chapter 2.

There are also several higher derivative modifications of the theory of gravity. The simplest modification is the $f(R)$ gravity [3–5], which is a higher derivative extension of general relativity and does *not* suffer from the Ostrogradski instability because of its degeneracy. The instability in $f(R)$ gravity is a gauge degree of freedom and the physical Hamiltonian is bounded from below if we choose a suitable gauge. However, the general higher derivative modification of general relativity, such as Weyl gravity or Stelle's higher derivative theory, admits universal instability, as the Ostrogradski instability is omnipresent.

On the other hand, the dynamics of the constrained systems have been studied for more than 150 years because of the considerable development of gauge theories. The Dirac analysis is now viewed as the standard formalism to investigate the constrained dynamical system [6], encoding the information about the constraints into the generalized Poisson bracket, and one can treat the theory as unconstrained once the new bracket (Dirac bracket) is found. The Dirac analysis of the constrained system is the main methodology employed in this thesis, which will be discussed in Section 1.2.

One should note that, even though the Ostrogradski instability is universal, there is an important condition for its existence — *non-degeneracy*. This is the blind spot on which we are going to focus. In this thesis, we will demonstrate a standard formalism to introduce constraints for non-degenerate higher derivative theories in such a way that the unstable degrees of freedom are removed while the healthy degrees of freedom as well as the positive characteristics of the higher derivative theories are retained.

In this chapter, we will first review the higher derivative gravitational theories in Section 1.1. We will discuss the motivation of modifying gravity, the $f(R)$ gravity, and the general higher derivative gravity up to the quadratic level in curvature invariants in four-dimensional spacetime. We then review in detail the Dirac analysis of constrained Hamiltonian systems in Section 1.2, introducing all of the technical terms and the analysis we will employ in this thesis. We close this chapter by presenting the thesis outline.

1.1 Higher derivative gravity

The general theory of relativity and quantum field theory are usually viewed as the most important accomplishments in the field of fundamental theoretical physics in the 20th century. While the quantization of gauge field theories was understood in the 1970s, the general theory of relativity was first proposed by Albert Einstein in 1915 and, 100 years later, it still amazingly retains its original form without any modification, i.e., the field equation written by Einstein

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{1.1}$$

still provides the best model for describing spacetime on a large scale [7–9]. In eq. (1.1), $G_{\mu\nu}$ is Einstein's tensor, $T_{\mu\nu}$ is the stress-energy tensor, G is Newton's constant, and c is the speed of light. Through using this equation, we can study the evolution of our Universe, the behavior of black holes, all of the structure in the Universe from the solar system up to super clusters of galaxies, and the gravitational waves produced by the Big Bang. By comparing simulations with experimental data, we find that the general theory of relativity is supported from the solar system to binary pulsars observational data with an extremely high degree of accuracy.

Even though the general theory of relativity is so powerful at the macroscopic scale, many alternative theories are still being investigated. A famous example of the modification of the general theory of relativity is conformal gravity [10–12], where the action is invariant under the Weyl transformation. Based on Weyl's work, Sakharov proposed a theory of modified gravity, whereby the Einstein-Hilbert action — from which Einstein's field equation can be derived — is the first approximation of a more complicated fundamental action [13]. In his approach, the perturbations of spacetime lead to corrections to the Einstein-Hilbert action, and the corrections refer in general to the higher power of the curvature or to higher order derivatives. In 1977, Kellogg Stelle [14, 15] proved that the theory is renormalizable at the one-loop level but suffers from the universal Ostrogradski instability. Due to this undesirable characteristic, people usually study the special class of this theory where the Ostrogradski instability is protected by gauge symmetry, i.e. $f(R)$ gravity.

One main reason why we investigate the modification of gravity is the "dark universe scenario". If the general theory of relativity governs the evolution of our Universe, there must be a substantial amount of dark energy in our Universe which causes the apparent accelerating expansion of the Universe. On the other hand, more than 30 years of data show that there should be dark matter haloes surrounding the galaxies in the Universe. It seems that more than 95% of the Universe consists of energy densities which do not emit radiation [16], which is very strange for the human beings living on earth. This weird energy composition suggests the possibility that, at the largest scale, general relativity might not be the final theory which governs the evolution of the Universe. In order to develop a better understanding of the dark universe, we might need theories of modified gravity. For example, in projectable Hořava–Lifshitz gravity, the dark matter component can be related appears as an integration constant [80].

In this section, we briefly review a class of modification of general relativity — higher derivative gravity. We start by introducing the conventions and notations we use in this thesis in Section 1.1.1, and introduce the simplest modification of general relativity, $f(R)$ gravity and its cosmological consequence in Section 1.1.2. We will then discuss further the general higher derivative gravity up to the quadratic level in curvature invariants in four-dimensional spacetime and the issue of the Ostrogradski instability in Section 1.1.3.

1.1.1 Conventions and notations

In this thesis, we will use the spacelike convention for the metric, i.e. in Minkowski spacetime, the line element is

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = -dt^2 + dx^2 + dy^2 + dz^2. \tag{1.2}$$

We will choose to write spacetime indices using the Greek alphabet and space indices using the Latin alphabet; we will also use units such that the speed of light is equal to 1 throughout most of the thesis. For de Sitter spacetime, we write the line element in the Friedmann-Lemaître-Robertson-Walker (FLRW) metric

$$ds^2 = a^2(t)\left[-dt^2 + dx^2 + dy^2 + dz^2\right], \tag{1.3}$$

with the understanding that the Hubble parameter is a constant in de Sitter spacetime, i.e.

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{\Lambda}{3}, \tag{1.4}$$

where the '$\dot{}$' denotes the derivative with respect to physical time and Λ is the cosmological constant in the action. Since there is no ambiguity in Sections 3.5 and 3.6, we will use t to denote 'conformal time' in these sections.

For the curvature terms, we adopt the convention of Misner, Thorne and Wheeler where the Christoffel connection, Riemann and Einstein tensor read

$$\Gamma^\alpha_{\mu\nu} = \frac{1}{2} g^{\alpha\beta}(\partial_\mu g_{\nu\beta} + \partial_\nu g_{\mu\beta} - \partial_\beta g_{\mu\nu}), \tag{1.5}$$

$$R^\mu{}_{\nu\alpha\beta} = \partial_\alpha \Gamma^\mu_{\nu\beta} - \partial_\beta \Gamma^\mu_{\nu\alpha} + \Gamma^\mu_{\sigma\alpha}\Gamma^\sigma_{\nu\beta} - \Gamma^\mu_{\sigma\beta}\Gamma^\sigma_{\nu\alpha}, \tag{1.6}$$

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R, \tag{1.7}$$

where $R_{\mu\nu} = R^\alpha{}_{\mu\alpha\nu}$ and $R = R^\alpha_\alpha$. Einstein's field equation (1.1) can be derived from a variation of the Einstein-Hilbert action with respect to the metric tensor, where the action is defined by

$$S_{EH} = \int d^4x\sqrt{-g}\left(\frac{M_P^2}{2}R + \mathcal{L}_M\right), \tag{1.8}$$

where g is the determinant of the metric tensor $g_{\mu\nu}$ and the stress-energy tensor is defined by

$$T_{\mu\nu} = \frac{2}{\sqrt{-g}}\frac{\delta\mathcal{L}_M}{\delta g^{\mu\nu}}. \tag{1.9}$$

In Chapter 4, we define the most general form of spherically symmetric static background metric by

$$\bar{g}_{\mu\nu} = diag(-e^{2\alpha(r)}, e^{2\beta(r)}, r^2, r^2\sin^2\theta). \tag{1.10}$$

The covariant derivatives are denoted by either a semicolon or ∇_μ. The four-dimensional d'Alembert operator will then be defined as $\Box = g_{\mu\nu}\nabla^\mu\nabla^\nu$.

1.1.2 f(R) gravity

In this section, we review the most enduring theories of modified gravity — $f(R)$ gravity — [4, 5, 17], which are derived from the easiest generalization of Einstein-Hilbert action

$$S = \frac{M_P^2}{2}\int d^4x\sqrt{-g}f(R) + \int d^4x\mathcal{L}_M(g_{\mu\nu}, \Psi_M). \tag{1.11}$$

The field equations can be derived by varying the action (1.11) with respect to the metric tensor $g_{\mu\nu}$

$$f'(R)R_{\mu\nu} - \frac{1}{2}fg_{\mu\nu} - \nabla_\mu\nabla_\nu f'(R) + g_{\mu\nu}\Box f'(R) = -\frac{T_{\mu\nu}}{M_P^2}, \tag{1.12}$$

where $f'(R)$ means that the functional derivative of $f(R)$ with respect to R, and $T_{\mu\nu}$ is the stress-energy tensor defined by a variation of the matter action with respect to $g_{\mu\nu}$ in the usual way.

One can see if $f(R) = R$, eq. (1.12) reduces to Einstein's field equations, while in all other cases, the equations are fourth order in terms of their derivatives. The existence of the higher order derivative terms might raise concerns about the Ostrogradski instability. However, the $f(R)$ gravity is free from the instability because the gauge constraints automatically remove the unstable degrees of freedom. We can see that the theory is stable by reformulating the action as general relativity with a minimally coupled scalar. We start by introducing a new field ψ and a dynamically equivalent action [18]

$$S_{ST} = \frac{M_P^2}{2}\int d^4x\sqrt{-g}\left[f(\psi)+f'(\psi)(R-\psi)\right]+\int d^4x\mathcal{L}_M(g_{\mu\nu},\Psi_M). \tag{1.13}$$

If $f''(\psi)\neq 0$, the variation of the action with respect to ψ will give us the equation $\psi = R$, which reproduces eq. (1.11). The action can now be written in the scalar-tensor form

$$S_{ST} = \frac{M_P^2}{2}\int d^4x\sqrt{-g}\left[\phi R - V(\phi)\right]+\int d^4x\mathcal{L}_M(g_{\mu\nu},\Psi_M), \tag{1.14}$$

with the definition $\phi = f'(\psi)$ and the potential

$$V(\phi) = \psi(\phi)\phi - f[\psi(\phi)]. \tag{1.15}$$

The action (1.14) can be cast into the Einstein frame by defining the conformal metric $g^E_{\mu\nu} = \phi g_{\mu\nu}$ and the field redefinition $\phi = \exp(\sqrt{2/(3M_P^2)}\varphi)$, with the action describing general relativity with a minimally coupled scalar

$$S_{ST} = \int d^4x\sqrt{-g^E}\left[\frac{M_P^2}{2}R^E - \frac{1}{2}(\nabla\varphi)^2 - U(\varphi)\right]+\int d^4x\mathcal{L}_M(\phi^{-1}(\varphi)g^E_{\mu\nu},\Psi_M), \tag{1.16}$$

where

$$U(\varphi) = \frac{M_P^2 V[\phi(\varphi)]}{2[\phi(\varphi)]^2}. \tag{1.17}$$

One can see from the matter action that the scalar field φ is directly coupled to matter in the Einstein frame. If we assume that there is no other fields from the matter Lagrangian in eq. (1.16), $f(R)$ gravity is nothing but general relativity with a minimally coupled scalar, where the scalar field has the right sign of its kinetic term, and it is stable if we choose its potential bounded from below.

Cosmological solutions

Most of the interest in $f(R)$ gravity emerges from the study of the accelerating expansion of the Universe, inflation and late-time accelerating expansion. Here, we review how to use this class of higher derivative gravity to model these phenomena. We start by writing down the line-element for the flat FLRW metric

$$ds^2 = -dt^2 + a^2(t)d\mathbf{x}^2. \tag{1.18}$$

The Friedmann equations with the action (1.11) can be written as

$$H^2 = \frac{8\pi G}{3}\left(\frac{Rf' - f - 6H\dot{R}f''}{2f'}\right), \tag{1.19}$$

$$2\dot{H} + 3H^2 = -\frac{8\pi G}{f'}\left[\dot{R}^2 f''' + 2H\dot{R}f'' + \ddot{R}f'' + \frac{1}{2}(f - Rf')\right], \tag{1.20}$$

where f' denotes the derivative of f with respect to the Ricci scalar. One can identify the effective energy density and pressure parameters by analogy to the Friedmann equations of general relativity and find

$$\rho_{eff} = \frac{Rf' - f - 6H\dot{R}f''}{2f'}, \tag{1.21}$$

$$P_{eff} = \frac{\dot{R}^2 f''' + 2H\dot{R}f'' + \ddot{R}f'' + \frac{1}{2}(f - Rf')}{f'}, \tag{1.22}$$

where ρ_{eff} has to be non-negative in a spatially flat FLRW spacetime due to the requirement in eq. (1.19). The effective equation of state can thus be found

$$w_{eff} = \frac{2\left[\dot{R}^2 f''' + 2H\dot{R}f'' + \ddot{R}f'' + \frac{1}{2}(f - Rf')\right]}{Rf' - f - 6H\dot{R}f''}. \tag{1.23}$$

From eq. (1.23), in general, if we want to use it to mimic the de Sitter solution with $w_{eff} = -1$, it must be

$$\frac{f'''}{f''} = \frac{\dot{R}H - \ddot{R}}{(\dot{R}^2)}. \tag{1.24}$$

Inflationary dynamics

Consider the model with

$$f(R) = R + \frac{R^2}{6M^2}, \tag{1.25}$$

where the constant M has a dimension of mass. This is the famous Starobinsky model which can generate an inflationary period while the inflation is eventually ended by the presence of the linear term R [3]. This model is the simplest extension of Einstein Hilbert action, and one might think that the theory is coming from the first and second order of expansion of general $f(R)$ gravity with respect to the derivatives, and thus is a effective theory. However, if we take this point of view, we assume that the first term in eq. (1.25) is larger than the second term, and since we know that an action dominated by the Einstein Hilbert term cannot generate a period of inflation, this will not work. Therefore, if we want to have inflationary dynamics, we assume that the Starobinsky model is a given model describing our universe.

The Friedmann equations (1.19) and (1.20) of this model can be written as

$$\ddot{H} - \frac{\dot{H}^2}{2H} + \frac{1}{2}M^2 H = -3H\dot{H}, \tag{1.26}$$

$$\ddot{R} + 3H\dot{R} + M^2 R = 0. \tag{1.27}$$

During the inflation, we assume that the slow-roll parameters are small i.e.

$$\varepsilon \equiv -\frac{\dot{H}}{H^2} \ll 1,$$

$$|\eta| \equiv \frac{|\dot{\varepsilon}|}{H\varepsilon} \ll 1,$$

and we can thus neglect the first two terms in eq. (1.26) and solve

$$H \simeq H_i - \frac{M^2}{6}(t - t_i), \tag{1.28}$$

$$a \simeq a_i \exp\left[H_i(t - t_i) - \left(\frac{M^2}{12}\right)(t - t_i)^2\right], \tag{1.29}$$

$$R \simeq 12H^2 - M^2, \tag{1.30}$$

where H_i and a_i are the Hubble parameter and the scale factor at the beginning of the inflation ($t = t_i$). The solution is an attractor solution while the accelerated expansion continues if

$$\varepsilon \simeq \frac{M^2}{6H^2} < 1. \tag{1.31}$$

If we transform this theory into the corresponding scalar-tensor theory, one can find in Einstein's frame that it corresponds to a scalar field ϕ with potential

$$V(\phi) = \frac{3M^2 M_{Pl}^2}{4}\left(1 - e^{-\sqrt{\frac{2}{3}}\frac{\phi}{M_{Pl}}}\right)^2. \tag{1.32}$$

The potential is displayed in Figure 1.1, where the slow-roll inflation occurs in the region $\phi \gg M_{Pl}$ and reheating occurs during the oscillations around the minimum of the potential.

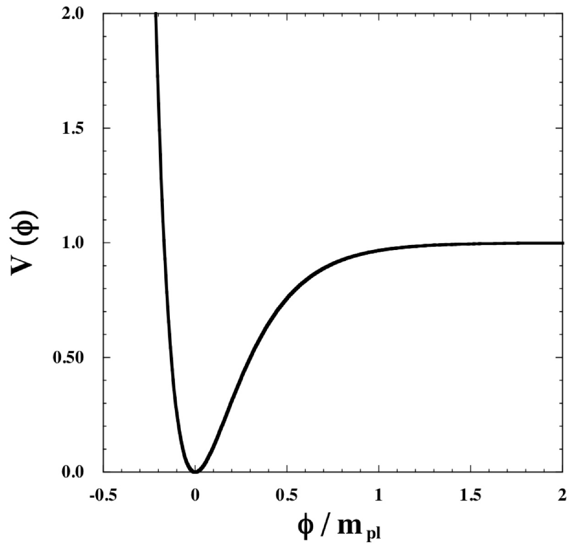

Fig. 1.1 The field potential in the Einstein frame corresponding to eq. (1.32). Inflation is realized in the regime $\phi \gg M_{Pl}$ [4].

1.1.3 Higher derivative gravity up to quadratic curvature invariant in four-dimensional spacetime

In the previous section, we considered theories which generalize the Einstein-Hilbert action by replacing the Ricci scalar R with an arbitrary function $f(R)$. In this section, we further generalize this by considering the action not only dependent on R but also the quadratic contractions of the Riemann curvature tensor: $R_{\mu\nu}R^{\mu\nu}$ and $R_{\mu\nu\sigma\rho}R^{\mu\nu\sigma\rho}$. In this thesis, we are interested in the theory where the action is dependent up to the quadratic curvature invariant

$$\mathcal{L} = \chi^{-1}\sqrt{-g}(R + \alpha R^2 + \beta R_{\mu\nu}R^{\mu\nu} + \gamma R_{\mu\nu\sigma\rho}R^{\mu\nu\sigma\rho}), \tag{1.33}$$

where α, β, and γ are constants. This model can be further simplified if we realize that, in four-dimensional spacetime, the Gauss-Bonnet term of the curvature invariants is a total divergence

$$4R_{\mu\nu}R^{\mu\nu} - R^2 - R_{\mu\nu\sigma\rho}R^{\mu\nu\sigma\rho} = \text{total divergence}, \tag{1.34}$$

which is a boundary term that has no effect on the equations of motion, and so can be ignored. The action is thus reduced to

$$S = \frac{M_P^2}{2}\int d^4x\,\sqrt{-g}(R - 2\Lambda + \alpha R^2 + \beta R_{\mu\nu}R^{\mu\nu}), \tag{1.35}$$

where we add the cosmological constant term and choose χ such that the theory reduces to GR when the linear term dominates.

This model was first studied by Stelle in 1977, where the theory was proved to be power counting renormalizable at the one-loop order [14]. The main reason for this is because the equations of motion are fourth order derivatives and the gravitation propagators thus behave like k^{-4} for large momenta. The integral of each Feynman diagram is suppressed by the extra k^{-2} compared with the gravitational propagators in GR, which behave like k^{-2}.

On the other hand, there is an extremely severe problem associated with this model — the inescapable Ostrogradski instability. The theory contains eight degrees of freedom which include the usual massless spin-2 excitation (the graviton), as well as the new massive spin-2 and scalar excitations. Of these, the massive and massless spin-2 excitations always have different signs for their kinetic terms, and one of them thus has a negative spectrum. This fact can be seen by first parameterizing the perturbations around the Minkowski metric and substituting them into the action; after some field redefinition, the action becomes

$$S = \frac{M_{Pl}^2}{2}\int d^4x \left[\mathcal{L}_E(\phi_{\mu\nu}) - \mathcal{L}_E(\psi_{\mu\nu}) + \frac{m_2^2}{4}\left(\psi_{\mu\nu}\psi^{\mu\nu} - \psi^\lambda_\lambda\psi^\rho_\rho\right) - \frac{3}{2}\partial_\mu\Phi\partial^\mu\Phi - \frac{3m_0^2}{2}\Phi^2\right], \tag{1.36}$$

where $\phi_{\mu\nu}$, $\psi_{\mu\nu}$, Φ are the massless spin-2, massive spin-2, and massive scalar excitations respectively, with

$$m_0^2 = \frac{1}{2(3\alpha+\beta)} \qquad \text{and} \qquad m_2^2 = -\frac{1}{\beta}, \tag{1.37}$$

are the mass of the scalar and the spin-2 excitations. The linearized Einstein Lagrangian is defined by

$$\mathcal{L}_E(\phi_{\mu\nu}) = \frac{1}{4}\phi_{\mu\nu}\Box\phi^{\mu\nu} - \frac{1}{4}\phi^\lambda_\lambda\Box\phi^\rho_\rho + \frac{1}{2}\phi^{\mu\nu}\partial_\mu\partial_\nu\phi^\lambda_\lambda - \frac{1}{2}\phi^{\mu\nu}\partial_\rho\partial_\nu\phi^\rho_\mu. \tag{1.38}$$

One can see from eq. (1.36) that the massless spin-2 and massive scalar excitations have the “right" signs of the kinetic terms while the massive spin-2 excitation has the “wrong" sign of the kinetic term and thus has a negative energy spectrum.

One of the main issues in this thesis is to specify this instability in different helicity sectors and introduce suitable constraints for removing the unstable degrees of freedom while keeping the stable degrees of freedom as well as preserving the improved renormalizable properties.

1.2 Analysis of constrained Hamiltonian systems

Gauge invariance—Constraints

Every gauge theory contains at least one dynamical variable which is specified by the "reference frame" which cannot be determined at every instant of time. This implies that we cannot find a unique solution to the equation of motion with an initial condition in a gauge theory, since a future change in the reference frame might induce a change in the variable (i.e. a gauge transformation) while keeping the initial condition fixed. We can thus conclude that the general solution to the equation of motion in a gauge theory contains arbitrary functions of time and, since the physically important variables should not depend on the choice of reference frame, they should be gauge invariant.

In this section, we will introduce a thorough analysis of gauge systems through the Hamiltonian formalism — the Dirac analysis. It will emerge from the discussion that the arbitrary functions of time in the general solution are related to the fact that not every canonical variable is independent. We will find that all gauge theories are constrained Hamiltonian systems but that the converse is untrue. The constrained Hamiltonian systems are more general and we will demonstrate how to apply the analysis to all types of constraint.

1.2.1 The Lagrangian and primary constraints

We start by studying the action principle of the action

$$S = \int_{t_1}^{t_2} dt\ L(q_n, \dot{q}_n), \qquad n = 1, \cdots, N, \tag{1.39}$$

where the Lagrangian is only dependent on the position q_n and the velocity $\dot{q}_n$ of the one-dimensional point particle. The classical motions of the system are those where the action is stationary under arbitrary variation δq_n and the conditions that need to be satisfied are the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^n}\right) - \frac{\partial L}{\partial q^n} = 0. \tag{1.40}$$

Equation (1.40) can be expanded into

$$\ddot{q}^m \frac{\partial^2 L}{\partial \dot{q}^m \partial \dot{q}^n} = \frac{\partial L}{\partial q^n} - \dot{q}^m \frac{\partial^2 L}{\partial q^m \partial \dot{q}^n}, \tag{1.41}$$

and the accelerations $\ddot{q}^n$ can be uniquely solved by the positions and velocities if and only if the matrix $\frac{\partial^2 L}{\partial \dot{q}^m \partial \dot{q}^n}$ is invertible, i.e. the determinant of it is not zero and the theory is called *non-degenerate*. If the determinant is zero, the acceleration cannot be uniquely determined, so we cannot solve the equations of motion with initial condition q_0, $\dot{q}_0$. There is at least one gauge degree of freedom in the theory and we will see that this is related to constraints in the phase space.

In order to study the problem with the Hamiltonian formalism, we first define the canonical momenta of the theory

$$p_n = \frac{\partial L}{\partial \dot{q}^n}. \tag{1.42}$$

If the matrix $\frac{\partial^2 L}{\partial \dot{q}^m \partial \dot{q}^n}$ is noninvertible, which implies that not all of the velocities $\dot{q}_n$ can be expressed by the canonical momenta p_m, i.e. not all of the momenta p_m are independent, there are some relations between the canonical coordinates

$$\varphi_m(q,p) = 0, \quad m = 1,2,\cdots,M. \tag{1.43}$$

These relations (1.43) are called primary constraints in this theory, and they define a submanifold that is smoothly embedded in the 2N-dimensional phase space where the physical degrees of freedom live on. If there are M independent primary constraints, the submanifold would be $(2N-M)$-dimensional.

For example, consider the following Lagrangian

$$L = \frac{1}{2}\left(\dot{q}_1 - \dot{q}_2\right)^2, \tag{1.44}$$

where $N = 2$ in this case, and the manifold is four-dimensional. The canonical momenta can be found by eq. (1.42)

$$\begin{aligned} p_1 &= \dot{q}_1 - \dot{q}_2, \\ p_2 &= \dot{q}_2 - \dot{q}_1, \end{aligned}$$

with one primary constraint ($M = 1$)

$$\varphi : p_1 + p_2 = 0, \tag{1.45}$$

here we introduce the notation ':' to denote "functional form given by". One can see in Figure 1.2 that the mapping from the phase space (q_n, p_n) to the configuration space $(q_n, \dot{q}_n)$ is multivalued. Every point in the $(q_n, \dot{q}_n)$ space is mapped on the straight line $p_1 + p_2 = 0$, and every point on the straight line $\dot{q}_2 - \dot{q}_1 = c$ is mapped to a single point in the phase space. This is because the manifold in the phase space is $(2N - M)$-dimensional, which is smaller than the manifold in the configuration space ($2N$-dimensional).

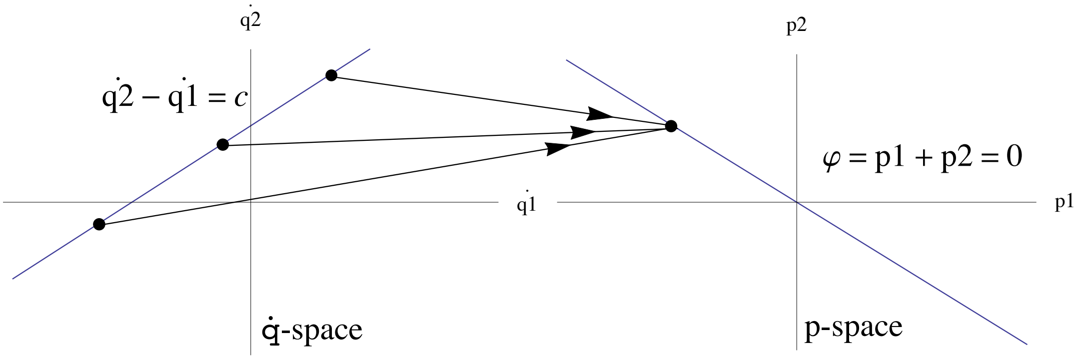

Fig. 1.2 The figure shows the configuration space and the phase space of the theory with Lagrangian $L = \frac{1}{2}(\dot{q}_1 - \dot{q}_2)^2$, which is an example of a theory with primary constraints. One can see the transformation $\dot{q} \to p$ is neither one-to-one nor onto mapping, so we cannot invert and write $\dot{q}$ as a function of (q, p).

1.2.2 The Hamiltonian

The second step in the formalism is to introduce the Hamiltonian. The canonical Hamiltonian is defined by

$$H = \dot{q}^n p_n - L, \tag{1.46}$$

which is a function of positions and the canonical momenta. We can vary the Hamiltonian with respect to the positions and the canonical momenta

$$\begin{aligned} \delta H &= \dot{q}^n \delta p_n + \delta \dot{q}^n p_n - \delta \dot{q}^n \frac{\partial L}{\partial \dot{q}^n} - \delta q^n \frac{\partial L}{\partial q^n} \\ &= \dot{q}^n \delta p_n - \delta q^n \frac{\partial L}{\partial q^n}. \end{aligned} \tag{1.47}$$

If the theory admits primary constraints, the δp_n in eq. (1.47) are not all independent but restricted in order to preserve the primary constraints $\varphi_m = 0$. In this case, the canonical Hamiltonian is only well defined on the submanifold where the primary constraints are satisfied, but can be extended arbitrarily off the manifold, i.e. we can extend the Hamiltonian as

$$H \to H_T = H + u^m(q,p)\varphi_m(q,p), \tag{1.48}$$

with m arbitrary functions $u^m(q,p)$, but retain the dynamics of the physical degrees of freedom. From now on, we will refer to the new Hamiltonian H_T as the total Hamiltonian. With this extension, eq. (1.47) can be written as

$$\left(\frac{\partial H}{\partial q^n} + \frac{\partial L}{\partial q^n} + u^m \frac{\partial \varphi_m}{\partial q^n}\right)\delta q^n + \left(\frac{\partial H}{\partial p_n} - \dot{q}^n + u^m \frac{\partial \varphi_m}{\partial p_n}\right)\delta p_n + \varphi_m \delta u^m = 0, \tag{1.49}$$

where we have used the fact that the primary constraints vanish throughout the whole physical phase space. The equations of motion are the evolution in which the Hamiltonian is invariant under arbitrary variations $\delta q^n, \delta p_n, \delta u^m$

$$\dot{q}^n = \frac{\partial H}{\partial p_n} + u^m \frac{\partial \varphi_m}{\partial p_n}, \tag{1.50}$$

$$\dot{p}_n = -\frac{\partial H}{\partial q^n} - u^m \frac{\partial \varphi_m}{\partial q^n}, \tag{1.51}$$

$$\varphi_m(q,p) = 0, \tag{1.52}$$

where we have used eqs. (1.40) and (1.42) to replace $\frac{\partial L}{\partial q^n}$ by $\dot{p}_n$. The equations of motion of arbitrary function of phase coordinates $F(q_n, p_n)$ can be generated by the total Hamiltonian with the Poisson bracket

$$\frac{dF(q_n,p_n)}{dt} = [F, H_T]_P, \tag{1.53}$$

where

$$[F,G]_P = \frac{\partial F}{\partial q^i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q^i}. \tag{1.54}$$

We can thus generate the time evolution of any functions of q, p with the total Hamiltonian.

1.2.3 Secondary constraints

The constrained theory always admits primary constraints. If the theory is consistent, we need the primary constraints to be preserved with time evolution, i.e. if we take the F in eq. (1.53) to be the primary constraints, we would expect $\dot{\varphi}_m = 0$ for a consistent theory. This requirement generates the so-called consistency relation

$$\dot{\varphi}_m = [\varphi_m, H_T]_P = [\varphi_m, H]_P + u^{m'}[\varphi_m, \varphi_{m'}]_P \approx 0. \tag{1.55}$$

Here, we introduce the *weak equality symbol* $\approx$ to emphasize that the quantity is numerically restricted to zero but does not vanish throughout the phase space; the equality is only satisfied on the submanifold defined by the constraints.

Usually, three possible relations emerge from equation (1.55). The first case is that the relation can be independent of u^m, but involve only q_m and p^m. In this case, if the relation is independent of primary constraints, we will find that the theory admits further constraints $\chi_k(q,p) \approx 0$, $(k = M+1,\cdots,M+K)$, which we call *secondary* constraints. We can iteratively generate more constraints through

$$[\chi_k, H_T]_P \approx 0 \tag{1.56}$$

until the relation is dependent on u^m or the primary constraints. The second case of eq. (1.55) is that the relation can be dependent on u^m, which we can use to find the expression of u^m. Finally, in the third case, if eq. (1.55) is dependent on primary constraints, it would be an identity.

After we have found all of the constraints through the consistency relation, operationally, there is no difference between the primary and secondary constraints, which we will combine as

$$\varphi_j \approx 0 \qquad j = 1,\cdots,M+K \equiv J. \tag{1.57}$$

1.2.4 First class and second class constraints

The more important classification of constraints is the concept of *first class* and *second class* constraints. A function $F(q,p)$ is said to be first class if the Poisson bracket of itself with every constraint vanishes weakly,

$$[F, \varphi_j]_P \approx 0, \qquad j = 1,\cdots,J. \tag{1.58}$$

If a function of the canonical variables is not first class, it is said to be second class, i.e. if $F(q,p)$ is second class, there is at least one constraint φ_1 such that the Poisson bracket $[F,\varphi_1]_P \not\approx 0$.

First and second class constraints are different in physics. Second class constraints are "physical" in the sense that the solutions to the equations of motion differ according to whether or not the constraints are present — e.g. a train restricted to move on a fixed rail-track enforces a second class constraint. Operationally, the second class constraints always emerge in pairs, which means that there is at least one redundant degree of freedom q_i, p_i

that can be removed from the Hamiltonian by using a pair of second class constraints. On the other hand, first class constraints are associated with some gauge freedom in the theory, i.e. the solutions of the equations of motion contain some arbitrary functions of time and hence describe physically equivalent systems. We can view the first class constraints as the generators of gauge transformation.

Proof: First class constraints generate gauge transformations

Considering a Lagrangian $L(q,\dot{q})$ with constraint

$$\varphi(q,p)=0. \tag{1.59}$$

The Hamiltonian can be derived from the Legendre transformation, and we can write the Lagrangian in the first order form

$$L(q,\dot{q},p)=p\dot{q}-H(q,p), \tag{1.60}$$

which can be cast into the original form if we replace p in eq. (1.60) by the definition of canonical momentum. We consider the following transformation with an arbitrary function $\alpha(t)$

$$\delta q=\alpha(t)[q(t),\varphi(q,p)]_P=\alpha(t)\frac{\partial\varphi(q,p)}{\partial p}, \tag{1.61}$$

$$\delta p=\alpha(t)[p(t),\varphi(q,p)]_P=-\alpha(t)\frac{\partial\varphi(q,p)}{\partial q}, \tag{1.62}$$

which we claim to be a gauge transformation. If this is the case, it will not change the action; we thus substitute eqs. (1.61) and (1.62) into the action

$$S[q,p]=\int dt\; L(q,\dot{q},p,\dot{p}). \tag{1.63}$$

The change in the action can be derived by

$$\delta S=\int dt\;\left[-\alpha\frac{\partial\varphi(q,p)}{\partial q}\dot{q}+p\frac{d}{dt}\left(\alpha\frac{\partial\varphi(q,p)}{\partial p}\right)+\alpha[\varphi,H]_P\right]$$

$$=\int dt\;\alpha\left[-\frac{\partial\varphi(q,p)}{\partial q}\dot{q}-\frac{\partial\varphi(q,p)}{\partial p}\dot{p}+[\varphi,H]_P\right]$$

$$= \int dt\ \alpha \left[-\frac{d\varphi(q,p)}{dt} + [\varphi, H]_P \right]. \tag{1.64}$$

If the constraint $\varphi(q,p)$ is not explicitly dependent on time, we can use the Hamiltonian (1.53) to reach the conclusion $\delta S = 0$ and hence finish the proof.

One should notice here that we assume that the theory only contains one constraint and it is thus a first class one. For the second class constraint $[\varphi_1, \varphi_2]_P \not\approx 0$, one can prove that eq. (1.64) still vanishes, but the transformation generated by φ_1 (eqs. (1.61) and (1.62)) does not preserve the constraint $\varphi_2 \approx 0$. The transformation generated by a second class constraint will make the theory inconsistent and the transformation is not a gauge transformation.

1.2.5 Second class constraints and the Dirac bracket

In order to continue analyzing the constrained system in the Hamiltonian formalism, we need to generalize the Poission bracket to the Dirac bracket, which is well-defined for the constrained systems. The spirit of the Dirac bracket is as follows: if the theory is constrained, the constraints reveals that some of the degrees of freedom are unimportant and so should be discarded in the modified Poisson bracket. The Dirac bracket is thus defined in such a way that it removes all of the virtual degrees of freedom from the Poisson bracket. Following this, there is no longer any ambiguity in the theory and we can set all of the constraints strongly equal to zero before evaluating the bracket.

Let us begin the Dirac analysis of the constrained system. For simplicity, we assume that there are only second class constraints in the system. The prescription of the first class constraints will be discussed in the next subsection.

We first define a matrix $C_{ab} \equiv [\varphi_a, \varphi_b]_P$. With this definition, we would like to prove the following theorem

Theorem: *If* $\det C_{ab} \approx 0$*, there exists at least one first class constraint among the* φ_a'*s.*

Proof: If $\det C_{ab} \approx 0$, one can find a nonzero solution λ^a such that $\lambda^a C_{ab} \approx 0$. We can thus find a constraint $\lambda^a \varphi_a$ to be first class, which proves the theorem.

If the theory only contains second class constraints, the matrix C_{ab} possesses an inverse C^{ab} whereby

$$C_{ab}C^{bc} = \delta_a^c. \tag{1.65}$$

The Dirac bracket is defined by

$$[F, G]_D = [F, G]_P - [F, \varphi_a]_P C^{ab} [\varphi_b, G]_P. \tag{1.66}$$

One can prove that the Dirac bracket has all of the good properties we need in the Hamiltonian formalism for the constrained system (see, e.g. [23])

$$[F,G]_D = -[G,F]_D, \tag{1.67}$$

$$[F,GR]_D = [F,G]_D R + G[F,R]_D, \tag{1.68}$$

$$[[F,G]_D,R]_D + [[R,F]_D,G]_D + [[G,R]_D,F]_D = 0. \tag{1.69}$$

Moreover, for any second class constraint φ_a and arbitrary function F

$$[\varphi_a,F]_D = 0, \tag{1.70}$$

and for a first class function χ and arbitrary function F

$$[F,\chi]_D \approx [F,\chi]_P. \tag{1.71}$$

We can see from eq. (1.70) that the second class constraints φ_a can be set to zero before or after the evaluation of the bracket. From eq. (1.71), since the total Hamiltonian is first class,[3] we can conclude that the total Hamiltonian along with the Dirac bracket generates the correct equations of motion for arbitrary function F, where

$$\dot{F} = [F,H_T]_D, \tag{1.72}$$

and all of the second class constraints $\varphi_a = 0$.

1.2.6 First class constraints and gauge fixing

We have mentioned that the presence of first class constraints indicates that there is gauge freedom in the theory and that a physical state can be represented by more than one set of canonical variables. In practice, we usually impose further constraints on the canonical variables to eliminate ambiguity, which is the so-called “gauge fixing". The geometric description of the process of gauge fixing can be seen in Figure 1.3, where the gauge fixing hypersurface should intersect every gauge orbit once and only once.

The gauge fixing conditions are a set of functions of canonical variables which are restricted to zero. After we choose a gauge, there should be a one-to-one mapping between each value of the canonical variables and the physical state. The gauge fixing functions are

[3] This fact comes from the requirement of the consistency relation: eqs. (1.55) and (1.56).

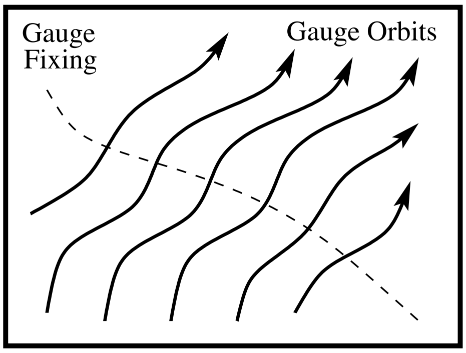

Fig. 1.3 In a gauge theory, all points that lie on a given gauge orbit (solid lines) can be mapped with each other via a gauge transformation and correspond to the same physical state. A good set of gauge fixing conditions (dashed lines) should intersect all of the gauge orbits once and only once[24].

in general like

$$C_j(q,p) \approx 0. \tag{1.73}$$

A satisfactory gauge fixation must satisfy two properties:

(a) The chosen gauge must be accessible. That is, for any set of canonical variables, a gauge transformation must exist which maps the given set onto a set which satisfies eq. (1.73). This requirement guarantees that eq. (1.73) does not affect the physics of the system but only imposes a restriction on the gauge freedom. Since the transformation generated by the first class constraints is of the form $\alpha^i[F,\varphi_i]_P$ and the number of independent parameters α^i is equal to the number of first class constraints φ_i, we conclude that the number of independent gauge fixing conditions eq. (1.73) cannot exceed that of the independent first class constraints.

(b) The gauge fixation eq. (1.73) must fix the gauge completely. That is, there is no longer any gauge transformation other than the identity which preserves eq. (1.73). This means that, if the equation

$$\alpha^i[C_j,\varphi_i]_P \approx 0, \tag{1.74}$$

this must imply

$$\alpha^i = 0, \tag{1.75}$$

which only happens when the number of independent equations is equal to or greater than the unknown parameters α^i. From (a) and (b), we conclude that the number of independent gauge fixing conditions must be equal to the number of independent first class constraints. The Poisson brackets $[\varphi_i, C_j]_P$ thus form a square matrix, which must be invertible in order for eq. (1.74) to imply eq. (1.75) and thus the condition

$$\det[C_j, \varphi_i]_P \neq 0. \tag{1.76}$$

However, this condition implies that φ_i, C_j are second class constraints, so we conclude that, following complete gauge fixing, there is no first class constraint left and all of the constraints are second class. We can thus use the Dirac formalism in Section 1.2.5 to analyze the system, and at this point can treat the system as free of constraints since all of the constraints can be regarded as identities which express some canonical variables in terms of others.

1.2.7 Counting the degrees of freedom

We summarize this subsection by counting the physical canonical variables of a theory containing first and second class constraints:

$$\begin{aligned}
2\times&\begin{pmatrix}\text{Number of physical}\\ \text{degrees of freedom}\end{pmatrix} = \begin{pmatrix}\text{Number of independent}\\ \text{canonical variables}\end{pmatrix}\\
&= \begin{pmatrix}\text{Total number of}\\ \text{canonical variables}\end{pmatrix} - \begin{pmatrix}\text{Number of original}\\ \text{second class constraints}\end{pmatrix}\\
&- \begin{pmatrix}\text{Number of first}\\ \text{class constraints}\end{pmatrix} - \begin{pmatrix}\text{Number of gauge}\\ \text{fixing conditions}\end{pmatrix}\\
&= \begin{pmatrix}\text{Total number of}\\ \text{canonical variables}\end{pmatrix} - \begin{pmatrix}\text{Number of original}\\ \text{second class constraints}\end{pmatrix} - 2\times\begin{pmatrix}\text{Number of first}\\ \text{class constraints}\end{pmatrix}.
\end{aligned} \tag{1.77}$$

As the number of the second class constraints is always even, the number of independent canonical variables is also even, which corresponds to an integer number of physical degrees of freedom.

1.3 Outline of the thesis

In this dissertation, we examine the higher derivative theories with constraints by using Dirac's analysis of constrained systems; we apply these methods to study the higher derivative gravitational theory with quadratic curvature invariants in four-dimensional spacetime. We also apply a similar methodology to a class of constrained scalar field theory and find a classical spherically symmetric static solution. The outline of the thesis is as follows.

In Chapter 2, we study higher derivative theories in the context of a one-dimensional point particle where the position of the particle is a function only dependent on time. We show that the inevitable linear instability, the Ostrogradski instability, exists in all of the non-degenerate higher derivative theories and can only be removed by introducing appropriate constraints if the original theory's phase space is reduced.

In Chapter 3, we study the most general higher derivative gravitational theory with the action containing quadratic curvature invariants in four-dimensional spacetime perturbatively. This theory was first studied by Stelle [14, 15] in the 1970s and is interesting because of its renormalizability. However, like all of the other non-degenerate higher derivative theories, it suffers from the Ostrogradski instability. We generalize the method developed in Chapter 2 and apply it to the linearized version of higher derivative gravity. We show that, given suitable parameters, the instability in different helicity sectors can be removed at the same time if the effective dimensionality is reduced, which confirms the conclusion outlined in Chapter 2. The constrained higher derivative theory retains the renormalization properties of higher derivative gravity at the cost of the Lorentz symmetry being explicitly broken.

In Chapter 4, we investigate a new class of two scalar-field theory, where one of them is a Lagrange multiplier enforcing a constraint relating the value of the other scalar field to the norm of its derivative. The fluid can possess non-zero pressure while the energy always flows along timelike geodesics, same as normal dust [25]. The theory is known to admit a cosmological solution whereby the effective equation of state evolves from 0 to -1 and thus can unify dark energy and dark matter into a model with one degree of freedom. Motivated by the cosmological solution, we study the spherically symmetric static solution of this theory and show that this can be found with appropriate exponential potential. By using Regge-Wheeler-Zerilli decomposition, we also study the perturbation around the solution and conclude that the modes with even symmetry are unstable.

We conclude in Chapter 5 and present the technical details in the appendices.

Chapter 2

Higher derivative theories with constraints: exorcising the Ostrogradski ghost

We begin our study of higher derivative theories by investigating the one-dimensional point particle case, whereby the variable $q(t)$ is the position of the particle and a function only dependent on time. In this chapter, we prove that the linear instability in a non-degenerate higher derivative theory, the Ostrogradski instability, can only be removed by the addition of constraints if the original theory's phase space is reduced.

2.1 Introduction

When Newton wrote his second Law of Motion

$$\ddot{q} = \frac{F(q)}{m}, \tag{2.1}$$

i.e. motion is described by an equation second order in the time derivative of the fundamental dynamical variable position q, he chose wisely. As is now well-known, almost 200 years later, Ostrogradski [1] proved a theorem which showed that, in any non-degenerate theory whose fundamental dynamical variable is higher than second order in time derivative, there exists linear instability.

Consider the possibility that, Newton had, instead, chosen the fourth order theory

$$q^{(4)} = \frac{dV(q)}{dq} \tag{2.2}$$

with V being some function of q, i.e. a potential. This equation of motion can be obtained from a higher derivative action of the following form

$$S = \int dt \, \left(\frac{1}{2}\ddot{q}^2 - V(q) \right). \tag{2.3}$$

Since eq. (2.2) is fourth order, the phase space is four-dimensional. Without going into too much detail here — we will get there soon enough — we can describe the phase space by a pair of canonical variables and their momenta (P_1, Q_1) and (P_2, Q_2), with the Hamiltonian

$$H = P_1 Q_2 + \frac{P_2^2}{2} + V(Q_1). \tag{2.4}$$

One can always choose $V(Q_1)$ to be some function which is bounded from below, say $V(Q_1) = Q_1^2$. More problematic, however, is the first term which signals the famous Ostrogradski linear instability. The word "linear" in "linear instability" refers to the linearity of the P_1 in this term — since P_1 is free to roam the phase space, there is no barrier preventing some degrees of freedom of the theory from probing arbitrarily negative energies. In other words, the Hamiltonian is not bounded from below.[1]

This instability *per se* is no bad thing, but it becomes severe when interactions with other degrees of freedom, whose Hamiltonians are bounded from below, are introduced. The presence of these negative energy states means that there exists a vast phase space where the Hamiltonian is negative; hence the modes will begin to populate them by entropic argument alone while, through the conservation of energy, thereby creating an equally large number of positive energy modes in the interacting degrees of freedom [26, 27]. This is the onset of the instability. Note that, while this is a *classical* instability, in quantum theory, negative energy modes are particularly sick — attempts to quantize them canonically will either lead to negative norm (and hence undefined) states or negative energy states (and hence runaway particle production). Since negative norm states are often called "ghosts" in quantum theory, higher derivative theories are often called "ghost-like".[2]

Recently, there has been a resurgence of interest in higher derivative theories, particularly within attempts to modify gravity [15, 17, 28–37]. It is well-known that higher deriva-

[1] Technically, it is single side boundedness that is important; a Hamiltonian that is bounded from above is equally good — one can simply flip its sign.

[2] However, the Hamiltonian of the non-degenerate higher derivative model is bounded neither from below nor above. With suitable canonical transformation, the unstable degree of freedom can be "ghosts" or "tachyons". This may be checked from the wave function of the theory. If the wave function is oscillatory (exponentially growing/decaying), it is a ghost (tachyon). The difference between ghosts and tachyons is discussed in Appendix A.

tive theories of the $f(R)$ form are secretly healthy as they are *degenerate* — a technically important distinction which means that the highest derivative term cannot be written as a function of canonical variables and the theory is thus constrained. The naive unstable degree of freedom is rendered harmless by a gauge constraint. In fact, $f(R)$ can be recast as an (interacting) theory of a scalar and two graviton modes (see Section 1.1.2 or references [3–5, 38–41]).

Furthermore, there is also great interest in the so-called "higher derivative" scalar field theories, such as Galileon or Lovelock gravity [42–47] which, when coupled non-trivially with the metric, result in interesting scalar field dynamics which cannot be reproduced by simple $f(R)$-type modifications. These theories, while naively looking like "higher derivative" theories (in the sense that, in the Lagrangian, there are terms of second order and higher in time derivatives), are secretly completely healthy *non*-higher-derivative theories; their equations of motion are second order in time derivatives and so the phase space of the theories are two-dimensional. These properties have been achieved by the addition of structure in the Lagrangian — usually by the clever cancellation of higher derivative terms in the equations of motion — as seen in Galileon theory. We do not consider this class of theories as higher derivative theories and they do not suffer from the Ostrogradski instability.

On the other hand, in true non-degenerate higher derivative theories, the Ostrogradski instability is ubiquitous — as we will review below (also see [41, 48]). Theories employing curvature invariants such as $R_{\mu\nu}R^{\mu\nu}$, $R_{\mu\nu\sigma\gamma}R^{\mu\nu\sigma\gamma}$ [14, 32–35, 49–53] or the Weyl invariant $C_{\mu\nu\sigma\gamma}C^{\mu\nu\sigma\gamma}$ [37, 54, 55],[3] are non-degenerate higher derivative theories and hence suffer from the problem of Ostrogradski instability. The generic non-degenerate higher derivative theories are inevitably unstable and often avoided in the literature. Since Ostrogradski's theorem is so simple to prove and requires very few initial assumptions, it is incredibly powerful [26, 57–61].[4]

One way to deal with the instability of the generic non-degenerate higher derivative theories is to impose boundary conditions in such a way that the unstable modes vanish. For example, the modes with the wrong sign of the kinetic terms are "turned off" by imposing suitable boundary conditions [54, 65]. However, this is not a satisfactory solution: as we explained above regarding the simple example of Newton's law of motion, in the presence

[3]In 4D, the Weyl invariant $C_{\mu\nu\sigma\rho}C^{\mu\nu\sigma\rho}$ can be written as $\frac{1}{2}(R_{\mu\nu}R^{\mu\nu} - \frac{1}{3}R^2)$ because the Gauss-Bonnet term $\sqrt{-g}(R_{\mu\nu\sigma\rho}R^{\mu\nu\sigma\rho} - 4R_{\mu\nu}R^{\mu\nu} + R^2)$ is a total divergence if the coupling is fixed by demanding a condition on the asymptotic curvature [56]. The boundary term does not contribute to the classical equations of motion.

[4]There are further exceptions: the higher derivative theories can be made degenerate by interacting with extra fields [62], or the theories with infinite order time derivatives (i.e. the nonlocal theory) are free of the Ostrogradski instability [27, 41, 63, 64]

of higher order interaction terms beyond the quadratic power of the field, the *vacuum* states will rapidly decay (even classically) into states with positive and negative energy modes by the entropic argument [26, 27, 41, 66]. The "removed" instability is thus revived.[5]

On the other hand, one might try to eliminate the instability by imposing constraints (for example, those suggested by [25, 27, 70]), i.e. one selectively restricts the trajectories of the degrees of freedom such that the Hamiltonian becomes bounded from below. The implementation of constraints into the theory requires the introduction of auxiliary variables and hence the enlargement of the total phase space (the dimensionality of the *reduced* phase space is still the same or smaller since the trajectories are constrained). As a consequence, one may hope to change the orbits of the trajectories of the theory to a degree which is sufficient to cure it of instability.

Using the fourth order theory example above, one can imagine a modification

$$S = \int dt \, \left(\frac{1}{2}\ddot{q}^2 - V(q) + \lambda f(q,\dot{q},\ddot{q}) \right), \tag{2.5}$$

where λ is an auxiliary field which enforces the constraint $f(q,\dot{q},\ddot{q}) = 0$. We emphasize that the action (2.5) is a *different* physical theory from the action (2.3) provided that the constraint cannot be gauged away. Can we cleverly choose the function f such that this theory, despite being a higher derivative theory, is free of linear instability?

In this chapter, we will prove that, in order to remove the instability through the imposition of constraints, the constraints must *reduce* the effective dimensionality of the phase space of the original theory. For example, an unstable theory with a six-dimensional phase space can be rendered stable by reducing the phase space to dimension four or less by employing Lagrange multiplier or auxiliary fields.

This chapter is organized as follows. In Section 2.2, we review Ostrogradski's theorem in the context of the famous Pais-Uhlenbeck oscillator. We demonstrate how the Ostrogradski instability appears in this particular model and this result can be generalized to the most general non-degenerate higher derivative model. In Section 2.3, we show that, for the simple case of a second order (in the action) theory, the addition of Lagrange multipliers which do not reduce the original phase space makes the theory unstable. In Section 2.4, we prove in general the previous statement in the context of the Nth order higher derivative theory with M auxiliary variables. In Section 2.5, we show how an unstable non-degenerate higher derivative theory can be rendered stable by reducing the dimensionality of the original phase

[5] In the literature, some authors claim that, under special initial conditions, the self-interacting higher derivative theory (Pais-Uhlenbeck oscillator) is stable [67].

space. In Section 2.6, we apply the general procedure of the stabilization of non-degenerate higher derivative theory to the Pais-Uhlenbeck oscillator as a final example. A summary is provided in Section 2.7.

In Appendix A, we demonstrate how the technical difference between the tachyon and the ghost.

2.2 Ostrogradski's theorem: an example

Ostrogradski's theorem [1, 41] can be stated as follows:

If the higher order time derivative Lagrangian is non-degenerate, there is at least one linearly unstable degree of freedom in the Hamiltonian of this system.

As noted in Section 1.2, *non-degeneracy* is a technical term which states that there is a *one-to-one* mapping between the configuration space $(q, \dot{q}, \cdots)$ and the phase space $(Q, P, \cdots)$, which also means that the highest time derivative term can be expressed in terms of canonical variables. For example, in a theory with Lagrangian

$$L = \frac{1}{2}\dot{q}^2 - \frac{1}{2}q^2, \tag{2.6}$$

the canonical momentum is

$$P = \frac{\delta L}{\delta \dot{q}} = \dot{q}, \tag{2.7}$$

which allows us trivially to invert $\dot{q} = F(P, Q) = P$ and the theory is non-degenerate.

In a Nth order higher derivative theory, this translates into expressing $q^{(N)}$ as a function of the canonical variables Q_i and P_i. Degenerate theories, on the other hand, are non-invertible and either stable on their own or may be made stable through the introduction of constraints [41] — such theories will not be discussed in this chapter.

A famous example of a non-degenerate higher derivative theory is the Pais-Uhlenbeck (PU) oscillator [71]. Here, we demonstrate the characteristic of this model by following the discussion in [68, 69]. The PU action is given by

$$S_{PU} = \int dt L_{PU} = \frac{\gamma}{2}\int dt \left[\ddot{q}^2 - \left(w_1^2 + w_2^2\right)\dot{q}^2 + w_1^2 w_2^2 q^2\right], \tag{2.8}$$

where γ, w_1, and w_2 are positive constants and, without any loss of generality, we assume $w_1 \geq w_2$. The equation of motion of the Pais-Uhlenbeck oscillator containing terms up to the fourth time derivative is

$$\frac{d^4q}{dt^4}+(w_1^2+w_2^2)\frac{d^2q}{dt^2}+w_1^2w_2^2q=0, \tag{2.9}$$

and hence requires four initial value data (q_0, $\dot{q}_0$, $\ddot{q}_0$, $q_0^{(3)}$), allowing us to solve $q(t)$, to obtain

$$\begin{aligned} q(t)= &-\frac{w_2^2q_0+\ddot{q}_0}{w_1^2-w_2^2}\cos(w_1t)-\frac{w_2^2\dot{q}_0+q_0^{(3)}}{w_1(w_1^2-w_2^2)}\sin(w_1t) \\ &+\frac{w_1^2q_0+\ddot{q}_0}{w_1^2-w_2^2}\cos(w_2t)+\frac{w_1^2\dot{q}_0+q_0^{(3)}}{w_2(w_1^2-w_2^2)}\sin(w_2t). \end{aligned} \tag{2.10}$$

Since the solution depends on four initial value data, the phase space must be four-dimensional, and Ostrogradski's choice for the canonical coordinates is

$$\begin{aligned} Q_1\equiv q \longleftrightarrow\ P_1 &\equiv \frac{\delta L_{PU}}{\delta\dot{q}} = -\gamma(w_1^2+w_2^2)\dot{q}-\gamma q^{(3)} \\ Q_2\equiv \dot{q} \longleftrightarrow\ P_2 &\equiv \frac{\delta L_{PU}}{\delta\ddot{q}} = \frac{\partial L_{PU}}{\partial\ddot{q}} = \gamma\ddot{q}. \end{aligned} \tag{2.11}$$

Non-degeneracy implies that $\ddot{q}$ can be inverted and written as a function of the canonical variables Q_i and P_i — here, this is clearly the case. On the other hand, a degenerate model is always guaranteed to have constraints. For example, if the model is degenerate, say if $P_2 = \delta L_{PU}/\delta\ddot{q}$ is an arbitrary function $f(q,\dot{q})$ but not dependent on $\ddot{q}$, then, from the definition of canonical coordinates in eq. (2.11), there will be a primary constraint $P_2-f(Q_1,Q_2)=0$, which will reduce the number of physical degrees of freedom and the final phase space will be smaller.

The Hamiltonian of the Pais-Uhlenbeck oscillator is, as usual, obtained by Legendre transforming

$$\begin{aligned} H_{PU} &= P_1\dot{q}+P_2\ddot{q}-L_{PU} \\ &= P_1Q_2+\frac{P_2^2}{2\gamma}+\frac{\gamma}{2}(w_1^2+w_2^2)Q_2^2-\frac{\gamma}{2}w_1^2w_2^2Q_1^2, \end{aligned} \tag{2.12}$$

where the time evolution of any function of canonical variables $F(Q_i,P_i)$ are generated by the Hamiltonian via the Poisson Bracket $\dot{F}(Q_i,P_i)=[F(Q_i,P_i),H_{PU}]_P$. As we discussed in Section 1.2, the evolution equations in this Hamiltonian formalism reproduce the Euler-Lagrange equation in the Lagrangian formalism, so it is the right Hamiltonian of the system.

The Hamiltonian is conserved if the Pais-Uhlenbeck Lagrangian is not explicitly dependent on t, thus we can view the Hamiltonian as "energy."

As in eq. (2.12), the Hamiltonian is linearly dependent on P_1 and it signals that the system is unstable. The $P_1 Q_2$ term can be arbitrarily negative when $P_1 \to -\infty$, $Q_2 > 0$, or vice versa, and the Hamiltonian is thus unbounded from below, which means that there is no well-defined vacuum state in the theory. Ostrogradski's result is that all of the Hamiltonians of non-degenerate higher time derivative theory suffer from the Ostrogradski instability.

2.3 Constraints do not cure Ostrogradski's instability if the dimensionality is not reduced

In this section, we will show that the Ostrogradski instability in general cannot be cured by adding constraints to the theory if the dimensionality of the phase space is not reduced by the constraints, i.e. one can only possibly selectively constrain the unstable degrees of freedom and remove them from the physical phase space if the dimension of the phase space is less than that of the original unstable higher derivative theory. We will introduce the constraints through auxiliary variables λ_i in such a way that there is no time derivative on λ_i in the Lagrangian and primary constraints are thus introduced through their canonical momenta, $P_{\lambda_i} = 0$. We will use Dirac's method introduced in Section 1.2 to analyze the higher order theory with constraints [6, 23, 72–74].

First, we will show that the most general non-degenerate second time derivative Lagrangian with one extra auxiliary field (and hence a pair of second class constraints) does not cure its instability without the dimensionality of the phase space being reduced. We then apply this result to the Pais-Uhlenbeck model. We generalize our result to any Nth order non-degenerate higher derivative theory in Section 2.4.

2.3.1 General second order non-degenerate theory with second class constraints

In this subsection, we will show that the general second order non-degenerate theory with a pair of second class constraints is unstable. As discussed in Section 1.2, we can "gauge fix" theories with first class constraints and these 'gauge fixing" functions appear as new (primary) constraints in the theory and, once introduced, the original first class constraint and the new gauge fixing constraint both become second class constraints. Hence, when considering instability, it is clear that, once a general proof for second class constraints is shown, it is complete — physics does not depend on gauge choices, after all.

The most general second order time derivative Lagrangian with one auxiliary field, λ, is given by the Lagrangian

$$L = f(q, \dot{q}, \ddot{q}, \lambda). \tag{2.13}$$

The equations of motion of this Lagrangian are

$$\frac{\partial f}{\partial \lambda} = 0 \tag{2.14}$$

$$\frac{\partial f}{\partial q} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{q}}\right) + \frac{d^2}{dt^2}\left(\frac{\partial f}{\partial \ddot{q}}\right) = 0, \tag{2.15}$$

where we assume that the theory is non-degenerate and we can solve $q(t)$ and $\lambda(t)$ with six initial value data (q_0, $\dot{q}_0$, $\ddot{q}_0$, $q_0^{(3)}$, λ_0, $\dot{\lambda}_0$). The phase space is thus six-dimensional and, following Ostrogradski's spirit, the choice of canonical variables is

$$Q_1 \equiv q \quad \longleftrightarrow \quad P_1 \equiv \frac{\delta L}{\delta \dot{q}} = -\frac{d}{dt}\frac{\partial f}{\partial \ddot{q}} + \frac{\partial f}{\partial \dot{q}} \tag{2.16}$$

$$Q_2 \equiv \dot{q} \quad \longleftrightarrow \quad P_2 \equiv \frac{\delta L}{\delta \ddot{q}} = \frac{\partial f}{\partial \ddot{q}} \tag{2.17}$$

$$Q_3 \equiv \lambda \quad \longleftrightarrow \quad P_3 \equiv \frac{\delta L}{\delta \dot{\lambda}} = 0, \tag{2.18}$$

where $\varphi_1 : P_3 = 0$ is the primary constraint, as in section 1.2, we introduce the notation ':' to denote "functional form given by". We can invert $\ddot{q} = h(Q_1, Q_2, Q_3, P_2)$ by using eq. (2.17) and the total Hamiltonian H_T of this system is defined by

$$H_T = P_1 Q_2 + P_2 h(Q_1, Q_2, Q_3, P_2) - f(Q_1, Q_2, Q_3, h) + u_1 \varphi_1, \tag{2.19}$$

where u_1 is a function of canonical variables which can be found later,[6] but since we are only interested in the stability of the physical degrees of freedom on the reduced phase space where $\varphi_1 = 0$, we will not write u_1 explicitly.

With the same procedure we introduced in Section 1.2, we can use the *consistency relations* to find further secondary constraints. In this case, only one further secondary constraint

[6]That is, by imposing the consistency relations.

exists as expected, which is

$$\begin{aligned}
\varphi_2 : [\varphi_1, H_T]_P &\equiv \sum_i \left(\frac{\partial \varphi_1}{\partial Q_i} \frac{\partial H_T}{\partial P_i} - \frac{\partial H_T}{\partial Q_i} \frac{\partial \varphi_1}{\partial P_i} \right) \\
&= -P_2 \frac{\partial h}{\partial Q_3} + \frac{\partial h}{\partial Q_3} \left[\frac{\partial f}{\partial \ddot{q}} \right]_{\ddot{q}=h} + \left. \frac{\partial f}{\partial \lambda} \right|_{\lambda=Q_3} \\
&= \left. \frac{\partial f}{\partial \lambda} \right|_{\lambda=Q_3} (Q_1, Q_2, Q_3, h) \approx 0. \qquad (2.20)
\end{aligned}$$

As in Section 1.2, "$\approx$" is the weak equality. If φ_2 is dependent on Q_3, $[\varphi_1, \varphi_2]_P \not\approx 0$, then both φ_1 and φ_2 are second class constraints and there are no further constraints from the consistency relations; further consistency relations only tell us the form of the arbitrary function u_1. Using the two second class constraints, we can rewrite (Q_3, P_3) as functions of other canonical variables ($Q_3 \approx \mathcal{F}_1(Q_1, Q_2, P_2)$, $P_3 = 0$). The reduced Hamiltonian H_R of the physical degree of freedom becomes

$$H_R = P_1 Q_2 + P_2 h(Q_1, Q_2, \mathcal{F}_1(Q_1, Q_2, P_2), P_2) - f(Q_1, Q_2, h, \mathcal{F}_1). \qquad (2.21)$$

The reduced Hamiltonian is always linearly dependent on P_1 for any conceivable Lagrangian $L = f(q, \dot{q}, \ddot{q}, \lambda)$, which is the signal of instability.

On the other hand, if φ_2 is not dependent on Q_3, φ_1 and φ_2 commute with each other. In this case, we can find further constraints from the consistency relations (and hence a reduction in the effective dimensionality of the phase space) and we should check whether the constraints are first or second class after we find them all. We will provide examples in the following sections.

Even in this simple example, one can quickly see that the only possible way to cure the instability comes from the further constraints generated by the consistency relation (φ_3 and φ_4). The instability's root cause is the pesky linear term $P_1 Q_2$ and, to fix this instability, one must find a constraint where Q_2 must be some function of P_1 — although it is clear that, when generating the constraint φ_2 with the consistency relation, P_1 never enters the equation (2.20).

2.3.2 The Pais-Uhlenbeck model with constraint

We will now apply the above result to the Pais-Uhlenbeck model as an example. We consider the constraint $\ddot{q}^2 = \dot{q}^2$ to present a flavor of how instability is unavoidable if the dimensionality is not reduced. In this case, the dimensionality of the phase space remains the same (i.e. 4), with or without the constraint term.

Constraint: $\ddot{q}^2 - \dot{q}^2 = 0$

The Lagrangian of Pais-Uhlenbeck model with constraint $\ddot{q}^2 - \dot{q}^2 = 0$ is given by

$$L_{PUC} = \frac{\gamma}{2}[\ddot{q}^2 - (w_1^2 + w_2^2)\dot{q}^2 + w_1^2 w_2^2 q^2] + \frac{\lambda}{2}(\ddot{q}^2 - \dot{q}^2). \tag{2.22}$$

This model is an example where φ_2 is dependent on Q_3.

The equations of motion generated from varying the action with respect to λ and q becoming differential equations of both variables

$$\ddot{q}^2 - \dot{q}^2 = 0, \tag{2.23}$$

$$\gamma\frac{d^4 q}{dt^4} + \gamma(w_1^2 + w_2^2)\frac{d^2 q}{dt^2} + \gamma w_1^2 w_2^2 + \frac{d^2}{dt^2}(\lambda \ddot{q}) + \frac{d}{dt}(\lambda \dot{q}) = 0, \tag{2.24}$$

and the functions $q(t)$ and $\lambda(t)$ can be solved with four initial value data q_0, $\dot{q}_0$, λ_0, and $\dot{\lambda}_0$. The phase space of the physical degrees of freedom is thus dimension four. Following the same procedure employed in the last section, the choice of canonical variables is

$$Q_1 \equiv q \longleftrightarrow P_1 \equiv \frac{\delta L}{\delta \dot{q}} = -(\gamma + \lambda)q^{(3)} - \dot{\lambda}\ddot{q} - [\lambda + \gamma(w_1^2 + w_2^2)]\dot{q} \tag{2.25}$$

$$Q_2 \equiv \dot{q} \longleftrightarrow P_2 \equiv \frac{\delta L}{\delta \ddot{q}} = (\gamma + \lambda)\ddot{q} \tag{2.26}$$

$$Q_3 \equiv \lambda \longleftrightarrow P_3 \equiv \frac{\delta L}{\delta \dot{\lambda}} = 0. \tag{2.27}$$

From eq. (2.26), we can invert $\ddot{q} = P_2/(\gamma + Q_3)$, and the total Hamiltonian is

$$H_{PUCT} = P_1 Q_2 + \frac{P_2^2}{2(\gamma + Q_3)} + \frac{1}{2}[Q_3 + \gamma(w_1^2 + w_2^2)]Q_2^2 - \frac{\gamma}{2}w_1^2 w_2^2 Q_1^2 + u_1\varphi_1. \tag{2.28}$$

The primary constraint is $\varphi_1 : P_3 = 0$, and there is only one secondary constraint

$$\begin{aligned} \dot{\varphi}_1 = [\varphi_1, H_T]_P = \frac{1}{2}\left[\frac{P_2^2}{(\gamma + Q_3)^2} - Q_2^2\right] \\ \Rightarrow \varphi_2 : \frac{P_2}{(\gamma + Q_3)} \pm Q_2 \approx 0. \end{aligned} \tag{2.29}$$

Here, the constraint algorithm bifurcates, and we choose $P_2/(\gamma + Q_3) - Q_2 \approx 0$ instead of

$P_2/(\gamma+Q_3)+Q_2 \approx 0$ — one can check that choosing the other branch does not change the results.[7] The constraints are both second class and we can use them to rewrite Q_3, P_3 as the functions of other canonical variables. The reduced Hamiltonian of the Pais-Uhlenbeck model with primary constraint $\ddot{q}^2-\dot{q}^2$ is thus

$$H_{PUCR}=P_1Q_2+P_2Q_2+\frac{\gamma}{2}(w_1^2+w_2^2-1)Q_2^2-\frac{\gamma}{2}w_1^2w_2^2Q_1^2. \tag{2.30}$$

This Hamiltonian remains linearly dependent on P_1 and P_2, hence it still suffers from the Ostrogradski instability.

2.4 *N*th order theory with *M* auxiliary variables

It is straightforward to generalize our result from the previous section to an *N*th order derivative theory (with $N>2$) with *M* auxiliary variables.

When we introduce constraints with *M* auxiliary variables into an *N*th order theory, it is clear that, since the *M* variables are non-dynamical, they will not enlarge the effective dimensionality of the original unconstrained phase space, which is $2N$. We consider the case where the number of constraints generated by *M* auxiliary variables is $2M$ in Section 2.4.1.

2.4.1 *M* auxiliary variables with $2M$ constraints

Consider the most general *N*th order theory with *M* auxiliary variables

$$L_N=f(q,\dot{q},\ddot{q},\ldots,q^{(N)},\lambda_1,\lambda_2,\ldots,\lambda_M). \tag{2.31}$$

There are $M+1$ Euler-Lagrange equations from varying L_N with respect to λ_a and q

$$\frac{\partial f}{\partial \lambda_a}=0 \;\; (a=1,2,\ldots,M) \tag{2.32}$$

$$\sum_{i=0}^{N}(-\frac{d}{dt})^i\frac{\partial f}{\partial q^{(i)}}=0 \;\; (i=0,1,2,\ldots,N). \tag{2.33}$$

The total (unconstrained) phase space is $2(N+M)$-dimensional, and the canonical vari-

[7] Bifurcation simply means that more than one constraint surface exists that is associated with the same variable. Operationally, one chooses a bifurcation by specifying initial conditions.

ables are chosen as follows

$$Q_1 \equiv q \longleftrightarrow P_1 \equiv \sum_{j=1}^{N} (-\frac{d}{dt})^{j-1} \frac{\partial f}{\partial q^{(j)}} \tag{2.34}$$

$$\vdots$$

$$Q_i \equiv q^{(i-1)} \longleftrightarrow P_i \equiv \sum_{j=i}^{N} (-\frac{d}{dt})^{j-i} \frac{\partial f}{\partial q^{(j)}} \tag{2.35}$$

$$\vdots$$

$$Q_N \equiv q^{(N-1)} \longleftrightarrow P_N \equiv \frac{\partial f}{\partial q^{(N)}} \tag{2.36}$$

$$Q_{N+1} \equiv \lambda_1 \longleftrightarrow P_{N+1} \equiv P_{\lambda_1} = 0 \tag{2.37}$$

$$\vdots$$

$$Q_{N+M} \equiv \lambda_M \longleftrightarrow P_{N+M} \equiv P_{\lambda_M} = 0. \tag{2.38}$$

The non-degeneracy assumption means that we can solve for $q^{(N)}$ as a function of P_N and Q_i, i.e. $q^{(N)} = h(Q_1, \dots, Q_N, Q_{N+1}, \dots, Q_{N+M}, P_N)$. The total Hamiltonian takes the form

$$\begin{aligned} H_T = {} & P_1 Q_2 + \cdots + P_{N-1} Q_N + P_N h(Q_1, \dots, Q_{N+M}, P_N) \\ & - f(Q_1, \dots, Q_{N+M}, h) + u_a \varphi_a, \end{aligned} \tag{2.39}$$

where $\varphi_a : P_{N+a} = 0$ are M primary constraints, with $1 \leq a \leq M$. We use the consistency relation to find the associated secondary constraints

$$\tilde{\varphi}_a = [\varphi_a, H_T]_P : \left.\frac{\partial f}{\partial \lambda_a}\right|_{\lambda_a = Q_{N+a}} \approx 0. \tag{2.40}$$

If $[\varphi_a, \tilde{\varphi}_b]_P \not\approx 0$ for $1 \leq a, b \leq M$, both φ_a and $\tilde{\varphi}_b$ are second class constraints and thus there are no further constraints which can be generated by using consistency relations — we will consider in the next section the case when further constraints are present. We can reduce M pairs of canonical variables Q_{N+a}, P_{N+a} by using the constraints, i.e. $Q_{N+a} = F_a(Q_1, \dots, Q_N, P_N)$, $P_{N+a} = 0$ and the reduced Hamiltonian on the $2N$-dimensional phase space becomes

$$
\begin{aligned}
H_R = P_1 Q_2 + \cdots + P_{N-1} Q_N + P_N h(Q_1, \ldots, Q_N, F_a, \ldots, F_M, P_N) \\
- f(Q_1, \ldots, Q_N, h, F_a, \ldots, F_M),
\end{aligned} \tag{2.41}
$$

which is linearly dependent on $P_1, \ldots, P_{N-1}$ and thus necessarily unstable. Therefore, we conclude that the Ostrogradski instability survives if the auxiliary variables do not introduce enough constraints to reduce the dimensionality of the phase space. Since each auxiliary variable generates here only a pair of constraints, the dimensionality of the reduced phase space is the same as that for the original theory without constraints

$$
\text{Total } 2(N+M) - 2M \text{ Constraints} = 2N. \tag{2.42}
$$

An example of this case is considered in Section 2.3.2 above.

2.5 Exorcising the Ostrogradski ghost by reducing the dimensionality of the phase space

In the last section, we demonstrated that the Ostrogradski ghost[8] could not be exorcised unless the effective dimensionality is reduced. In this section, we demonstrate that such a reduction can render the theory stable. We will first introduce an example of higher derivative theory which is stabilized by the constraints, then demonstrate under which general conditions such stabilization can occur.

2.5.1 An example of stable non-degenerate higher derivative theory

Consider the following Lagrangian

$$
L = \frac{\dot{q}^2}{2} + \frac{(\ddot{q} - \lambda)^2}{2}, \tag{2.43}
$$

[8] I.e., instability.

which is a non-degenerate higher derivative Lagrangian but secretly stable, as we will now show. As usual, the canonical variables are defined by

$$Q_1 \equiv q \longleftrightarrow P_1 \equiv \dot{q} - q^{(3)} + \dot{\lambda} \tag{2.44}$$
$$Q_2 \equiv \dot{q} \longleftrightarrow P_2 \equiv \ddot{q} - \lambda \tag{2.45}$$
$$Q_3 \equiv \lambda \longleftrightarrow P_3 \equiv 0, \tag{2.46}$$

where $\varphi_1 : P_3 = 0$ is the primary constraint, and the total Hamiltonian is

$$H_T = P_1 Q_2 + P_2 Q_3 + \frac{P_2^2}{2} - \frac{Q_2^2}{2} + u_1 \varphi_1. \tag{2.47}$$

The secondary constraints are again generated by the consistency relation $\dot{\varphi}_i \equiv [\varphi_i, H_T] \approx 0$; the secondary constraints of the theory are thus

$$\varphi_2 : \quad -P_2 \approx 0, \tag{2.48}$$
$$\varphi_3 : \quad P_1 - Q_2 \approx 0, \tag{2.49}$$
$$\varphi_4 : -Q_3 - P_2 \approx 0. \tag{2.50}$$

One can check that all of the constraints are second class. Now, if we use (φ_1, φ_4) to reduce (Q_3, P_3) and use (φ_2, φ_3) to reduce (Q_2, P_2), the reduced Hamiltonian will become

$$H_R = P_1^2 - \frac{P_1^2}{2} = \frac{P_1^2}{2}, \tag{2.51}$$

which is bounded from below and hence is free of the ghost. One can see that the effective dimensionality of the phase space is reduced from four (Q_1, Q_2, P_1, P_2) to two (Q_1, P_1).

2.5.2 General conditions for the stabilization of a class of non-degenerate higher derivative theory

It turns out that the above procedure is not general — a willy-nilly reduction of the phase space does not necessary lead to a stable theory. In this section we will find the condition such that the ghost is removed.

Consider the most general second order derivative theory with an auxiliary field λ

$$L = A_{ij} \lambda^i \ddot{q}^j, \tag{2.52}$$

where A_{ij} are functions of q and $\dot{q}$. Note that, here, we have used subscripts, i.e. A_{ij}, to *label* the functions, and superscripts on variables, i.e. λ^i, to denote their *power*. We will use Einstein summation convention. In principle we can investigate the conditions for the stabilization of the theory with i, j run from 0 to arbitrary finite integers. However, for $j > 2$, the mapping from the configuration space to the phase space is a *many-to-one* mapping, we thus need the intial conditions to specify which branch we are studying, since all the consistent theory needs to be *one-to-one* mapping from the configuration space to the phase space. Similarly, we choose $i \leq 2$, in order to prevent the bifurcation of the constraint algorithm.[9] Therefore, for simplicity, we restrict ourselves to the case where the auxiliary fields are, at most, quadratic, $i, j = 0, 1, 2$, which guarantees a one to one mapping from the configuration space to the phase space without any intial conditions. Again, we follow Dirac's analysis of constrained systems, by defining the canonical variables

$$Q_1 \equiv q \longleftrightarrow P_1 \equiv \frac{\delta L}{\delta \dot{q}} = \frac{\partial L}{\partial \dot{q}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}} \tag{2.53}$$

$$Q_2 \equiv \dot{q} \longleftrightarrow P_2 \equiv A_{i1}\lambda^i + 2A_{i2}\lambda^i \ddot{q} \tag{2.54}$$

$$Q_3 \equiv \lambda \longleftrightarrow P_3 \equiv 0. \tag{2.55}$$

We can invert $\ddot{q}$ in R.H.S. of eq. (2.54) as a function of canonical variables

$$\ddot{q} \equiv h(Q_1, Q_2, Q_3, P_2) = \frac{P_2 - A_{i1}Q_3^i}{2A_{j2}Q_3^j}. \tag{2.56}$$

The total Hamiltonian thus becomes

$$\begin{aligned} H_T = {} & P_1 Q_2 + P_2 h(Q_1, Q_2, Q_3, P_2) + u_1 \varphi_1 \\ & - A_{ij}(Q_1, Q_2) Q_3^i h^j (Q_1, Q_2, Q_3, P_2), \end{aligned} \tag{2.57}$$

where $\varphi_1 : P_3 = 0$ is the primary constraint of this theory and generates a secondary constraint φ_2 by the consistency relation,

$$\begin{aligned} & \dot{\varphi}_1 \equiv [P_3, H_T]_P \approx 0 \\ \Rightarrow & -\left[P_2 - \sum_{k,l} l A_{kl} Q_3^k h^{l-1}\right] \frac{\partial h}{\partial Q_3} + \sum_{i,j} i A_{ij} Q_3^{i-1} h^j \approx 0 \end{aligned}$$

[9] Again, this can be fixed by the initial conditions

$$\Rightarrow \varphi_2 : \sum_{i,j} iA_{ij}Q_3^{i-1}h^j = A_{1j}h^j + 2A_{2j}Q_3h^j \approx 0. \tag{2.58}$$

From the second to the third weak equality, the coefficient of $\partial h/\partial Q_3$ vanishes, by virtue of eq. (2.54).

Rendering the theory stable requires a reduction in the dimensionality of the original phase space. To ensure this, the consistency relations must continue to generate constraints beyond the first pair, which algebraically requires φ_2 to be independent of Q_3. The stable theory hence needs to obey the condition $\partial\varphi_2/\partial Q_3 = 0$, i.e.

$$\frac{\partial\varphi_2}{\partial Q_3} = 2A_{2i}h^i + jkA_{jk}Q_3^{j-1}h^{k-1}\frac{\partial h}{\partial Q_3} = 0, \tag{2.59}$$

where

$$\frac{\partial h}{\partial Q_3} = -\frac{1}{2(A_{i2}Q_3^i)^2}[(A_{j2}Q_3^j)(A_{11}+2A_{21}Q_3) + (P_2 - A_{k1}Q_3^k)(A_{12}+2A_{22}Q_3)]. \tag{2.60}$$

From eqs. (2.59) and (2.60), one can see that $\partial\varphi_2/\partial Q_3$ is a quadratic function of P_2. In order to obtain vanishing $\partial\varphi_2/\partial Q_3$, we set the coefficients of P_2^0, P_2^1, and P_2^2 at zero. This leads to the following most general conditions on A_{ij} one can have with φ_2 independent of Q_3

$$A_{ij} = \begin{bmatrix} A & B & a \\ c & \pm\sqrt{4ab} & 0 \\ b & 0 & 0 \end{bmatrix},$$

where A, B, a, b, c are all functions of Q_1 and Q_2. Furthermore, the coefficient a is non-vanishing by construction or else the Lagrangian will not describe a higher derivative theory. The most general Lagrangian with more than two constraints can now be written as

$$L = A + B\ddot{q} + a\ddot{q}^2 + c\lambda + b\lambda^2 \pm \sqrt{4ab}\lambda\ddot{q}, \tag{2.61}$$

where all of the coefficients are functions of q and $\dot{q}$, and the “acceleration" $\ddot{q}$ can be inverted by the definition of canonical momentum P_2 using eq. (2.54)

$$\ddot{q} = h = \frac{P_2 - B \mp \sqrt{4ab}Q_3}{2a}. \tag{2.62}$$

The total Hamiltonian (2.57) and the secondary constraint φ_2 can now be rewritten as

$$H_T = P_1 Q_2 + P_2 h - A - Bh - ah^2 - cQ_3 - bQ_3^2 \mp \sqrt{4ab} Q_3 h + u_1 \varphi_1, \tag{2.63}$$

$$\varphi_2 : c \pm \sqrt{4ab}(\frac{P_2 - B}{2a}) \approx 0. \tag{2.64}$$

Since the instability comes from the linear term $P_1 Q_2$, to fix this instability, we need to generate a constraint whereby Q_2 must be some function of P_1. To develop a nontrivial theory, we need P_1 to enter the constraint equations either at φ_3 or φ_4.[10] We will now show that the latter condition will not lead to a stable theory, and then show the condition for the former leading to stability.

P_1 entering φ_4 does not lead to a stable theory

To pick up P_1 in the constraint φ_4 requires P_2 to be in φ_3 but not before, i.e. φ_2 has to be independent of P_2. This can be achieved by specifying $b = 0$ such that $\varphi_2 : c \approx 0$ and $h = (P_2 - B)/2a$. Using the consistency relation, φ_3 thus becomes

$$\varphi_3 : \frac{\partial c}{\partial Q_1} Q_2 + \frac{\partial c}{\partial Q_2} \frac{(P_2 - B)}{2a}. \tag{2.65}$$

If $\partial c / \partial Q_2 = 0$, we will be unable to pick up P_1 at φ_4, which means that the reduced Hamiltonian is either unstable (no constraint picks P_1 up) or trivial (theory with six constraints, all of the variables are constants). We thus require $\partial c / \partial Q_2 \neq 0$ in order to have a possibly stable theory, with P_1 appearing in φ_4. One can see that this requirement also applies for Q_3 to be in φ_4, since φ_4 can be generated again by the consistency relation

$$\varphi_4 : -\frac{\partial \varphi_3}{\partial P_2} \left(P_1 - \frac{\partial A}{\partial Q_2} - \frac{\partial B}{\partial Q_2} h - \frac{\partial a}{\partial Q_2} h^2 - \frac{\partial c}{\partial Q_2} Q_3 \right) + \frac{\partial \varphi_3}{\partial Q_1} Q_2 + \frac{\partial \varphi_3}{\partial Q_2} h \approx 0. \tag{2.66}$$

Using φ_1, φ_3, and φ_4 to eliminate P_3, P_2 and Q_3 and then substituting them into the total Hamiltonian (2.63), the semi-reduced Hamiltonian becomes

$$H_{SR} = F(Q_1, Q_2) + P_1 \left(Q_2 - \frac{c}{\frac{\partial c}{\partial Q_2}} \right). \tag{2.67}$$

[10] If P_1 enters the constraint equations φ_5 or φ_6, there will be six constraints, through which all of the canonical variables will be some constants, and thus a trivial theory.

If we substitute the last constraint $\varphi_2 = c \approx 0$ which relates Q_2 to some function of Q_1, we will have the final reduced Hamiltonian

$$H_R = F_1(Q_1) + P_1 F_2(Q_1), \tag{2.68}$$

where F_1, F_2 are functions of Q_1 only. It is clear that the final reduced Hamiltonian is always unstable unless $F_2 = 0$, implying $c = Q_2 F_3(Q_1)$, which means that the Lagrange multiplier constrains Q_1 to be a constant and the theory is thus trivial. We conclude that, if we wish P_1 to appear only in the constraint φ_4, the theory is either unstable or trivial.

P_1 entering φ_3 and the conditions for stability

Finally, we consider the case where P_1 enters at φ_3. This means that P_2 enters at φ_2 which requires that $b \neq 0$. Replacing h in the total Hamiltonian (2.63), we get

$$H_T = P_1 Q_2 + \frac{(P_2 - B \mp \sqrt{4ab} Q_3)^2}{4a} - A - cQ_3 - bQ_3^2 + u_1 \varphi_1, \tag{2.69}$$

which we can use to calculate the awkward looking φ_3

$$\begin{aligned} \varphi_3 :\ & \pm (P_2 - B)\left((P_2 - B)\frac{\partial b}{\partial Q_2} - 2bQ_2 \frac{\partial a}{\partial Q_1}\right) \\ & \pm a \left(2Q_2(P_2 - B)\frac{\partial b}{\partial Q_1} - 4b(P_1 - \frac{\partial A}{\partial Q_2} + Q_2 \frac{\partial B}{\partial Q_1})\right) \\ & + 2(P_2 - B)\sqrt{ab}\frac{\partial c}{\partial Q_2} + 4aQ_2\sqrt{ab}\frac{\partial c}{\partial Q_1} \approx 0, \end{aligned} \tag{2.70}$$

which is always independent of Q_3 and, because $a, b \neq 0$, we can use φ_3 to express P_1 as other canonical variables on the constraint surface,

$$\begin{aligned} P_1 \approx \pm & \left(\frac{c\, Q_2}{2\sqrt{ab}}\frac{\partial a}{\partial Q_1} - \frac{c\, Q_2\sqrt{ab}}{2b^2}\frac{\partial b}{\partial Q_1} + \frac{a\, Q_2}{\sqrt{ab}}\frac{\partial c}{\partial Q_1}\right) \\ & + \frac{\partial A}{\partial Q_2} - Q_2 \frac{\partial B}{\partial Q_1} + \frac{c^2}{4b^2}\frac{\partial b}{\partial Q_2} - \frac{c}{2b}\frac{\partial c}{\partial Q_2}. \end{aligned} \tag{2.71}$$

If we use φ_1 and φ_2 to eliminate P_3 and P_2 in the total Hamiltonian, we can write the semi-reduced Hamiltonian as

$$H_{SR} = P_1 Q_2 + \frac{c^2}{4b} - A. \tag{2.72}$$

The last step in finding a stable reduced Hamiltonian for the physical degrees of freedom is to reverse eq. (2.71) as $Q_2 = g(Q_1, P_1)$ and substitute it into eq. (2.72). Since there are five arbitrary functions (A, B, a, b, and c), we simply have to choose them as functions of q and $\dot{q}$ such that the reduced Hamiltonian is stable. For example, in Section 2.5.1, we chose $A = Q_2^2/2$, $B = c = 0$, and $a = b = 1/2$.

2.6 An example of a stable constrained non-degenerate Pais-Uhlenbeck oscillator

In this section, we demonstrate how to use this procedure to stabilize the Pais-Uhlenbeck model by introducing constraints in such a way that the dimensionality of the phase space is reduced. We consider the Lagrangian of the Pais-Uhlenbeck model with the auxiliary field λ

$$\begin{aligned} L = &\frac{\gamma}{2}[\ddot{q}^2 - (w_1^2 + w_2^2)\dot{q}^2 + w_1^2 w_2^2 q^2] \\ &+ 2\gamma\lambda^2 q^2 + \gamma(w_1 + w_2)^2 \lambda q^2 + 2\gamma\lambda q\ddot{q}, \end{aligned} \tag{2.73}$$

and the canonical variables are chosen by

$$Q_1 \equiv q \longleftrightarrow P_1 \equiv -\gamma[q^{(3)} + 2(\lambda\dot{q} + \dot{\lambda}q)] - \gamma(w_1^2 + w_2^2)\dot{q} \tag{2.74}$$

$$Q_2 \equiv \dot{q} \longleftrightarrow P_2 \equiv \gamma\ddot{q} + 2\gamma\lambda q \tag{2.75}$$

$$Q_3 \equiv \lambda \longleftrightarrow P_3 \equiv 0, \tag{2.76}$$

where the primary constraint is $\varphi_1 : P_3 = 0$, and the total Hamiltonian is

$$\begin{aligned} H_T = &P_1 Q_2 + \frac{P_2^2}{2\gamma} - \frac{\gamma}{2} w_1^2 w_2^2 Q_1^2 + \frac{\gamma}{2}(w_1^2 + w_2^2) Q_2^2 \\ &- \gamma(w_1 + w_2)^2 Q_1^2 Q_3 - 2Q_1 Q_3 P_2 + u_1 \varphi_1. \end{aligned} \tag{2.77}$$

The secondary constraints are generated by the consistency relation

$$\varphi_2 : \qquad \gamma(w_1 + w_2)^2 Q_1 + 2P_2 \approx 0 \tag{2.78}$$

$$\varphi_3 : \qquad \gamma(w_1 - w_2)^2 Q_2 + 2P_1 \approx 0 \tag{2.79}$$

$$\varphi_4 : 16 w_1 w_2 Q_3 - (w_1^4 - 6w_1^2 w_2^2 + w_2^4) \approx 0. \tag{2.80}$$

One can check that all of the constraints are second class. Now, if we use (φ_1, φ_4) to reduce (Q_3, P_3) and use (φ_2, φ_3) to reduce (Q_2, P_2), the reduced Hamiltonian will become

$$H_R = \frac{4w_1w_2}{(w_1-w_2)^4\gamma}P_1^2 + \frac{\gamma}{8}(w_1^2+w_2^2)(w_1^2+4w_1w_2+w_2^2)Q_1^2,$$

which is positive definite. One can see that the effective dimensionality of the phase space is reduced from four (Q_1, Q_2, P_1, P_2) to two (Q_1, P_1).

2.7 Conclusion

In this chapter, we proved that the linear instability, i.e. Ostrogradski ghost, in a non-degenerate higher derivative theory can be exorcised by the addition of constraints, at the cost of reducing the dimensionality of the phase space. We show this procedure in a class of second order time derivative theories with one Lagrange multiplier to illustrate how this is possible in principle. Generalization to arbitrary higher order derivative theory with multiple Lagrange multipliers is straightforward and we will explore how it may be generalized to higher derivative gravity in the next chapter.

Chapter 3

Stabilization of Linear Higher Derivative Gravity with Constraints

As emphasized in the two preceding chapters, all of the non-degenerate higher derivative theories suffer from the Ostrogradski instability. In this chapter, we will investigate the Ostrogradski instability of higher derivative gravity models with quadratic curvature invariant $\alpha R^2 + \beta R_{\mu\nu}R^{\mu\nu}$ perturbatively and show that the instability can be removed by the judicious addition of constraints at the quadratic level of metric fluctuations around Minkowski/de Sitter background. By making a suitable parameter choice, we find that the instability of the helicity-0, 1, and 2 modes can be removed while reducing the dimensionality of the original phase space. To retain the renormalization properties of higher derivative gravity, the Lorentz symmetry in the constrained theory is explicitly broken.

3.1 Introduction

We have seen that every *non-degenerate* higher derivative theory suffers from the Ostrogradski instability (see Chapter 2 and the references [1, 41, 48, 58, 66]).

In this chapter, we will consider the following action first investigated by Stelle [14][1]

$$S = \frac{M_P^2}{2}\int d^4x\,\sqrt{-g}(R - 2\Lambda + \alpha R^2 + \beta R_{\mu\nu}R^{\mu\nu}). \tag{3.1}$$

This action with mass dimension -2 parameters α and β in general contains eight degrees of freedom [15]. By expanding it around Minkowski spacetime, we can see that two of which

[1] Here we have turned on the bare cosmological constant since the theory admits a constant curvature background solution with $R_{\mu\nu} = \Lambda g_{\mu\nu}$.

correspond to the massless graviton in general relativity, five of which correspond to the massive graviton, and the last one is a massive scalar. As we will see later in section 3.3.1, the helicity-2 sector is a non-degenerate higher derivative theory and thus suffers from the Ostrogradski instability.

Nevertheless, this action is interesting, as it is power-counting renormalizable [14] – the presence of higher derivative terms in the action means that higher *spatial* derivatives exist in the propagator of the graviton modes. These spatial derivatives suppress the UV divergences in the loops, rendering the theory naively renormalizable. The price we pay for this is the presence of the higher *time* derivative terms which leads to the Ostrogradski instability.

One way to take advantage of this insight is to impose different scaling dimensions on the time and space coordinates – a stratagem utilized by Hořava [75–80]. The low energy limit of this theory is then a generic first order time derivative graviton action with higher order Lorentz violating spatial derivative terms, which is both stable and power-counting renormalizable.

In this chapter, we pursue a different track, and ask whether we can selectively remove the linear instability by imposing constraints on the theory. This idea is motivated by the proof outlined in Chapter 2 [66], that the linearly unstable phase space can be excised from the theory through the judicious choice of additional constraints (i.e. the final dimensionality of the phase space will be smaller). We will show that, at least in the linear theory, we can stabilize the theory with additional constraint terms, while simultaneously preserving its improved renormalizable features. Roughly speaking, we add a constraint whereby the higher timelike derivative terms in the equation of motion are constrained to some lower timelike derivative or higher order spacelike derivative term, i.e.

$$g^{(4)} \sim \partial^2 \ddot{g}, \partial^4 g, \cdots . \tag{3.2}$$

We will show that the final form of this constrained theory is, at least linearly, that of a second order time derivative equation of motion with higher order spatial derivatives that are very similar in spirit to the Hořava model, at the price of relinquishing the Lorentz invariance. Such an artificial addition of constraints changes the dynamics and thus the general theory. However, as we have simply worked in linear theory, we are unaware of what the non-linear completion of the theory would be. One possibility which we plan to explore in future work is that the full non-linear theory suffers from no such technical issues or at least that they appear in a more natural way.

This chapter is structured as follows. In Section 3.2, we show how to perturb the action up to second order in metric perturbation in the general background, which will be used in the Minkowski/de Sitter backgrounds. In Section 3.3, we obtain the action quadratic in the metric fluctuation by parameterizing the metric fluctuation around the Minkowski background. Since, up to the quadratic order, the action can be separated into the helicity-0, 1, and 2 sectors, we demonstrate how the instability appears in each sector. In Section 3.4, we show how the helicity-0, 1, and 2 instability can be rendered stable by introducing suitable constraints. In Sections 3.5 and 3.6, we study the behavior and how to remove the instability in the de Sitter background. We hasten to add that we have chosen *different* constraints for the two different backgrounds considered in this work (Minkowski and de Sitter) – while it is possible to stabilize the theory with the same constraints, in the helicity-0 mode in the de Sitter background, we encounter the difficulty that the resultant action is non-local. This non-locality could be a pathology that arises because we are dealing with linear theory and may be resolved when a full non-linear completion is obtained, but we will postpone its resolution for future work. We conclude and conjecture a possible way of making the procedure viable in the full non-linear theory in Section 3.7.

3.2 Higher derivative gravity: quadratic action

In order to study how the instability appears in action (3.1) at the quadratic order in the metric fluctuation, it is necessary to expand every curvature invariant up to the second order in the metric perturbation $h_{\mu\nu}$, which is defined by

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}, \tag{3.3}$$

where $\bar{g}_{\mu\nu}$ at this stage can be a general background metric and $h_{\mu\nu} \ll \bar{g}_{\mu\nu}$ [81]. The inverse metric up to the second order in h can be written as

$$g^{\mu\nu} = \bar{g}^{\mu\nu} - h^{\mu\nu} + h^{\mu\rho} h^{\nu}_{\rho} + O\left(h^3\right). \tag{3.4}$$

Assuming a constant curvature background of either Minkowski ($\Lambda = 0$), de Sitter ($\Lambda > 0$), or Anti-de Sitter ($\Lambda < 0$), we compute the second order action

$$S = -\frac{M_P^2}{4} \int d^4x \sqrt{-\bar{g}} h^{\mu\nu} \left[(1 + 8\alpha\Lambda + \frac{4}{3}\beta\Lambda)\mathcal{G}^L_{\mu\nu} + \beta \left(\Box \mathcal{G}^L_{\mu\nu} - \frac{2\Lambda}{3} \bar{g}_{\mu\nu} R_L \right) \right.$$

$$+(\beta+2\alpha)(\bar{g}_{\mu\nu}\Box-\bar{\nabla}_\mu\bar{\nabla}_\nu+\Lambda\bar{g}_{\mu\nu})R_L\Big], \tag{3.5}$$

where $\Box$ is the d'Alembert operator and the linearized Ricci tensor, Ricci scalar, and Einstein tensor are defined by[2]

$$\begin{aligned}
R^L_{\mu\nu} &= \frac{1}{2}(\bar{\nabla}_\rho\bar{\nabla}_\mu h^\rho_\nu+\bar{\nabla}_\rho\bar{\nabla}_\nu h^\rho_\mu-\Box h_{\mu\nu}-\bar{\nabla}_\mu\bar{\nabla}_\nu h),\\
R_L &= \bar{g}^{\mu\nu}R^L_{\mu\nu}-\bar{R}^{\mu\nu}h_{\mu\nu},\\
\mathcal{G}^L_{\mu\nu} &= R^L_{\mu\nu}-\frac{1}{2}\bar{g}_{\mu\nu}R_L-\Lambda h_{\mu\nu}.
\end{aligned} \tag{3.6}$$

Note that the indices are raised and lowered by the background metric $\bar{g}_{\mu\nu}$.

3.3 Quadratic action around the Minkowski background

In this section, we will study how the instability appears in the action at the quadratic level of perturbation around Minkowski background ($\Lambda=0$). We parameterize the metric fluctuation by

$$ds^2=-(1+2\phi)dt^2+2B_idx^idt+[(1-2\psi)\delta_{ij}+2E_{ij}]dx^idx^j, \tag{3.7}$$

where E_{ij} is a symmetric, traceless tensor and the index i,j are raised and lowered by δ_{ij}. We can further decompose B_i and E_{ij} into the helicity-0, 1, and 2 modes

$$B_i=\partial_iB+B^{\mathrm{T}}_i \tag{3.8}$$

$$E_{ij}=\partial_{\langle i}\partial_{j\rangle}E+\partial_{(i}E^{\mathrm{T}}_{j)}+E^{\mathrm{TT}}_{ij}, \tag{3.9}$$

where B and B^{T}_i are longitudinal and transverse parts of vector B_i, E^{T}_i is transverse, and E^{TT}_{ij} is symmetric, trace-free and transverse, and the angled bracket indices component

$$\partial_{\langle i}\partial_{j\rangle}E\equiv\partial_i\partial_jE-\frac{1}{3}\delta_{ij}\nabla^2E \tag{3.10}$$

is trace-free. Through this decomposition, we can separate the action into the helicity-0, 1, and 2 sectors since, at the quadratic level, there is no mixing between different helicities.

[2] see, for example, [82]

3.3.1 Helicity-2 sector

The second order action of the helicity-2 sector is

$$S = \frac{M_P^2}{2}\int d^4x\ \left\{\beta\left[(\ddot{E}_{ij}^{\mathrm{TT}})^2 + 2\dot{E}^{\mathrm{TT}ij}\nabla^2\dot{E}_{ij}^{\mathrm{TT}} + (\nabla^2 E_{ij}^{\mathrm{TT}})^2\right] + (\dot{E}_{ij}^{\mathrm{TT}})^2 + E^{\mathrm{TT}ij}\nabla^2 E_{ij}^{\mathrm{TT}}\right\}, \tag{3.11}$$

which describes two massless helicity-2 degrees of freedom originating from the massless graviton and two massive helicity-2 degrees of freedom emanating from the quadratic invariant term $\beta R_{\mu\nu}R^{\mu\nu}$. Since there is no first class (i.e. gauge) constraint in the helicity-2 modes, there are four helicity-2 degrees of freedom in the theory. Note that only the β term contributes to the helicity-2 sector, since αR^2 is a class of $f(R)$ gravity and can be written as an extra scalar field which only contributes to the helicity-0 sector.

Ostrogradski's choice of canonical coordinates is the pair of canonical variables (E_{ij}, π_{ij}) and (q_{ij}, p_{ij}), defined by

$$\begin{aligned} E_{ij} &\equiv E_{ij}^{\mathrm{TT}} \longleftrightarrow \pi^{ij} = 2\dot{E}^{\mathrm{TT}ij} + \beta(-2\dddot{E}^{\mathrm{TT}ij} + 4\nabla^2\dot{E}^{\mathrm{TT}ij}) \\ q_{ij} &\equiv \dot{E}_{ij}^{\mathrm{TT}} \longleftrightarrow p^{ij} = 2\beta\ddot{E}^{\mathrm{TT}ij}. \end{aligned} \tag{3.12}$$

One might notice that, in Ostrogradski's formalism, the two canonical variables E_{ij}, q_{ij} have different dimensionalities: the field E_{ij} is dimensionless while q_{ij} has a mass dimension of 1 and thus the dimension of canonical momenta are different. The dimensionality is not particularly important – in principle, one can rescale $q_{ij} = M_P^{-1}\dot{E}_{ij}$ to place the two canonical variables on the same footing.

Using the Legendre transform, we construct the Hamiltonian following the usual method

$$H = \frac{M_P^2}{2}\int d^3x\ \left(\frac{p^{ij}p_{ij}}{4\beta} + \pi^{ij}q_{ij} - 2\beta q^{ij}\nabla^2 q_{ij} - q^{ij}q_{ij} - \beta\nabla^2 E^{ij}\nabla^2 E_{ij} - E^{ij}\nabla^2 E_{ij}\right). \tag{3.13}$$

It is easy to check that the Hamiltonian (3.13) generates the equations of motion for the 4 canonical variables via the Poisson Bracket $d(\cdot)/dt = [\cdot, H]$. It is important to note that the Hamiltonian is linearly dependent on π^{ij} in the second term and hence unbounded from below – the $\pi^{ij}q_{ij}$ term can be arbitrarily negative when $q_{ij} > 0$, $\pi^{ij} \to -\infty$ or vice versa.

As noted in Chapter 2, this instability is a ghost. In order to see this, we can explicitly diagonalize the Hamiltonian by the following canonical transformation

$$\begin{aligned}
\psi_{ij} &= \sqrt{2}\left(\frac{p_{ij}}{2} - \beta\nabla^2 E_{ij}\right) \\
\phi_{ij} &= \sqrt{2}\left(-\frac{p_{ij}}{2} + \beta\nabla^2 E_{ij} + E_{ij}\right) \\
p_\psi^{ij} &= \frac{1}{\sqrt{2}}\left(\pi^{ij} - 2\beta\nabla^2 q^{ij} - 2q^{ij}\right) \\
p_\phi^{ij} &= \frac{1}{\sqrt{2}}\left(\pi^{ij} - 2\beta\nabla^2 q^{ij}\right),
\end{aligned} \tag{3.14}$$

in which case the Hamiltonian becomes

$$H = \frac{M_P^2}{2}\int d^3x \left[\frac{p_{\phi ij}p_\phi^{ij}}{2} - \frac{\phi_{ij}\nabla^2\phi^{ij}}{2} - \left(\frac{p_{\psi ij}p_\psi^{ij}}{2} - \frac{\psi_{ij}\nabla^2\psi^{ij}}{2} - \frac{\psi_{ij}\psi^{ij}}{2\beta}\right)\right], \tag{3.15}$$

where the (ψ, p_ψ) pair is ghostlike. In the classical theory, the Hamiltonian is clearly unbounded from below since the arbitrary choice of (ψ, p_ψ) in the phase space renders the second term in the Hamiltonian (3.15) arbitrarily negative.[3] In quantum theory, while this instability does not prevent us from identifying a vacuum state and then constructing the Fock space of many particle states, the imposition of positivity in the energy of all of the particle states will lead to some states possessing negative norms, i.e. ghosts. One can further excise these unphysical negative norm states from the Fock space, but this generically leads to violations of unitarity. For a review of the quantization issues associated with such theories, see Appendix B.

3.3.2 Helicity-1 sector

The second order action of the helicity-1 modes can be written by the gauge invariant variable $v_i = \sqrt{-\nabla^2}(B_i^{\mathrm{T}} - \dot{E}_i^{\mathrm{T}})$[4]

$$S = \frac{M_P^2}{2}\int d^4x\, \frac{\beta}{2}\left(\dot{v}_i\dot{v}^i + v_i\nabla^2 v^i + \frac{1}{\beta}v_i v^i\right). \tag{3.16}$$

[3] The choice $\beta < 0$ corresponds to the oscillating ghost while $\beta > 0$ corresponds to the exponential ghost. See Appendices A and B on the difference between the tachyon and the ghost quantum mechanically.

[4] One should not be unduly worried by the appearance of the non-local square root of the Laplace operator. Recall that the Laplace operator $-\nabla^2$ has zero or positive eigenvalues λ_k, e.g. $-\nabla^2\phi_k = \lambda_k\phi_k$ with $\lambda \geq 0$. Formally, $\sqrt{-\nabla^2}u = \sum_k c_k\lambda_k^{1/2}\phi_k$ (as long as both u and ϕ_k vanish at the boundary), i.e. $u = \sum_k c_k\phi_k$.

The action describes a vector with mass $m_2^2 = -\frac{1}{\beta}$ and the sign of β also decides the overall sign of the action, i.e. if $\beta < 0$, the helicity-1 modes are ghostlike. The Euler-Lagrange equation of action (3.16) is

$$\left[\beta\left(\frac{d^2}{dt^2} - \nabla^2\right) - 1\right] v_i = 0, \tag{3.17}$$

which can be solved by the Fourier transform, and the solutions are harmonic oscillators with frequency $w_{\mathbf{p}}^2 = \mathbf{p}^2 - \frac{1}{\beta}$. The canonical momentum conjugate to v_i is, as usual, defined by

$$p_{vi} = \frac{\delta S}{\delta \dot{v}_i} = \beta \dot{v}_i, \tag{3.18}$$

and, since we use the gauge invariant variable to write the action, there is no constraint in the helicity-1 sector and the Hamiltonian is

$$H = \frac{M_P^2}{2}\int d^3x \left(\frac{p_{vi}p_v^i}{2\beta} - \frac{\beta}{2} v_i \nabla^2 v^i - \frac{1}{2} v_i v^i\right). \tag{3.19}$$

If we choose $\beta > 0$, then $m_2^2 < 0$, which means that the theory is tachyonic. On the other hand, if we choose $\beta < 0$ in eq. (3.19), the Hamiltonian will be negative definite and thus ghostlike. One can see that, if $\beta < 0$, we can perform a canonical transformation of the variables into "canonically normalized" form $\sqrt{-\beta} v_i \to v_i$, $(-\beta)^{-\frac{1}{2}} p_{vi} \to p_{vi}$ with the Hamiltonian

$$H = \frac{M_P^2}{2}\int d^3x \left(-\frac{p_{vi}p_v^i}{2} + \frac{1}{2} v_i \nabla^2 v^i + \frac{1}{2\beta} v_i v^i\right), \tag{3.20}$$

where the mass of the helicity-1 ghost is $m_2^2 = -\frac{1}{\beta}$. In summary, the helicity-1 modes are either tachyonic ($\beta > 0$) or ghostlike ($\beta < 0$).

3.3.3 Helicity-0 sector

The second order action for the helicity-0 modes is more complicated. With the help of two gauge invariant variables

$$\begin{aligned} \Phi &= \phi + \dot{B} - \ddot{E}, \\ \Psi &= \psi + \frac{1}{3}\nabla^2 E, \end{aligned} \tag{3.21}$$

the action can be written as

$$S = \frac{M_P^2}{2}\int d^4x\ \big[(-6\dot{\Psi}^2 - 2\Psi\nabla^2\Psi + 4\Psi\nabla^2\Phi) + 4(\beta+3\alpha)(3\ddot{\Psi}^2 + 4\dot{\Psi}\nabla^2\dot{\Psi} + 2\ddot{\Psi}\nabla^2\Phi) \\ + 2(3\beta+8\alpha)(\nabla^2\Psi)^2 + 2(\beta+2\alpha)(\nabla^2\Phi)^2 - 4(\beta+4\alpha)\nabla^2\Psi\nabla^2\Phi\big]\,. \tag{3.22}$$

There are two scalar functions in the action and, because of the second order time derivatives on Ψ, there are naively three degrees of freedom.[5] One degree of freedom will eventually be removed by a gauge constraint and the helicity-0 modes sector in general consists of two degrees of freedom. Note that all of the second order time derivatives appear on the first line with the coefficient $(\beta+3\alpha)$ – it is the well-known fact [14] that, if we choose $\beta+3\alpha=0$, the massive scalar will be frozen and removed from the theory because of its infinite mass. The only degree of freedom in this sector is the helicity-0 mode of massive graviton.

On the other hand, we know that $\beta=0$ is simply an $f(R)$ type theory which is degenerate and hence also ghost-free – this fact is not manifest in eq. (3.22) above, if we simply set $\beta=0$. However, when $\beta=0$, the action can be rearranged as

$$S = \frac{M_P^2}{2}\int d^4x\ \big[(-6\dot{\Psi}^2 - 2\Psi\nabla^2\Psi + 4\Psi\nabla^2\Phi) + 4\alpha(3\ddot{\Psi} - 2\nabla^2\Psi + \nabla^2\Phi)^2\big]\,, \tag{3.23}$$

where we have suggestively written the second term in the action (3.23) as a complete square. By varying Φ, we obtain

$$\nabla^2\Phi = \left(-\frac{1}{2\alpha} + 2\nabla^2\right)\Psi - 3\ddot{\Psi}. \tag{3.24}$$

Inserting eq. (3.24) back into the action (3.23), we obtain the action of a single non-ghostlike massive scalar field

$$S = \frac{M_P^2}{2}\int d^4x\ \left(6\dot{\Psi}^2 + 6\Psi\nabla^2\Psi - \frac{2}{\alpha}\Psi^2\right), \tag{3.25}$$

as we would expect for $f(R)$ type theories. It is clear that, since the action is only dependent on Ψ and $\dot{\Psi}$, there is only one ghost-free degree of freedom. Notice that, if $\alpha<0$, this scalar is a tachyonic unstable degree of freedom, which is consistent with the general $f(R)$ gravity

[5] As in eq. (2.3), an extra time derivative in the action will generate two further dimensions of the phase space, i.e., one more degree of freedom.

theory, approaching general relativity limit at high curvatures, where we require $f''(R) > 0$ to avoid tachyonic instability [4, 20, 21, 83]. By setting $\alpha \to 0$, the mass term blows up and thus renders this degree of freedom non-dynamical, i.e. it reduces to simple General Relativity.

Harking back to the action for general α and β, eq. (3.22), Ostrogradski's choice of canonical coordinates is

$$\begin{aligned}
\Phi &\equiv \Phi \longleftrightarrow p_\Phi \equiv 0 \\
\Psi &\equiv \Psi \longleftrightarrow p_\Psi \equiv \frac{\delta S}{\delta \dot{\Psi}} \\
\chi &\equiv \dot{\Psi} \longleftrightarrow p_\chi \equiv 8(\beta+3\alpha)(3\ddot{\Psi}+\nabla^2\Phi),
\end{aligned} \tag{3.26}$$

where the choice $\beta + 3\alpha = 0$ means that $p_\chi = 0$ becomes a *primary* constraint instead of an additional degree of freedom.

The Hamiltonian can be expressed by the canonical coordinates

$$\begin{aligned}
H = \frac{M_P^2}{2}\int d^3x \Bigg[& p_\Psi \chi + \frac{p_\chi^2}{48(\beta+3\alpha)} - \frac{p_\chi \nabla^2\Phi}{3} + (6\chi^2 + 2\Psi\nabla^2\Psi - 4\Psi\nabla^2\Phi) \\
& - 16(\beta+3\alpha)\chi\nabla^2\chi - 2(3\beta+8\alpha)(\nabla^2\Psi)^2 + 4(\beta+4\alpha)\nabla^2\Psi\nabla^2\Phi - \frac{2\beta}{3}(\nabla^2\Phi)^2 \Bigg].
\end{aligned} \tag{3.27}$$

The primary constraint is $\varphi_1 : p_\Phi = 0$ and all of the constraints can be generated by the consistency relation

$$\varphi_2 : \nabla^2\left(\frac{p_\chi}{3} + 4\Psi - 4(\beta+4\alpha)\nabla^2\Psi + \frac{4\beta}{3}\nabla^2\Phi\right) \approx 0, \tag{3.28}$$

where $\approx$ means "weak equality" (i.e. the equality is numerically restricted to be satisfied but not identically valid throughout the whole phase space.) – see [23] for a discussion on this point.

Since φ_1, φ_2 are second class,[6] we can use them to reduce the phase space (Φ, p_Φ), and

[6] Note that, in the case of $\beta = 0$, the constraints φ_1 and φ_2 are not second class and the theory will contain two more constraints. The reduced phase space is then two-dimensional and the Hamiltonian is bounded below if $\alpha > 0$, as in the conclusion of the full $f(R)$ theory.

the reduced Hamiltonian is

$$\begin{aligned}
H_R = \frac{M_P^2}{2}\int d^3x \Big\{ & p_\Psi \chi + \frac{(\beta+2\alpha)}{16\beta(\beta+3\alpha)} p_\chi^2 + \frac{1}{\beta} p_\chi [1-(\beta+4\alpha)\nabla^2]\Psi \\
& + \frac{6}{\beta}\Psi^2 - (10+\frac{48\alpha}{\beta})\Psi\nabla^2\Psi + \frac{32\alpha(\beta+3\alpha)}{\beta}(\nabla^2\Psi)^2 \\
& + 6\chi^2 - 16(\beta+3\alpha)\chi\nabla^2\chi \Big\}. \qquad (3.29)
\end{aligned}$$

The linear dependence of p_Ψ again renders the Hamiltonian unbounded from below.

In order to see the mass content of the helicity-0 modes, we will need to further diagonalize the Hamiltonian by the following canonical transformation

$$\begin{aligned}
Q_1 &= \sqrt{3}\left(\frac{p_\chi}{6} - \frac{8(\beta+3\alpha)}{3}\nabla^2\Psi + 2\Psi\right), \\
Q_2 &= \sqrt{3}\left(\frac{p_\chi}{6} - \frac{8(\beta+3\alpha)}{3}\nabla^2\Psi\right), \\
P_1 &= \frac{1}{\sqrt{3}}\left(\frac{p_\Psi}{2} - 8(\beta+3\alpha)\nabla^2\chi\right), \\
P_2 &= \frac{1}{\sqrt{3}}\left(-\frac{p_\Psi}{2} + 8(\beta+3\alpha)\nabla^2\chi - 6\chi\right). \qquad (3.30)
\end{aligned}$$

The diagonalized Hamiltonian is then

$$\begin{aligned}
H_R = \frac{M_P^2}{2}\int d^3x \Big[& -\frac{P_1^2}{2} + \frac{1}{2}Q_1\nabla^2 Q_1 + \frac{1}{2\beta}Q_1^2 \\
& + \frac{P_2^2}{2} - \frac{1}{2}Q_2\nabla^2 Q_2 + \frac{1}{2}\frac{1}{2(\beta+3\alpha)}Q_2^2 \Big]. \qquad (3.31)
\end{aligned}$$

The reduced Hamiltonian of the helicity-0 sector contains two massive degrees of freedom. One is a massive ghost coming from massive graviton with mass $m_2^{-2} = -\beta$ and the other is massive scalar with positive definite kinetic energy, with mass $m_0^2 = \frac{1}{2(\beta+3\alpha)}$.[7]

Let us combine the results from all sectors. In Section 3.3.1, we saw that there are four helicity-2 degrees of freedom, two of which suffer from ghostlike instability. In Section 3.3.2, the two helicity-1 degrees of freedom are either ghostlike or tachyonic, depending on the sign of β. In Section 3.3.3, one of the two scalar degrees of freedom is ghostlike. With $\beta < 0$, one can see that the unstable modes in the helicity-0, 1, and 2 sectors are massive with mass $m_2^2 = -\frac{1}{\beta}$, which corresponds to the massive graviton. This result is de-

[7] Again, the sign of β would change (Q_1, P_1) from a massive oscillating ghost to an exponential ghost, while the sign of $(\beta+3\alpha)$ would change (Q_2, P_2) from the normal mode to the tachyonic mode.

rived by Stelle in his seminal work on higher derivative gravity [15] using an auxiliary field methodology. Here, we rederive the results using the usual Hamiltonian formalism and the equivalence between the two formalisms is given in Appendix C.

There are two special choices of parameter in the linearized theory. With $\alpha \neq 0, \beta = 0$, the massive graviton sector gains an infinite mass and hence becomes non-dynamical. In this case, the theory consists of one massless graviton with one massive scalar field (i.e. an $f(R)$ theory). On the other hand, by taking the limit $\beta + 3\alpha = 0$, the massive scalar field becomes infinitely massive and hence non-dynamical. In this case, the theory's particle content is reduced to one massive and one massless graviton. With the latter choice and a total minus sign, at the linear level, one can have a theory with a healthy massive graviton [65], since this choice is consistent with the Fierz-Pauli tuning. However, one should expect that the Boulware-Deser ghost [84] would enter at the nonlinear level.

3.4 Stabilization by constraints in the Minkowski background

In this section, we will demonstrate how to remove the unstable degrees of freedom by introducing constraints via auxiliary fields. As shown in [66], this will result in the effective dimensionality of the phase space being reduced. Roughly speaking, we impose the constraints such that the auxiliary fields are related to the second order time derivative of the unstable fields, resulting in the final equations of motion being second order in time derivatives yet up to fourth order in spatial derivatives. The advantage of preserving the spatial part of the "higher derivative" component is that we retain the improved renormalization properties of such theories, at the price of relinquishing the Lorentz invariance.

One might ask: what if we remove the instability without explicitly breaking the Lorentz invariance? Here, we emphasize that we can insert constraints to remove the higher spatial as well as the higher time derivatives by treating them on an equal basis, with the end result being a stable second order theory in both space and time derivatives. For example, the unconstrained helicity-2 action (3.11) can be written as

$$S = \frac{M_P^2}{2} \int d^4x \, \left[\beta(\Box E_{ij}^{\mathrm{TT}})^2 + E^{\mathrm{TT}ij} \Box E_{ij}^{\mathrm{TT}}\right].$$

Without the full theory, we do not know how to impose constraints by introducing the auxiliary field λ into the action without breaking the Lorentz invariance while removing the highest time derivative in the equations of motion. The best thing we can do is to couple λ_{ij} with $\Box E_{ij}$, and the Lorentz invariance is not explicitly broken by extra terms. We can modify the action as

$$S = \frac{M_P^2}{2}\int d^4x\ \left[\beta(\Box E_{ij}^{\mathrm{TT}} - \lambda_{ij})^2 + E^{\mathrm{TT}ij}(\Box E_{ij}^{\mathrm{TT}} - a\lambda_{ij})\right].$$

The choice $a=1$ corresponds to forcing λ_{ij} coupling to every $\Box E_{ij}$ while the choice $a=0$ corresponds to forcing λ_{ij} only coupling to those $\Box E_{ij}$ where the $\Box$ cannot be removed through integration by parts. The equations of motion of the theory are

$$\begin{aligned}
&\delta\lambda : 2\beta(\Box E_{ij}^{\mathrm{TT}} - \lambda_{ij}^{\mathrm{TT}}) + aE_{ij}^{\mathrm{TT}} = 0,\\
&\delta E : 2\beta\Box(\Box E_{ij}^{\mathrm{TT}} - \lambda_{ij}) + (\Box E_{ij}^{\mathrm{TT}} - a\lambda_{ij}^{\mathrm{TT}}) + \Box E_{ij}^{\mathrm{TT}} = 0,
\end{aligned}$$

which can be written as a single equation of E_{ij}

$$2(1-a)\Box E_{ij}^{\mathrm{TT}} - \frac{a^2}{2\beta}E_{ij}^{\mathrm{TT}} = 0.$$

The solution to the equation of motion is either trivial if $a=1$ or a travelling wave solution to the Klein-Gordon equation with mass $m^2 = a^2/4\beta(1-a)$ if $a \neq 1$. In both cases, the equations of motion will have the same order of time derivatives and spatial derivatives and the improved renormalization properties will not be retained.

For notational simplicity, from now on we drop the traceless notation B_i^{T}, E^{TT}, which should be clear from the context.

3.4.1 Helicity-2 sector

We begin by introducing a helicity-2 auxiliary tensor field λ_{ij} into the action (3.11)

$$\begin{aligned}
S = \frac{M_P^2}{2}\int d^4x\ \{&\beta[(\ddot{E}_{ij} - \lambda_{ij})^2 + 2\dot{E}^{ij}\nabla^2\dot{E}_{ij} + (\nabla^2 E_{ij})^2]\\
&+\dot{E}^{ij}\dot{E}_{ij} + E^{ij}\nabla^2 E_{ij} + 4\beta\lambda^{ij}\nabla^2 E_{ij}\},
\end{aligned} \tag{3.32}$$

where λ_{ij} is transverse traceless, which also explicitly breaks Lorentz invariance. The canonical coordinates are

$$E_{ij} \equiv E_{ij} \longleftrightarrow \pi^{ij} = 2\dot{E}^{ij} + \beta(-2\dddot{E}^{ij} + 2\dot{\lambda}^{ij} + 4\nabla^2\dot{E}^{ij})$$

$$q_{ij} \equiv \dot{E}_{ij} \longleftrightarrow p^{ij} = 2\beta(\ddot{E}^{ij} - \lambda^{ij})$$
$$\lambda_{ij} \equiv \lambda_{ij} \longleftrightarrow p^{ij}_{\lambda} = 0, \tag{3.33}$$

and the Hamiltonian is

$$H = \frac{M_P^2}{2} \int d^3x \left[\pi^{ij} q_{ij} + \frac{1}{4\beta} p^{ij} p_{ij} - E^{ij}(\beta \nabla^2 \nabla^2 + \nabla^2) E_{ij} \right.$$
$$\left. - q^{ij}(1 + 2\beta \nabla^2) q_{ij} + \lambda^{ij}(p_{ij} - 4\beta \nabla^2 E_{ij}) \right]. \tag{3.34}$$

The Poisson bracket of a pair of transverse traceless canonical coordinates can be found as

$$[E_{ij}(\mathbf{x}), \pi_{kl}(\mathbf{y})]_P = \hat{\Lambda}_{ij,kl} \delta^{(3)}(\mathbf{x} - \mathbf{y}), \tag{3.35}$$

where $\hat{\Lambda}_{ij,kl}$ is the transverse traceless projection operator defined by $\hat{\Lambda}_{ij,kl} \equiv 1/2(\hat{\theta}_{ik}\hat{\theta}_{jl} + \hat{\theta}_{il}\hat{\theta}_{jk} - \hat{\theta}_{ij}\hat{\theta}_{kl})$, while $\hat{\theta}_{ij} \equiv \delta_{ij} - \frac{\partial_i \partial_j}{\partial^2}$ is the transverse projection operator. Since the equations of motion in the Hamiltonian picture are generated by the Poisson bracket, the projection operator will preserve the transverse traceless characteristic.

It is clear that $p_{\lambda ij} = 0$ is a primary constraint as it is an auxiliary field. Via the consistency relation, we can generate further (traceless and transverse) secondary constraints as follows

$$\varphi_1 : p_{\lambda ij} = 0,$$
$$\varphi_2 : p_{ij} - 4\beta \nabla^2 E_{ij} \approx 0,$$
$$\varphi_3 : \pi_{ij} - 2q_{ij} \approx 0,$$
$$\varphi_4 : 2(\beta \nabla^2 \nabla^2 + \nabla^2) E_{ij} - \frac{1}{\beta} p_{ij} + 2(-1 + 2\beta \nabla^2) \lambda_{ij} \approx 0. \tag{3.36}$$

We can use the constraints φ_1, φ_4 to eliminate the degree of freedom (λ, p_λ), and use φ_2, φ_3 to eliminate (q, p). The coefficients in the action (3.32) are chosen such that there are at least four constraints in the theory and there is no ∇^2 in φ_3 which will generate nonlocal terms in the reduced Hamiltonian.

Using the constraints, (q_{ij}, p_{ij}) can be written as follows

$$q_{ij} = \frac{\pi_{ij}}{2},$$
$$p_{ij} = 4\beta \nabla^2 E_{ij}, \tag{3.37}$$

and the reduced Hamiltonian becomes

$$H_R = \frac{M_P^2}{2}\int d^3x \left[\frac{1}{4}\pi^{ij}(1-2\beta\nabla^2)\pi_{ij} + E^{ij}(-\nabla^2+3\beta\nabla^2\nabla^2)E_{ij}\right]. \tag{3.38}$$

To check whether the reduced Hamiltonian is bounded from below, we will explicitly quantize the theory. Similar to quantum electrodynamics in the Coulomb gauge, one can follow Dirac's method to quantize the constrained system. We first write down the generalized version of the Poisson bracket (i.e. the Dirac bracket), which generates the time evolution of any fields in constrained theory while preserving all of the constraints. We then promote all of the fields to operators and the commutators of two fields now become i times their Dirac bracket.

To write down the Dirac bracket, we first define a matrix $C_{ab} \equiv [\varphi_a, \varphi_b]_P$,

$$C_{ab;ij,kl}(\mathbf{x},\mathbf{y}) = \begin{bmatrix} 0 & 0 & 0 & -\hat{a} \\ 0 & 0 & -\hat{a} & 0 \\ 0 & \hat{a} & 0 & -\hat{b} \\ \hat{a} & 0 & \hat{b} & 0 \end{bmatrix} \hat{\Lambda}_{ij,kl}\delta^{(3)}(\mathbf{x}-\mathbf{y}),$$

where $\hat{a}$ and $\hat{b}$ are two operators $\hat{a} \equiv 2(-1+2\beta\nabla^2)$ and $\hat{b} \equiv 2(\beta\nabla^2\nabla^2+\nabla^2-1/\beta)$. The inverse of C_{ab} is

$$C^{-1;ab;ij,kl} = \begin{bmatrix} 0 & \hat{a}^{-2}\hat{b} & 0 & \hat{a}^{-1} \\ -\hat{a}^{-2}\hat{b} & 0 & \hat{a}^{-1} & 0 \\ 0 & -\hat{a}^{-1} & 0 & 0 \\ -\hat{a}^{-1} & 0 & 0 & 0 \end{bmatrix} \hat{\Lambda}^{ij,kl}\delta^{(3)}(\mathbf{x}-\mathbf{y}),$$

and the Dirac bracket of two fields X, Y is defined by

$$[X,Y]_D = [X,Y]_P - [X,\varphi_{a,ij}]_P C^{-1;ab;ij,kl}[\varphi_{b,kl},Y].$$

Equipped with the Dirac bracket, one can use the reduced Hamiltonian to write down the equations of motion of this system

$$\dot{E}_{ij} = [E_{ij},H_R]_D = \frac{1}{2}\pi_{ij} \tag{3.39}$$

$$\dot{\pi}_{ij} = [\pi_{ij},H_R]_D = \frac{(-2\nabla^2+6\beta\nabla^2\nabla^2)}{(-1+2\beta\nabla^2)}E_{ij}. \tag{3.40}$$

Using eqs. (3.39) and (3.40), we find

$$\ddot{E}_{ij} = \frac{(-\nabla^2 + 3\beta\nabla^2\nabla^2)}{(-1+2\beta\nabla^2)} E_{ij}, \tag{3.41}$$

which is the Euler-Lagrange equation of the action (3.32). We can solve eq. (3.41) by taking the Fourier transform

$$E_{ij}(\mathbf{x},t) = \int \frac{d^3p}{(2\pi)^3}\, e^{i\mathbf{p}\cdot\mathbf{x}} \tilde{E}_{ij}(\mathbf{p},t), \tag{3.42}$$

where $\tilde{E}_{ij}(\mathbf{p},t)$ satisfies

$$\left[\frac{d^2}{dt^2} - \frac{(\mathbf{p}^2 + 3\beta\mathbf{p}^4)}{(-1-2\beta\mathbf{p}^2)}\right] \tilde{E}_{ij}(\mathbf{p},t) = 0. \tag{3.43}$$

For any $\mathbf{p}$, $\tilde{E}_{ij}(\mathbf{p},t)$ is a harmonic oscillator with frequency

$$w_{\mathbf{p}} = \sqrt{\frac{(\mathbf{p}^2 + 3\beta\mathbf{p}^4)}{(1+2\beta\mathbf{p}^2)}}, \tag{3.44}$$

where $w_{\mathbf{p}}^2$ is positive definite if $\beta > 0$.

In order to quantize the theory, we write E_{ij}, π_{ij} as a linear summation of the creation and annihilation operators $a_{\mathbf{p}}^{r\dagger}$, $a_{\mathbf{p}}^{r}$,

$$\begin{aligned} E_{ij}(\mathbf{x}) &= \int \frac{d^3p}{(2\pi)^3}\, \frac{1}{\sqrt{2|w_{\mathbf{p}}|}} \sum_{r=1}^{2} \varepsilon_{ij}^r(\mathbf{p})(a_{\mathbf{p}}^r e^{i\mathbf{p}\cdot\mathbf{x}} + a_{\mathbf{p}}^{r\dagger} e^{-i\mathbf{p}\cdot\mathbf{x}}) \\ \pi_{kl}(\mathbf{x}) &= \int \frac{d^3p}{(2\pi)^3}\, -2i\sqrt{\frac{|w_{\mathbf{p}}|}{2}} \sum_{r=1}^{2} \varepsilon_{kl}^r(\mathbf{p})(a_{\mathbf{p}}^r e^{i\mathbf{p}\cdot\mathbf{x}} - a_{\mathbf{p}}^{r\dagger} e^{-i\mathbf{p}\cdot\mathbf{x}}), \end{aligned} \tag{3.45}$$

where the coefficients $\varepsilon_{ij}^r(\mathbf{p})$ are chosen in such a way that they solve the equations of motion eqs. (3.39) and (3.40), and the superscript r labels the polarizations. The symmetric transverse traceless tensor ε_{ij}^r satisfies $p^i\varepsilon_{ij}^r = \delta^{ij}\varepsilon_{ij}^r = 0$, and is normalized as

$$\sum_{i,j} \varepsilon_{ij}^r(\mathbf{p})\varepsilon^{s,ij}(\mathbf{p}) = \frac{\delta^{rs}}{2(1+2\beta\mathbf{p}^2)}, \tag{3.46}$$

with the completeness relation

$$\sum_{r=1}^{2} \varepsilon_{ij}^r(\mathbf{p})\varepsilon_{kl}^r(\mathbf{p}) = \frac{\Lambda_{ij,kl}(\mathbf{p})}{2(1+2\beta\mathbf{p}^2)}. \tag{3.47}$$

The operator $\Lambda_{ij,kl}(\mathbf{p})$ is defined by replacing all of the ∇^2 in the transverse traceless projection operator by $-\mathbf{p}^2$. One can calculate the Dirac bracket of (E_{ij}, π_{kl}) and the commutator of the two operators is thus

$$[E_{ij}(\mathbf{x}), \pi_{kl}(\mathbf{y})]_D = -i\Lambda_{ij,kl}\left[\frac{1}{(-1+2\beta\nabla^2)}\right]\delta^{(3)}(\mathbf{x}-\mathbf{y}). \tag{3.48}$$

With the normalization eq. (3.46) and the completeness relation eq. (3.47), the commutation relation eq. (3.48) is equivalent to

$$\begin{aligned}[a^r_{\mathbf{p}}, a^s_{\mathbf{q}}] &= [a^{r,\dagger}_{\mathbf{p}}, a^{s,\dagger}_{\mathbf{q}}] = 0,\\ [a^r_{\mathbf{p}}, a^{s,\dagger}_{\mathbf{q}}] &= (2\pi^3)\delta^{rs}\delta^{(3)}(\mathbf{p}-\mathbf{q}).\end{aligned} \tag{3.49}$$

One can thus rewrite the reduced Hamiltonian (3.38) as creation and annihilation operators

$$H_R = \frac{M_P^2}{2}\int \frac{d^3p}{(2\pi)^3}\,|w_p|\left(\sum_{r=1}^{2} a^{r,\dagger}_{\mathbf{p}} a^r_{\mathbf{p}} + \frac{1}{2}(2\pi)^3\delta^{rr}\delta^{(3)}(0)\right). \tag{3.50}$$

The energy spectrum is real and bounded from below if w_p^2 is positive definite, as long as $\beta > 0$.

3.4.2 Helicity-1 sector

We now turn to the helicity-1 unstable modes. As shown in Section 3.3.2, this sector is tachyonic if $\beta < 0$ and ghostlike if $\beta > 0$. As usual, we will remove it by modifying the action (3.16) with the introduction of a helicity-1 field λ_i

$$S = \frac{M_P^2}{2}\int d^4x\,\frac{\beta}{2}\left[(\dot{v}_i - \lambda_i)^2 + v_i\nabla^2 v^i + \frac{1}{\beta}v_i v^i\right]. \tag{3.51}$$

Ostrogradski's choice of canonical coordinates is

$$\begin{aligned} v_i \equiv v_i &\longleftrightarrow p^i_v = \beta(\dot{v}^i - \lambda^i)\\ \lambda_i \equiv \lambda_i &\longleftrightarrow p^i_\lambda = 0,\end{aligned} \tag{3.52}$$

and the Hamiltonian is

$$H = \frac{M_P^2}{2}\int d^3x\,\left(\frac{p^i_v p_{vi}}{2\beta} + p^i_v\lambda_i - \frac{\beta}{2}v_i\nabla^2 v^i - \frac{1}{2}v_i v^i\right). \tag{3.53}$$

There are four constraints in the theory, which can be found as

$$
\begin{aligned}
&\varphi_1 : p^i_\lambda = 0 \\
&\varphi_2 : p^i_v \approx 0 \\
&\varphi_3 : v^i + \beta\nabla^2 v^i \approx 0 \\
&\varphi_4 : \frac{p^i_v}{\beta} + \nabla^2 p^i_v + \lambda^i + \beta\nabla^2\lambda^i \approx 0.
\end{aligned} \tag{3.54}
$$

If we use the four constraints to eliminate (v_i, p^i_v), (λ_i, p^i_λ), the physical phase space will be zero-dimensional and the reduced Hamiltonian vanishes.

3.4.3 Helicity-0 sector

Finally, we introduce a helicity-0 field λ into the action (3.22)

$$
\begin{aligned}
S = \frac{M_P^2}{2}\int d^4x \Big[&(-6\dot{\Psi}^2 - 2\Psi\nabla^2\Psi + 4\Psi\nabla^2\Phi) + 4(\beta+3\alpha)(3\ddot{\Psi}^2 + 4\dot{\Psi}\nabla^2\dot{\Psi} + 2\ddot{\Psi}\nabla^2\Phi) \\
&+ 2(3\beta+8\alpha)(\nabla^2\Psi)^2 + 2(\beta+2\alpha)(\nabla^2\Phi)^2 - 4(\beta+4\alpha)\nabla^2\Psi\nabla^2\Phi \\
&+ 32(\beta+3\alpha)\lambda\nabla^2\Psi + 12(\beta+3\alpha)(\lambda^2 - 2\ddot{\Psi}\lambda - \frac{2}{3}\lambda\nabla^2\Phi) + A\lambda\Psi\Big],
\end{aligned} \tag{3.55}
$$

where A is some arbitrary real constant. There are four naive degrees of freedom, whose canonical variables are

$$
\begin{aligned}
\Phi \equiv \Phi &\longleftrightarrow p_\Phi = 0 \\
\Psi \equiv \Psi &\longleftrightarrow p_\Psi = \frac{\delta S}{\delta\dot{\Psi}} \\
\chi \equiv \chi &\longleftrightarrow p_\chi = 8(\beta+3\alpha)[3(\ddot{\Psi} - \lambda) + \nabla^2\Phi] \\
\lambda \equiv \lambda &\longleftrightarrow p_\lambda = 0.
\end{aligned} \tag{3.56}
$$

The Hamiltonian can be written as

$$
\begin{aligned}
H = \frac{M_P^2}{2}\int d^3x \Big[&p_\Psi\chi + \frac{p_\chi^2}{48(\beta+3\alpha)} - \frac{p_\chi\nabla^2\Phi}{3} + \big(6\chi^2 + 2\Psi\nabla^2\Psi - 4\Psi\nabla^2\Phi\big) \\
&- 16(\beta+3\alpha)\chi\nabla^2\chi - 2(3\beta+8\alpha)(\nabla^2\Psi)^2 + 4(\beta+4\alpha)\nabla^2\Psi\nabla^2\Phi
\end{aligned}
$$

$$-\frac{2\beta}{3}(\nabla^2\Phi)^2+\lambda(p_\chi-A\Psi)-32(\beta+3\alpha)\lambda\nabla^2\Psi\Big],\tag{3.57}$$

where the constraints of this theory are

$$\begin{aligned}
&\varphi_1 : p_\Phi = 0\\
&\varphi_2 : p_\lambda = 0\\
&\varphi_3 : p_\chi - A\Psi - 32(\beta+3\alpha)\nabla^2\Psi \approx 0\\
&\varphi_4 : \nabla^2\left[\frac{p_\chi}{3}+4\Psi-4(\beta+4\alpha)\nabla^2\Psi+\frac{4\beta}{3}\nabla^2\Phi\right]\approx 0\\
&\varphi_5 : p_\Psi+(12+A)\chi \approx 0\\
&\varphi_6 : \frac{(12+A)p_\chi}{24(\beta+3\alpha)}+2(6+A)\lambda+32(\beta+3\alpha)\nabla^2\lambda\\
&\quad +4(3\beta+8\alpha)\nabla^2\nabla^2\Psi-4(\beta+4\alpha)\nabla^2\nabla^2\Phi-4\nabla^2\Psi-\frac{A}{3}\nabla^2\Phi\approx 0.
\end{aligned}\tag{3.58}$$

We use the six constraints to eliminate *three* pairs of canonical coordinates (Φ, p_Φ), (λ, p_λ), and (χ, p_χ), reducing the Hamiltonian to

$$\begin{aligned}
H_R = \frac{M_P^2}{2}\int d^3x\,&\left\{\frac{-p_\Psi}{(12+A)^2}[(6+A)+16(\beta+3\alpha)\nabla^2]p_\Psi+\frac{1}{\beta}\left[(6+A)+\frac{A^2(\beta+2\alpha)}{16(\beta+3\alpha)}\right]\Psi^2\right.\\
&\left.+\left[(22+\frac{48\alpha}{\beta})+\frac{A}{\beta}(3\beta+4\alpha)\right]\Psi\nabla^2\Psi+\frac{32(\beta+3\alpha)(\beta+\alpha)}{\beta}(\nabla^2\Psi)^2\right\},
\end{aligned}\tag{3.59}$$

which generates the evolution of a single dynamical variable Ψ. It is clear that it can be made positive definite by certain parameter choices, for example ($A=-8, \alpha=0, \beta>0$). We can check that the quantum theory is also stable in the following manner. First, we find the Dirac bracket of the theory as usual, which can be used to find the equations of motion

$$\dot{\Psi} = -\frac{1}{(12+A)}p_\Psi \tag{3.60}$$

$$\begin{aligned}
\dot{p}_\Psi = &-\frac{(12+A)}{2(6+A)+32(\beta+3\alpha)\nabla^2}\times\left\{\frac{2}{\beta}\left[(6+A)+\frac{A^2(\beta+2\alpha)}{16(\beta+3\alpha)}\right]\Psi\right.\\
&\left.+2\left[(22+\frac{48\alpha}{\beta})+\frac{A}{\beta}(3\beta+4\alpha)\right]\nabla^2\Psi+\frac{64(\beta+3\alpha)(\beta+\alpha)}{\beta}\nabla^2\nabla^2\Psi\right\}.
\end{aligned}\tag{3.61}$$

One can check that eqs. (3.60) and (3.61) reproduce the Euler-Lagrange equation. We solve these equations by taking the Fourier transform

$$\Psi(\mathbf{x},t) = \int \frac{d^3p}{(2\pi)^3}\, e^{i\mathbf{p}\cdot\mathbf{x}}\tilde{\Psi}(\mathbf{p},t) \tag{3.62}$$

as usual, and the solution $\tilde{\Psi}(\mathbf{p},t)$ is a harmonic oscillator with frequency

$$\begin{aligned} w_{\mathbf{p}}^2 = \frac{1}{\beta[16(\beta+3\alpha)\mathbf{p}^2-(6+A)]} \times \Bigg\{ & 32(\beta+3\alpha)(\beta+\alpha)\mathbf{p}^4 \\ & -[\beta(3A+22)+\alpha(4A+48)]\mathbf{p}^2 + \left[6+A+\frac{A^2(\beta+2\alpha)}{16(\beta+3\alpha)}\right]\Bigg\}. \end{aligned} \tag{3.63}$$

Following the usual quantization rules, we define the commutator of Ψ and p_Ψ to be “i" times their classical Dirac bracket, i.e.

$$[\Psi, p_\Psi] \equiv \frac{(12+A)i}{2(6+A)+32(\beta+3\alpha)\nabla^2}\delta^{(3)}(\mathbf{x}-\mathbf{y}). \tag{3.64}$$

By expanding Ψ, p_Ψ as a linear summation of creation and annihilation operators $a_{\mathbf{p}}^\dagger$, $a_{\mathbf{p}}$

$$\Psi = \int \frac{d^3p}{(2\pi)^3}\, \frac{1}{\sqrt{2w_{\mathbf{p}}}} \frac{[a_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}} + a_{\mathbf{p}}^\dagger e^{i\mathbf{p}\cdot\mathbf{x}}]}{\sqrt{[32(\beta+3\alpha)\mathbf{p}^2-2(6+A)]}} \tag{3.65}$$

$$p_\Psi = \int \frac{d^3p}{(2\pi)^3} \sqrt{\frac{w_{\mathbf{p}}}{2}}\, \frac{i(12+A)[a_{\mathbf{p}}e^{i\mathbf{p}\cdot\mathbf{x}} - a_{\mathbf{p}}^\dagger e^{i\mathbf{p}\cdot\mathbf{x}}]}{\sqrt{[32(\beta+3\alpha)\mathbf{p}^2-2(6+A)]}}, \tag{3.66}$$

we obtain the usual result that the commutator eq. (3.64) is consistent with the Fock space commutators

$$\begin{aligned} [a_{\mathbf{p}}, a_{\mathbf{q}}] &= [a_{\mathbf{p}}^\dagger, a_{\mathbf{q}}^\dagger] = 0, \\ [a_{\mathbf{p}}, a_{\mathbf{q}}^\dagger] &= (2\pi^3)\delta^{(3)}(\mathbf{p}-\mathbf{q}). \end{aligned} \tag{3.67}$$

Using these operators, the reduced Hamiltonian can then be written as

$$H_R = \frac{M_P^2}{2}\int \frac{d^3p}{(2\pi)^3}\, |w_{\mathbf{p}}| \left(\sum_{r=1}^{2} a_{\mathbf{p}}^\dagger a_{\mathbf{p}} + \frac{1}{2}(2\pi)^3\delta^{(3)}(0)\right), \tag{3.68}$$

which is bounded from below as long as $w_{\mathbf{p}}^2$ is positive definite, a condition which is satisfied by suitable choices of parameters A, α and β.

3.5 Quadratic action around the de Sitter background

In this and the following section, we will show how to remove the unstable degrees of freedom in each helicity sector in the de Sitter background. In order to separate the action into the helicity-0, 1, and 2 sectors, we first parameterize the metric fluctuation as

$$ds^2 = a^2(t)\left\{-(1+2\phi)dt^2 + 2B_i dx^i dt + [(1-2\psi)\delta_{ij} + 2E_{ij}]dx^i dx^j\right\}, \tag{3.69}$$

where t is conformal time, and $a(t) = -\frac{1}{Ht}$, $\Lambda = 3H^2$ in the de Sitter background. We can again decompose B_i, E_{ij} into

$$B_i = \partial_i B + B_i^{\mathrm{T}} \tag{3.70}$$

$$E_{ij} = \partial_{\langle i}\partial_{j\rangle} E + \partial_{(i} E_{j)}^{\mathrm{T}} + E_{ij}^{\mathrm{TT}}. \tag{3.71}$$

3.5.1 Helicity-2 sector

The second order action of the helicity-2 modes in the de Sitter background is

$$S = \frac{M_P^2}{2}\int d^4x \, \left\{\beta[(\ddot{E}_{ij})^2 + 2\dot{E}^{ij}\nabla^2\dot{E}_{ij} + (\nabla^2 E_{ij})^2] + ca^2(t)[(\dot{E}_{ij})^2 + E^{ij}\nabla^2 E_{ij}]\right\}, \tag{3.72}$$

which reduces to the unconstrained Minkowski case if $a(t) \to 1$, and $H^2 \to 0$. We have also defined the dimensionless parameter

$$c \equiv 1 + 8H^2(\beta + 3\alpha). \tag{3.73}$$

From eq. (3.1), we note that in the de Sitter space $R \approx H^2$, and hence we would expect that $|\alpha|H^2 \ll 1$ and $|\beta|H^2 \ll 1$ if we assume that the higher derivative terms are corrections to general relativity i.e., $c > 0$ unless the higher derivative terms dominate. As an aside, note that, in the special case of $\beta + 3\alpha = 0$, $c = 1$.

Since there is no constraint in the helicity-2 modes, there are four helicity-2 degrees of freedom in the theory. Ostrogradski's choice of canonical coordinates is

$$\begin{aligned} E_{ij} &\equiv E_{ij} \longleftrightarrow \pi^{ij} \\ q_{ij} &\equiv \dot{E}_{ij} \longleftrightarrow p^{ij} = 2\beta\ddot{E}^{ij}, \end{aligned} \tag{3.74}$$

and the Hamiltonian is

$$H = \frac{M_P^2}{2}\int d^3x \left(\frac{p^{ij}p_{ij}}{4\beta} + \pi^{ij}q_{ij} - 2\beta q^{ij}\nabla^2 q_{ij} - ca^2(t)q^{ij}q_{ij} - \beta\nabla^2 E^{ij}\nabla^2 E_{ij} - ca^2(t)E^{ij}\nabla^2 E_{ij}\right). \quad (3.75)$$

As in eq. (3.75), the linear dependence of π^{ij} is the signal of the Ostrogradski instability. The $\pi^{ij}q_{ij}$ term can be arbitrarily negative when $q_{ij} > 0$, $\pi^{ij} \to -\infty$ or vice versa and hence the Hamiltonian is unbounded from below.

3.5.2 Helicity-1 sector

The action up to the quadratic level of the helicity-1 modes can be written using the gauge invariant variable $v_i = \sqrt{-\nabla^2}(B_i - \dot{E}_i)$,

$$S = \frac{M_P^2}{2}\int d^4x\, \frac{\beta}{2}\left(\dot{v}_i\dot{v}^i + v_i\nabla^2 v^i + \frac{ca^2(t)}{\beta}v_i v^i\right), \quad (3.76)$$

and the Euler-Lagrange equation of action (3.76) is

$$\left[\beta\left(\frac{d^2}{dt^2} - \nabla^2\right) - ca^2(t)\right] v_i = 0. \quad (3.77)$$

Through the Fourier transform, we find that the solutions are harmonic oscillators with frequency $w_{\mathbf{p}}^2 = \mathbf{p}^2 - \frac{ca^2(t)}{\beta}$. The canonical momentum conjugate to v_i is then defined by

$$p_v^i = \frac{\delta S}{\delta \dot{v}_i} = \beta\dot{v}^i \quad (3.78)$$

Since we use the gauge invariant variable to write the action, there is no constraint and the Hamiltonian is

$$H = \frac{M_P^2}{2}\int d^3x \left(\frac{p_{vi}p_v^i}{2\beta} - \frac{\beta}{2}v_i\nabla^2 v^i - \frac{ca^2(t)}{2}v_i v^i\right). \quad (3.79)$$

There is a subtle but important difference between the de Sitter and Minkowski backgrounds for the helicity-1 modes. In the Minkowski case, eq. (3.19), the helicity-1 mode is either tachyonic or ghostlike since the sign of $v_i v^i$ is always negative. However, in the de Sitter background, one may choose $c < 0$ to render the coefficient of $v_i v^i$ in eq. (3.79) to be positive. Nevertheless, as we have argued that, generically, $c > 0$ unless the higher derivative terms dominate, we will not consider this case further. The helicity-1 sector is thus either tachyonic or ghostlike, depending on the sign of β.

3.5.3 Helicity-0 sector

The second order action for the helicity-0 modes in the de Sitter space is far more complicated. We can use the usual two gauge invariant variables

$$\begin{aligned}\Phi &= \phi + \dot{B} - \ddot{E} - \frac{1}{t}(B - \dot{E}),\\ \Psi &= \psi + \frac{1}{3}\nabla^2 E + \frac{1}{t}(B - \dot{E}),\end{aligned} \tag{3.80}$$

to write the action as

$$\begin{aligned}S = \frac{M_P^2}{2}\int d^4x\ \{ & a^2(t)\left(-6\dot{\Psi}^2 - 2\Psi\nabla^2\Psi + 4\Psi\nabla^2\Phi - 12a(t)H\Phi\dot{\Psi} - 6a^2(t)H^2\Phi^2\right)\\ & + 4(\beta+3\alpha)(3\ddot{\Psi}^2 + 4\dot{\Psi}\nabla^2\dot{\Psi} + 2\ddot{\Psi}\nabla^2\Phi)\\ & + 2(3\beta+8\alpha)(\nabla^2\Psi)^2 + 2(\beta+2\alpha)(\nabla^2\Phi)^2 - 4(\beta+4\alpha)\nabla^2\Psi\nabla^2\Phi\\ & + 4(\beta+3\alpha)\left[12a^3(t)H^3\Phi\dot{\Psi} + a^2(t)H^2(6\dot{\Psi}^2 + 6\ddot{\Psi}\Phi + 3\dot{\Phi}^2 + 2\Psi\nabla^2\Psi\right.\\ & \left. + 7\Phi\nabla^2\Phi - 4\Psi\nabla^2\Phi) + 2a(t)H(5\dot{\Psi}\nabla^2\Phi + 3\ddot{\Psi}\dot{\Phi})\right]\}.\end{aligned} \tag{3.81}$$

Again, one can see that the action can be reduced to the Minkowski case eq. (3.22) if $a(t) \to 1$, $H \to 0$. However, this set of variables is rather unwieldy because of the nonlocal terms in the reduced Hamiltonian, so we choose the following pair of gauge invariant variables instead

$$\begin{aligned}\mathcal{A} &= \phi - t(\dot{\psi} + \frac{1}{3}\nabla^2\dot{E}),\\ \mathcal{B} &= \nabla^2\left[B - \dot{E} + t\left(\psi + \frac{1}{3}\nabla^2 E\right)\right],\end{aligned} \tag{3.82}$$

whereupon the action becomes far shorter, i.e.

$$\begin{aligned}S = \frac{M_P^2}{2}\int d^4x\ \Big\{ & 2(\beta+2\alpha)(\dot{\mathcal{B}} + \nabla^2\mathcal{A})^2 - 2a^4H^2\mathcal{A}(\frac{2\mathcal{B}}{aH} + 3\mathcal{A})\\ & + (\beta+3\alpha)\left\{8aH(\mathcal{B}\nabla^2\mathcal{A} + \dot{\mathcal{B}}\dot{\mathcal{A}}) + 16a^3H^3\mathcal{B}\mathcal{A}\right.\\ & \left. + a^2H^2[8\mathcal{B}^2 + 8\dot{\mathcal{B}}\mathcal{A} + 12\dot{\mathcal{A}}^2 + 28\mathcal{A}\nabla^2\mathcal{A}]\right\}\Big\}.\end{aligned} \tag{3.83}$$

The two sets of gauge invariant variables eqs. (3.80) and (3.82) are related through the following field redefinitions

$$\mathcal{A} = \frac{\dot{\Psi}}{aH} + \Phi\,,\ \mathcal{B} = -\frac{1}{aH}\nabla^2\Psi. \tag{3.84}$$

One can see that there is no corresponding field redefinitions in the Minkowski background, because, in the Minkowski case, $aH = 0$. In this and the following sections, we will use the new pair of fields $(\mathcal{A},\mathcal{B})$ to show how to remove the instability. We made this choice to evade the possible issues with the physical interpretation of the nonlocal terms – we emphasize that, in principle, one can use the old pair of fields (Φ,Ψ) and add the same set of auxiliary field terms to remove all of the instability, similar to the helicity-1, 2 cases.[8] Since there might not be any nonlocal terms in the full non-linear theory, this procedure could provide a better starting point – however, since the non-linear theory is beyond our scope, here we only want to show that the same idea can be applied to the de Sitter case. Using the new set of variables, the canonical momenta can be written as

$$\begin{aligned} p_{\mathcal{B}} &= (\beta+3\alpha)\left(8a^2H^2\mathcal{A}+8aH\dot{\mathcal{A}}\right)+4(\beta+2\alpha)\left(\dot{\mathcal{B}}+\nabla^2\mathcal{A}\right), \\ p_{\mathcal{A}} &= (\beta+3\alpha)\left(24a^2H^2\dot{\mathcal{A}}+8aH\dot{\mathcal{B}}\right), \end{aligned} \tag{3.85}$$

and the Hamiltonian thus becomes

$$\begin{aligned} H = \frac{M_P^2}{2}\int d^3x\, \Bigg\{ & \frac{3}{8\beta}\left[p_{\mathcal{B}}^2 - \frac{2p_{\mathcal{B}}p_{\mathcal{A}}}{3aH} + \frac{(\beta+2\alpha)p_{\mathcal{A}}^2}{6a^2H^2(\beta+3\alpha)}\right] \\ & - \frac{3}{\beta}(p_{\mathcal{B}} - \frac{p_{\mathcal{A}}}{3aH})[2a^2H^2(\beta+3\alpha)\mathcal{A} + (\beta+2\alpha)\nabla^2\mathcal{A}] \\ & - 8a^3H^3(\beta+3\alpha)\mathcal{B}\left\{\frac{\nabla^2\mathcal{A}}{a^2H^2} + \frac{\mathcal{B}}{aH} + \left[2 - \frac{1}{2(\beta+3\alpha)H^2}\right]\mathcal{A}\right\} \\ & + \frac{2}{\beta}\mathcal{A}\left[2(\beta+2\alpha)(\beta+3\alpha)\nabla^2\nabla^2 + 12a^4H^4(\beta+3\alpha)^2\right. \\ & \left. + 2a^2H^2(\beta+3\alpha)(12\alpha-\beta)\nabla^2 + 3a^4H^2\beta\right]\mathcal{A}\Bigg\}. \end{aligned} \tag{3.86}$$

To make the dynamics explicit, we perform a final canonical transformation

$$p_{\mathcal{B}} \longrightarrow p_{\mathcal{B}} - \frac{1}{3aH}p_{\mathcal{A}} + \frac{4}{3aH}[(\beta+2\alpha)\nabla^2 + 2a^2H^2(\beta+3\alpha)]\mathcal{B}$$

[8]We possibly need to add extra terms in both the Minkowski and the de Sitter cases if we want to remove the nonlocal terms by constraints.

$$\mathcal{A} \longrightarrow \mathcal{A} + \frac{1}{3aH}\mathcal{B}, \tag{3.87}$$

such that the Hamiltonian becomes

$$\begin{aligned} H = \frac{M_P^2}{2}\int d^3x\, \Bigg\{ & \frac{3}{8\beta}p_{\mathcal{B}}^2 + \frac{p_{\mathcal{A}}^2}{48a^2H^2(\beta+3\alpha)} - \frac{3}{\beta}p_{\mathcal{B}}\left[(\beta+2\alpha)\nabla^2 + 2a^2H^2(\beta+3\alpha)\right]\mathcal{A} \\ & - \frac{2}{9}\mathcal{B}\left[\frac{(\beta+2\alpha)}{a^2H^2}\nabla^2\nabla^2 + 2(\beta+3\alpha)\nabla^2 + (3a^2 + 12a^2H^2(\beta+3\alpha))\right]\mathcal{B} \\ & + \frac{4}{3aH}\mathcal{B}\left[(\beta+2\alpha)\nabla^2\nabla^2 + 8a^2H^2(\beta+3\alpha)\nabla^2 - 12a^4H^4(\beta+3\alpha)\right]\mathcal{A} \\ & + \frac{2}{\beta}\mathcal{A}\left[2(\beta+2\alpha)(\beta+3\alpha)\nabla^2\nabla^2 + 12a^4H^4(\beta+3\alpha)^2\right. \\ & \left. +\ 2a^2H^2(\beta+3\alpha)(12\alpha-\beta)\nabla^2 + 3a^4H^2\beta\right]\mathcal{A}\Bigg\}. \end{aligned} \tag{3.88}$$

Since there is no constraint in this theory, this is the Hamiltonian describing two physical degrees of freedom. Although the instability is not explicitly shown, one can see that, at some limit the Hamiltonian is unbounded from below. In order to have stable kinetic terms, we require that $(\beta+3\alpha) > 0$ and $\beta > 0$, which guarantees that the first two terms are positive definite. On the other hand, in the high frequency limit in Fourier space, one should expect those terms with the highest spatial derivatives to dominate. This requires that $(\beta+2\alpha) < 0$ so that the $\mathcal{B}^2$ term is stable (from the second line in eq. (3.88)), which cannot be satisfied simultaneously. We thus conclude that the Hamiltonian is unbounded from below with any parameter choice in the high frequency limit and there must be at least one unstable degree of freedom in the helicity-0 sector.

3.6 Stabilization by constraints in the de Sitter background

3.6.1 Helicity-2 sector

Similar to the Minkowski case, we first rewrite the action (3.72) by introducing a helicity-2 auxiliary tensor field λ_{ij}

$$\begin{aligned} S = \frac{M_P^2}{2}\int d^4x\, \Big\{ & \beta\left[(\ddot{E}_{ij} - \lambda_{ij})^2 + 2\dot{E}^{ij}\nabla^2\dot{E}_{ij} + (\nabla^2 E_{ij})^2 + 4\lambda^{ij}\nabla^2 E_{ij}\right] \\ & + ca^2(t)[(\dot{E}_{ij})^2 + E^{ij}\nabla^2 E_{ij}\Big\}, \end{aligned} \tag{3.89}$$

where λ_{ij} is transverse traceless, and its introduction explicity breaks the Lorentz invariance.

Ostrogradski's choice of canonical coordinates is

$$\begin{aligned}
E_{ij} &\equiv E_{ij} \longleftrightarrow \pi^{ij} = 2ca^2\dot{E}^{ij} + \beta(-2\dddot{E}^{ij} + 2\dot{\lambda}^{ij} + 4\nabla^2\dot{E}^{ij}), \\
q_{ij} &\equiv \dot{E}_{ij} \longleftrightarrow p^{ij} = 2\beta(\ddot{E}^{ij} - \lambda^{ij}), \\
\lambda_{ij} &\equiv \lambda_{ij} \longleftrightarrow p_{\lambda}^{ij} = 0,
\end{aligned} \tag{3.90}$$

and the Hamiltonian is

$$\begin{aligned}
H = \frac{M_P^2}{2}\int d^3x \, \Big[& \pi^{ij}q_{ij} + \frac{1}{4\beta}p^{ij}p_{ij} - q^{ij}(ca^2 + 2\beta\nabla^2)q_{ij} \\
& - E^{ij}(\beta\nabla^2\nabla^2 + ca^2\nabla^2)E_{ij} + \lambda^{ij}(p_{ij} - 4\beta\nabla^2 E_{ij})\Big].
\end{aligned} \tag{3.91}$$

The Poisson bracket of a pair of transverse traceless canonical coordinates is identical to their Minkowski counterparts

$$[E_{ij}(\mathbf{x}), \pi_{kl}(\mathbf{y})]_P = \hat{\Lambda}_{ij,kl}\delta^{(3)}(\mathbf{x} - \mathbf{y}). \tag{3.92}$$

To find the constraints, we apply the Dirac Bracket formalism as usual. It is clear that $p_{\lambda}^{ij} = 0$ is a primary constraint, and the rest of the (transverse and traceless) constraints of this theory are generated by the consistency relation

$$\begin{aligned}
&\varphi_1 : p_{\lambda ij} = 0, \\
&\varphi_2 : p_{ij} - 4\beta\nabla^2 E_{ij} \approx 0, \\
&\varphi_3 : \pi_{ij} - 2ca^2 q_{ij} \approx 0, \\
&\varphi_4 : 2(\beta\nabla^2\nabla^2 + ca^2\nabla^2)E_{ij} - \frac{ca^2}{\beta}p_{ij} \\
&\qquad + 2(-ca^2 + 2\beta\nabla^2)\lambda_{ij} \approx 0.
\end{aligned} \tag{3.93}$$

Armed with these, we can use φ_1, φ_4 to eliminate the degree of freedom (λ, p_λ), and φ_2, φ_3 to eliminate (q, p). The coefficients in the action (3.89) are again chosen such that there are at least four constraints in the theory and there is no ∇^2 in φ_3 which will generate nonlocal terms in the reduced Hamiltonian.

Using the constraints, (q_{ij}, p_{ij}) can be written as follows

$$\begin{aligned}
q_{ij} &= \frac{\pi_{ij}}{2ca^2}, \\
p_{ij} &= 4\beta\nabla^2 E_{ij},
\end{aligned} \tag{3.94}$$

and the reduced Hamiltonian becomes

$$H_R = \frac{M_P^2}{2} \int d^3x \left[\frac{1}{4c^2a^4} \pi^{ij} (ca^2 - 2\beta\nabla^2) \pi_{ij} + E^{ij}(-ca^2\nabla^2 + 3\beta\nabla^2\nabla^2) E_{ij} \right], \tag{3.95}$$

which is positive definite if $\beta > 0$, $c > 0$.

3.6.2 Helicity-1 sector

In Section 3.5.2, we showed that the helicity-1 modes are only stable if $\beta > 0$, $c < 0$. However, our imposition of the constraints to restore the stability of the helicity-2 sector requires that $c > 0$ in addition to the usual arguments on subdominant higher derivative terms. We thus choose $c > 0$ and remove the unstable helicity-1 modes altogether, as follows. Similar to the Minkowski case, we modify the action (3.76) by introducing a helicity-1 field λ_i

$$S = \frac{M_P^2}{2} \int d^4x \, \frac{\beta}{2} \left[(\dot{v}_i - \lambda_i)^2 + v_i \nabla^2 v^i + \frac{ca^2}{\beta} v_i v^i \right]. \tag{3.96}$$

Ostrogradski's choice of canonical coordinates is

$$\begin{aligned} v_i \equiv v_i &\longleftrightarrow p_v^i = \beta(\dot{v}^i - \lambda^i), \\ \lambda_i \equiv \lambda_i &\longleftrightarrow p_\lambda^i = 0, \end{aligned} \tag{3.97}$$

and the Hamiltonian is

$$H = \frac{M_P^2}{2} \int d^3x \left(\frac{p_v^i p_{vi}}{2\beta} + p_v^i \lambda_i - \frac{\beta}{2} v_i \nabla^2 v^i - \frac{ca^2}{2} v_i v^i \right). \tag{3.98}$$

The four constraints in the theory can be found as

$$\begin{aligned} &\varphi_1 : p_\lambda^i = 0 \\ &\varphi_2 : p_v^i \approx 0 \\ &\varphi_3 : ca^2 v^i + \beta \nabla^2 v^i \approx 0 \\ &\varphi_4 : \frac{ca^2 p_v^i}{\beta} + \nabla^2 p_v^i + ca^2 \lambda^i + \beta \nabla^2 \lambda^i \approx 0. \end{aligned} \tag{3.99}$$

If we use the four constraints to eliminate (v_i, p_v^i), (λ_i, p_λ^i), the physical phase space will be zero-dimensional and the reduced Hamiltonian vanishes.

3.6.3 Helicity-0 sector

Finally, we deal with the helicity-0 instability. We modify the action eq. (3.83) by introducing a helicity-0 field λ

$$S=\frac{M_P^2}{2}\int d^4x\,\Big\{2(\beta+2\alpha)(\dot{\mathcal{B}}-\lambda+\nabla^2\mathcal{A})^2-2a^4H^2\mathcal{A}\left(\frac{2\mathcal{B}}{aH}+3\mathcal{A}\right)$$
$$+(\beta+3\alpha)\left[8aH(\mathcal{B}\nabla^2\mathcal{A}+\dot{\mathcal{B}}\dot{\mathcal{A}}-\lambda\dot{\mathcal{A}})+16a^3H^3\mathcal{B}\mathcal{A}+28a^2H^2\mathcal{A}\nabla^2\mathcal{A}\right.$$
$$\left.+a^2H^2\left(8\mathcal{B}^2+8\dot{\mathcal{B}}\mathcal{A}-8\lambda\mathcal{A}+12\dot{\mathcal{A}}^2\right)\right]\Big\}. \tag{3.100}$$

As now must be familiar, the canonical coordinates are

$$p_\lambda=0,$$
$$p_{\mathcal{A}}=(\beta+3\alpha)(24a^2H^2\dot{\mathcal{A}}+8aH\dot{\mathcal{B}}-8aH\lambda),$$
$$p_{\mathcal{B}}=4(\beta+2\alpha)(\dot{\mathcal{B}}-\lambda+\nabla^2\mathcal{A})+(\beta+3\alpha)(8a^2H^2\mathcal{A}+8aH\dot{\mathcal{A}}), \tag{3.101}$$

and the Hamiltonian is

$$H=\frac{M_P^2}{2}\int d^3x\,\Big\{\frac{3}{8\beta}\left[p_{\mathcal{B}}^2-\frac{2p_{\mathcal{B}}p_{\mathcal{A}}}{3aH}+\frac{(\beta+2\alpha)p_{\mathcal{A}}^2}{6a^2H^2(\beta+3\alpha)}\right]+\lambda p_{\mathcal{B}}$$
$$-\frac{3}{\beta}(p_{\mathcal{B}}-\frac{p_{\mathcal{A}}}{3aH})[2a^2H^2(\beta+3\alpha)\mathcal{A}+(\beta+2\alpha)\nabla^2\mathcal{A}]$$
$$+\frac{2}{\beta}\mathcal{A}\left[2(\beta+2\alpha)(\beta+3\alpha)\nabla^2\nabla^2+12a^4H^4(\beta+3\alpha)^2\right.$$
$$\left.+2a^2H^2(\beta+3\alpha)(12\alpha-\beta)\nabla^2+3a^4H^2\beta\right]\mathcal{A}$$
$$-8a^3H^3(\beta+3\alpha)\mathcal{B}\left[\frac{\nabla^2\mathcal{A}}{a^2H^2}+\left(2-\frac{1}{2(\beta+3\alpha)H^2}\right)\mathcal{A}+\frac{\mathcal{B}}{aH}\right]\Big\}. \tag{3.102}$$

There are four constraints in the theory, which can be found as

$$\varphi_1: p_\lambda=0$$
$$\varphi_2: p_{\mathcal{B}}\approx 0$$
$$\varphi_3: \frac{2\mathcal{B}}{aH}+\frac{\nabla^2\mathcal{A}}{a^2H^2}+\left[2-\frac{1}{2H^2(\beta+3\alpha)}\right]\mathcal{A}\approx 0$$
$$\varphi_4: F(\lambda,\cdots)\approx 0. \tag{3.103}$$

Applying these constraints to remove the λ and $\mathcal{B}$ pairs and redefine

$$p_{\mathcal{A}} \to p_{\mathcal{A}} + \left[8aH(\beta+3\alpha)\nabla^2 + 16a^3H^3\frac{(\beta+3\alpha)^2}{\beta+2\alpha}\right]\mathcal{A} \tag{3.104}$$

we obtain the reduced Hamiltonian

$$\begin{aligned} H_R = \frac{M_P^2}{2}\int d^3x \,\bigg\{ & \frac{(\beta+2\alpha)}{16a^2H^2\beta(\beta+3\alpha)}p_{\mathcal{A}}^2 + 2(\beta+3\alpha)(\nabla^2\mathcal{A})^2 \\ & - 2a^2[1+6(\beta+3\alpha)H^2]\mathcal{A}\nabla^2\mathcal{A} + 2a^4H^2\mathcal{A}^2 \\ & + \left[\frac{8a^4H^4(\beta+3\alpha)(2\beta+5\alpha)}{(\beta+2\alpha)} + \frac{a^4}{2(\beta+3\alpha)}\right]\mathcal{A}^2\bigg\}. \end{aligned} \tag{3.105}$$

If we require $H_R > 0$, which means that every term in eq. (3.105) needs to be positive definite, there are two possibilities:

- $\beta < 0$, $-\frac{\beta}{3} < \alpha \leq -\frac{2\beta}{5}$, which is not compatible with the condition we require that the helicity-1, and 2 modes to be stable, $\beta > 0$.
- $\beta > 0$, $\beta + 3\alpha > 0$, we can use this parameter choice to have stable helicity-0, 1, and 2 modes.

Choosing either possibility will result in a stable helicity-0 sector.

3.7 Conclusion and future direction

We investigate the instability in higher derivative gravity models with quadratic curvature invariant R^2, $R^{\mu\nu}R_{\mu\nu}$ by expanding the action to the quadratic level of metric fluctuation around the Minkowski/de Sitter backgrounds. We show how the instability in the helicity-0, 1, and 2 sectors can be removed by choosing additional constraints. With the help of these constraints, the degrees of freedom are reduced from *two* helicity-0, *two* helicity-1, and *four* helicity-2 to *one* helicity-0, *zero* helicity-1, and *two* helicity-2 modes. The fact that the phase space has to be reduced – i.e. it is impossible to modify the theory via constraints such that the instabilities are "made stable" – is an expression of the theorem proven in [66] that the Ostrogradski instability can only be removed if the original theory's phase space is reduced.

We emphasize that adding constraints to remove instability is only valid in the linear theory. A full non-linear extension of this methodology is beyond the scope of this dissertation, and we have made no attempt to produce a covariant formalism. However, even in the linear theory, some features of a stable higher derivative gravity can be gleaned. First, it is clear that a general higher derivative theory which is stable and possesses the desirable renormalization properties breaks the Lorentz invariance. Indeed, the "stabilized" theory has the form of a low energy effective limit of a Lorentz violation, much like that of Hořava gravity.

Second, the stable higher derivative theory has no helicity-1 modes, at least in the Minkowski case, since this mode is unstable in the original theory and hence must be removed. The de Sitter case is less clear-cut – the helicity-1 sector may be made stable by the curvature term although we have chosen to remove it in order to be consistent with the stability of the helicity-2 sector.

Although the full non-linear extension of this procedure is not yet known, we would like to point out several directions for potential future research. First, one could expand the action (3.1) around the constant curvature background to the next order. The action would be cubic in the metric perturbation $h_{\mu\nu}$ and one can check if the stabilization procedure remains valid. If the procedure is valid for the next order, one can see whether there is a pattern that we can use to generate the full non-linear extension. The second possible thread of study is to stabilize the linear action covariantly; although the theory would not improve the renormalization properties, it would be easier to find the pattern and thus generalize the methodology to the full non-linear theory. It will be very interesting to check whether this result can be extended to the full non-linear regime.

Chapter 4

Stability of Constrained Dynamics in a Spherically Symmetric Static Metric

In this chapter, we study the spherically symmetric static solution of a new class of two scalar-field theory, where one of them is a Lagrange multiplier enforcing a constraint relating the value of the other scalar field to the norm of its derivative. Because of the constraint, there is only a single degree of freedom in the theory, and the spherically symmetric static solution can be found with an appropriate exponential potential. We also study the perturbation around the spherically symmetric static solution and find that the odd modes are stable against the perturbation while the even modes are not.

4.1 Introduction

The theory of cosmological perturbation is one of the most popular topics in the field of theoretical cosmology, which can be used to connect the primordial quantum fluctuation of any model of inflation and the metric perturbation during the inflation period with the Cosmic Microwave Background Radiation (CMB). This technique was developed for more than 30 years ago and has been applied to different candidates of inflation such as single scalar field inflation, multi-scalar field inflation, k-essence, and modified gravity in Friedmann-Lemaître-Robertson-Walker (FLRW) background metric.

On the other hand, it is also important to study the perturbation of the spherically symmetric static background solution – which describes a time-independent spherically symmetric object which can be a star or a dark matter halo. Although some authors have studied the spherically symmetric static solution of k-essence [85], Galileon [86] and Chameleon [87], the perturbation of these solutions has only been studied up to the equation of motion level, i.e. the authors only perturb the equations of motion and check the conditions for the absence of tachyonic instability rather than perturb the action itself, which provides information about the no-ghost condition. Nevertheless, it is natural to check whether any new models can form stable spherically symmetric static objects, i.e. dark matter haloes or "unknown stars" if the models can describe the inflation, dark energy or other prominent phenomena well.

In this chapter, we study a new class of constrained scalar-field theories with Lagrangian [25]

$$S = \int d^4x\sqrt{-g}\,[-K(\varphi,X) - \lambda(X+V(\varphi))]\,, \tag{4.1}$$

which we will introduce in Section 4.2 and refer it as $\lambda\varphi$-fluid through this chapter. The energy of $\lambda\varphi$-fluid always flows along timelike geodesics, like normal dust, except for the fact that the former has non-zero pressure, and it admits a cosmological solution, unifying the dark energy and cold dark matter into a single degree of freedom. Although it is accepted that all of the known particles cannot form stable self-gravitating static spherical systems, it is unclear whether or not the spherically symmetric static system formed by $\lambda\varphi$-fluid will be unstable because of its pressure support. Therefore, in this chapter, we check whether the system suffers from the tachyonic and ghost instability, and so need to demonstrate the formalism for this purpose.

In Section 4.2, we introduce the $\lambda\varphi$-fluid and the basic property of the cosmological solution. In Section 4.3, we review the Regge-Wheeler-Zerilli decomposition, which is the standard formalism for analyzing the perturbation around a spherically symmetric static metric. We will outline the definition of the “odd" and “even" modes according to their behaviors under parity transformation and identify them in scalar, vector, and tensor spherical harmonics. At the linear level, one can see that the odd and even modes do not interact with each other. In Section 4.4, we briefly review the recent work, where the Regge-Wheeler-Zerilli decomposition is used to study the stability of a spherically symmetric static solution in f(R, $\mathcal{G}$) theory, where R is the Ricci scalar and $\mathcal{G}$ the Gauss-Bonnet term. In Section 4.5, we apply the Regge-Wheeler-Zerilli decomposition to a special class of $\lambda\varphi$-fluid – *Dust of Dark Energy* – and demonstrate both how we find the background solution and why it is unstable against the even mode perturbation. A conclusion is provided in Section 4.6.

The technical details of the computations presented in this chapter are provided in Appendix D, where we explicitly include all of the coefficients discussed in Section 4.5. Throughout this chapter, we will use mostly plus metric convention, while the Greek indices are used to denote the spacetime label (0,1,2,3) and the Latin indices are used for the two-sphere label (2,3).

4.2 Dusty fluid with pressure

Dusty fluid with pressure [25] is a new class of scalar-field model which shares the same characteristic as normal dust, where energy always flows along timelike geodesics but can possess non-zero pressure. The model, which we will call $\lambda\varphi$-fluid, is described by a Lagrangian which contains two scalar fields, φ and λ. The former is a scalar field similar to the k-essence model but constrained by the latter, which plays the role of a Lagrange multiplier in the Lagrangian. Explicitly, the $\lambda\varphi$-fluid model is given by the action

$$S = \int d^4x\sqrt{-g}\left[-K(\varphi,X) - \lambda\left(X - \frac{1}{2}\mu^2(\varphi)\right)\right], \tag{4.2}$$

where the field λ does not have a kinetic term and is thus a Lagrange multiplier, which gives a constraint on φ and its first derivative, while

$$X = -\frac{1}{2}g^{\mu\nu}\nabla_\mu\varphi\nabla_\nu\varphi \tag{4.3}$$

is the standard kinetic term. The equations of motion and energy-momentum tensor are

$$\frac{1}{\sqrt{-g}}\frac{\delta S}{\delta\lambda} = -\left(X - \frac{1}{2}\mu^2(\varphi)\right) = 0, \tag{4.4}$$

$$\frac{1}{\sqrt{-g}}\frac{\delta S}{\delta\varphi} = -\left[K_\varphi + \nabla_\nu(K_X\nabla^\nu\varphi) - \lambda\mu\mu_\varphi + \nabla_\nu(\lambda\nabla^\nu\varphi)\right] = 0, \tag{4.5}$$

$$T_{\nu\rho} = \frac{2}{\sqrt{-g}}\frac{\delta S}{\delta g^{\nu\rho}} = (K_X+\lambda)\nabla_\nu\varphi\nabla_\rho\varphi + Kg_{\nu\rho}, \tag{4.6}$$

where K_φ, K_X denote functional derivative of $K(\varphi,X)$ with respect to φ and X.

In the case of timelike X, similar to k-essence, one can introduce the convenient four-velocity $u_\nu = -\frac{\nabla_\nu\varphi}{\sqrt{2X}} = -\mu^{-1}\nabla_\nu\varphi$, and express the energy density and the pressure as

$$\varepsilon(\lambda,\varphi) = \mu^2(K_X+\lambda) - K, \tag{4.7}$$

$$p(\varphi) = K(\varphi, \frac{\mu^2(\varphi)}{2}). \tag{4.8}$$

If we assume a background cosmology of pure $\lambda\varphi$-fluid in FLRW spacetime with metric

$$ds^2 = -dt^2 + a^2(t)d\mathbf{x}^2, \tag{4.9}$$

the equations of motions eqs. (4.4) and (4.5) become

$$\dot{\varphi} = \mu(\varphi), \tag{4.10}$$

$$\dot{\lambda} = -\mu^{-2}[\varepsilon_\varphi \mu + 3H(\varepsilon + p)]. \tag{4.11}$$

In order to find a scaling solution of the $\lambda\varphi$-fluid, the authors consider the following class of model

$$K = \sigma X, \quad \text{where} \quad \sigma = \pm 1, \tag{4.12}$$

$$\mu = \mu_0 \exp(-\frac{\varphi}{m}) \tag{4.13}$$

where m is the mass scale for φ

$$m = \sqrt{\frac{8}{3}} \frac{\sqrt{\sigma w_{\text{fin}}}}{1 + w_{\text{fin}}} M_{\text{Pl}}. \tag{4.14}$$

The dynamics of the $\lambda\varphi$-fluid-dominated cosmological background, with a constant equation of state w_{fin}, have a fixed point. The equation of state w_{fin} can have either sign and even be phantom-like.

The eqs. (4.7) and (4.8) in this setup become

$$\varepsilon = \mu^2 \left(\frac{\sigma}{2} + \lambda \right), \tag{4.15}$$

$$p = \frac{\sigma}{2} \mu^2, \tag{4.16}$$

and the instantaneous equation of state $w_{\text{X}} = \frac{p}{\varepsilon} = \frac{1}{1+2\sigma\lambda}$ is dependent on the value of λ. The equation of motion can thus be written as

$$\lambda_{w_{\text{X}}} = \frac{1}{2} \sigma (w_{\text{X}}^{-1} - 1), \tag{4.17}$$

$$\mu = \frac{m}{t} \quad \text{and} \quad \mu_\varphi = -\frac{\mu}{m} = -\frac{1}{t}, \tag{4.18}$$

where the integration constants of $\mu(\varphi)$ have been chosen such that the pressure p is singular at the Big Bang, $t = 0$. The equation of motion for the equation of state can be obtained by using the Friedmann equation and eqs. (4.10), (4.11), (4.16) and (4.17)

$$w'_{\text{X}} = 3 w_{\text{X}} \left[1 + w_{\text{X}} - \sqrt{\frac{w_{\text{X}}}{w_{\text{fin}}}} (1 + w_{\text{fin}}) \right], \tag{4.19}$$

with $()' = \partial_N()$ the derivative with respect to the number of e-folds ($N \equiv \ln a$). One can solve eq. (4.19) and obtain the scale factor a as a function of the equation of state w_X

$$\left(\frac{a}{a_0}\right)^{3(w_{fin}-1)} = \left[\frac{\left(\sqrt{\frac{w_{\rm fin}}{w_X}} - w_{\rm fin}\right)^{w_{\rm fin}}}{\sqrt{\frac{w_{\rm fin}}{w_X}} - 1}\right]^{-2}, \tag{4.20}$$

where a_0 is an integration constant. The evolution of the effective equation of state w_X with different final attractors $w_{\rm fin}$ can be obtained by the inverse relation, as illustrated in fig. 4.1. One can see that the $\lambda\varphi$-fluid admits a scaling solution which unifies the dark matter and dark energy with a single degree of freedom, which we call *Dust of Dark Energy*.

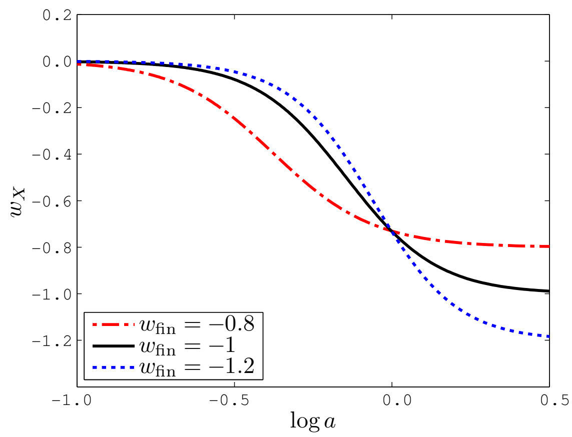

Fig. 4.1 The time evolution behavior of the total equation of state of the dark sector w_X with different fixed points $w_{\rm fin}$. The evolution is normalized such that the equation of state at $a = 1$ matches WMAP7's best-fit data for the Λ CDM cosmology, $w_0 = -0.74$. (Taken from [25])

The cosmological perturbations of *Dust of Dark Energy* are considered in [25] and the authors conclude that this model recovers the standard result for general hydrodynamics in the limit of the vanishing speed of sound. By writing down the Lagrangian for perturbations to the second order, one can find that there is no ghost in this theory when the equation of state for the $\lambda\varphi$-fluid is non-phantom, i.e. ($w \geq -1$).

4.3 Review of the Regge-Wheeler-Zerilli Decomposition of Perturbed Metric

The Regge-Wheeler-Zerilli's decomposition of metric perturbation [88, 89] is the standard way to study the stability of a spherically symmetric static system. The basic idea of this formalism is to write the metric perturbations as scalar, vector, and tensor spherical harmonics according to their transformation properties under the two-dimensional rotations. With this decomposition, one can separate the odd and even modes according to their behavior under the parity transformation and, at the linear level, they are decoupled from each other.

In order to study the stability issue of spherically symmetric static systems, we first write our total metric as a background spherically symmetric static metric plus metric perturbation

$$\begin{aligned} g_{\mu\nu} &= \bar{g}_{\mu\nu} + h_{\mu\nu}, \\ g^{\mu\nu} &= \bar{g}^{\mu\nu} - h^{\mu\nu}, \end{aligned} \tag{4.21}$$

where the background metric $\bar{g}_{\mu\nu}$ is the most general form of spherically symmetric static metric

$$\bar{g}_{\mu\nu} = diag(-e^{2\alpha(r)}, e^{2\beta(r)}, r^2, r^2\sin^2\theta). \tag{4.22}$$

Under the two-dimensional rotations on the sphere, the metric perturbations h_{tt} h_{tr}, h_{rr} transform like scalars, h_{ta}, h_{ra} transform like two-component vectors, and h_{ab} transforms as a tensor, where a, b denote either θ or ϕ.

We start by decomposing the scalar function into the sum of spherical harmonics

$$S(t,r,\theta,\phi) = \sum_{L,M} F_{LM}(t,r) Y^{LM}(\theta,\phi), \tag{4.23}$$

where $F_{LM}(t,r)$ is the arbitrary function of t,r, and $Y^{LM}(\theta,\phi)$ are the spherical harmonics. The parity operator, $\mathbf{P}$, is defined by $\mathbf{P}Y^{LM}(\theta,\phi) = Y^{LM}(\pi-\theta,\pi+\phi) = cY^{LM}(\theta,\phi)$, and it is well-known that the spherical harmonics $Y^{LM}(\theta,\phi)$ have “even" parity $c = (-1)^L$.

In order to build the vector spherical harmonics, one should know that there is more than one definition of it, and the one mostly tied to the rotation group is obtained by coupling the scalar spherical harmonics of order L' to the basis vectors

$$X^0 \equiv e_z, \quad X^{\pm 1} \equiv \mp(e_x \pm i e_y)/\sqrt{2}, \tag{4.24}$$

which transform under an irreducible representation of order 1. One can thus build vector spherical harmonics $Y^{L',LM}$ which transform under an irreducible representation of order $L = L'-1, L', L'+1$

$$Y^{L',LM}(\theta,\phi) = \sum_{M'=-L'}^{L'} \sum_{M''=-1}^{1} \langle 1M'', L'M'|LM\rangle X^{M''} Y^{L'M'}, \tag{4.25}$$

where $\langle L''M'', L'M'|LM\rangle$ are the Clebsch-Gordan coefficients

$$\begin{aligned}\langle L''M'', L'M'|LM\rangle = \delta_{M,M'+M''} \sqrt{\frac{(2L+1)(L+L'-L'')!(L-L'+L''|)!(L'+L''-L)!}{(L+L'+L''+1)!}} \times \\ \times \sqrt{(L+M)!(L-M)!(L'+M')!(L'-M')!(L''+M'')!(L''-M'')!} \times \\ \times \sum_K \frac{(-1)^K}{K!(L'+L''-K-L)!(L'-M'-K)!(L''+M''-K)!(L-L'-M''+K)!(L-L''+M'+K)!}.\end{aligned}$$

The vector spherical harmonics under the parity operator are $\mathbf{P}Y^{L',LM} = (-1)^{L'+1}Y^{LM}$, which can be obtained by the combination of scalar spherical harmonics ($\mathbf{P}Y^{LM} = (-1)^L Y^{LM}$) and the basis vector ($\mathbf{P}X^{M''} = -X^{M''}$), which gives it an extra minus sign for vector spherical harmonics. These so-called “pure-orbital" vector spherical harmonics are closely related to the solutions of Laplace’s equation and vector wave equation, but are not optimally designed for separating modes with different parities. We can use them to build the “pure-spin" vector spherical harmonics, since they are either purely radial or purely transverse and have either odd or even type parity. The “pure-spin" vector harmonics are defined by

$$\begin{aligned} Y^{E,LM} &= (2L+1)^{-\frac{1}{2}} \left[(L+1)^{\frac{1}{2}} Y^{L-1,LM} + L^{\frac{1}{2}} Y^{L+1,LM} \right] \\ &= [L(L+1)]^{-\frac{1}{2}} r \nabla Y^{LM}, \end{aligned} \tag{4.26}$$

$$Y^{B,LM} = iY^{L,LM} = [L(L+1)]^{-\frac{1}{2}} \mathbf{x} \times \nabla Y^{LM}, \tag{4.27}$$

$$\begin{aligned} Y^{R,LM} &= (2L+1)^{-\frac{1}{2}} \left[-(L+1)^{\frac{1}{2}} Y^{L-1,LM} + L^{\frac{1}{2}} Y^{L+1,LM} \right] \\ &= \mathbf{n} Y^{LM}, \end{aligned} \tag{4.28}$$

where $Y^{E,LM}$ and $Y^{B,LM}$ are purely transverse and $Y^{R,LM}$ is purely radial. $Y^{E,LM}$ and $Y^{R,LM}$ have even-type parity with $c = (-1)^L$ and $Y^{B,LM}$ has odd-type parity $c = (-1)^{L+1}$. Since the vector on the two-sphere is pure transverse, we can use $Y^{E,LM}$ and $Y^{B,LM}$ to write any two-vector function V_a as

$$V_a(t,r,\theta,\phi) = \nabla_a \Phi_1 + E_a^b \nabla_b \Phi_2, \tag{4.29}$$

where Φ_1, Φ_2 are scalar functions associated with $Y^{E,LM}$ and $Y^{B,LM}$ respectively, and defined by

$$\Phi_1 \equiv f_E(t,r)\,[L(L+1)]^{-\frac{1}{2}}\,Y^{LM}(\theta,\phi), \tag{4.30}$$

$$\Phi_2 \equiv f_B(t,r)\,[L(L+1)]^{-\frac{1}{2}}\,Y^{LM}(\theta,\phi), \tag{4.31}$$

where $f_E(t,r), f_B(t,r)$ are arbitrary functions of t,r and $E_{ab} \equiv \sqrt{\gamma}\varepsilon_{ab}$. γ_{ab} is the two-dimensional metric on the two-sphere, ε_{ab} is the totally anti-symmetric symbol where $\varepsilon_{\theta\phi} = 1$ while ∇_a is the covariant derivative for the metric on the two-sphere. Since the transverse vector function has only two components, it is completely determined by Φ_1 and Φ_2. One can also see that the first term in eq. (4.29) has even parity and the second term has odd parity.

Using similar decomposition, Mathews [90] constructed a set of “pure-orbital" tensor spherical harmonics by first coupling the basis vectors X^M to obtain five symmetric basis tensors t^M, which transform under an irreducible representation of the rotation group of order 2

$$t^M = \sum_{M'=-1}^{1}\sum_{M''=-1}^{1} \langle 1M'', 1M'|2M\rangle X^{M'} \otimes X^{M''}, \tag{4.32}$$

and a single basis tensor which is the unit tensor and gives a trivial representation

$$3^{-1/2}\delta = -\sum_{M'=-1}^{1}\sum_{M''=-1}^{1} \langle 1M'', 1M'|00\rangle X^{M'} \otimes X^{M''}, \tag{4.33}$$

where δ is the Euclidean metric. By coupling these basis tensors to the scalar spherical harmonics, one can obtain the six basis tensor spherical harmonics

$$\begin{aligned} T^{2L',LM} &= \sum_{M'=-L'}^{L'}\sum_{M''=-2}^{2} \langle L'M', 2M''|LM\rangle Y^{L'M'} t^{M''} \quad [L' = L \pm (0,1,\text{or } 2)], \\ T^{0L,LM} &= -Y^{LM} 3^{-1/2}\delta. \end{aligned} \tag{4.34}$$

Similarly, one can use them to define six “pure-spin" tensor harmonics

$$\begin{aligned} T^{R\,0,LM}_{\mu\nu} = &\left(\frac{(L+1)(L+2)}{(2L+1)(2L+3)}\right)^{\frac{1}{2}} T^{2\,L+2,LM}_{\mu\nu} - \left(\frac{2L(L+1)}{3(2L-1)(2L+3)}\right)^{\frac{1}{2}} T^{2\,L,LM}_{\mu\nu} \\ &+ \left(\frac{L(L-1)}{(2L-1)(2L+1)}\right)^{\frac{1}{2}} T^{2\,L-2,LM}_{\mu\nu} - \frac{1}{3^{1/2}} T^{0\,L,LM}_{\mu\nu} \end{aligned}$$

$$= n_\mu n_\nu Y^{LM}, \tag{4.35}$$

$$\begin{aligned}
T_{\mu\nu}^{T\ 0,LM} &= -\left(\frac{(L+1)(L+2)}{2(2L+1)(2L+3)}\right)^{\frac{1}{2}} T_{\mu\nu}^{2\ L+2,LM} + \left(\frac{L(L+1)}{3(2L-1)(2L+3)}\right)^{\frac{1}{2}} T_{\mu\nu}^{2\ L,LM} \\
&\quad -\left(\frac{L(L-1)}{2(2L-1)(2L+1)}\right)^{\frac{1}{2}} T_{\mu\nu}^{2\ L-2,LM} - \left(\frac{2}{3}\right)^{1/2} T_{\mu\nu}^{0\ L,LM} \\
&= 2^{-1/2}(\delta_{\mu\nu} - n_\mu n_\nu) Y^{LM},
\end{aligned} \tag{4.36}$$

$$\begin{aligned}
T_{\mu\nu}^{E\ 1,LM} &= -\left(\frac{2L(L+2)}{(2L+1)(2L+3)}\right)^{\frac{1}{2}} T_{\mu\nu}^{2\ L+2,LM} - \left(\frac{3}{(2L-1)(2L+3)}\right)^{\frac{1}{2}} T_{\mu\nu}^{2\ L,LM} \\
&\quad + \left(\frac{2(L-1)(L+1)}{(2L-1)(2L+1)}\right)^{\frac{1}{2}} T_{\mu\nu}^{2\ L-2,LM} \\
&= \left(\frac{2}{L(L+1)}\right)^{1/2} [n_\mu r \nabla_\nu Y^{LM}]^S,
\end{aligned} \tag{4.37}$$

$$\begin{aligned}
T_{\mu\nu}^{B\ 1,LM} &= i\left(\frac{L+2}{2L+1}\right)^{\frac{1}{2}} T_{\mu\nu}^{2\ L+1,LM} - i\left(\frac{L-1}{2L+1}\right)^{\frac{1}{2}} T_{\mu\nu}^{2\ L-1,LM} \\
&= [2\varepsilon_{\mu\sigma\rho} n_\sigma T_{\rho\nu}^{E\ 1,LM}]^S,
\end{aligned} \tag{4.38}$$

$$\begin{aligned}
T_{\mu\nu}^{E\ 2,LM} &= \left(\frac{L(L-1)}{2(2L+1)(2L+3)}\right)^{\frac{1}{2}} T_{\mu\nu}^{2\ L+2,LM} + \left(\frac{3(L-1)(L+2)}{(2L-1)(2L+3)}\right)^{\frac{1}{2}} T_{\mu\nu}^{2\ L,LM} \\
&\quad + \left(\frac{(L+1)(L+2)}{2(2L-1)(2L+1)}\right)^{\frac{1}{2}} T_{\mu\nu}^{2\ L-2,LM} \\
&= \left(2\frac{(L-2)!}{(L+2)!}\right)^{\frac{1}{2}} [r^2 \nabla_\mu \nabla_\nu Y^{LM}]^{STT},
\end{aligned} \tag{4.39}$$

$$\begin{aligned}
T_{\mu\nu}^{B\ 2,LM} &= -i\left(\frac{L-1}{2L+1}\right)^{\frac{1}{2}} T_{\mu\nu}^{2\ L+1,LM} - i\left(\frac{L+2}{2L+1}\right)^{\frac{1}{2}} T_{\mu\nu}^{2\ L-1,LM} \\
&= [\varepsilon_{\mu\sigma\rho} n_\sigma T_{\rho\nu}^{E\ 2,LM}]^S.
\end{aligned} \tag{4.40}$$

Here the superscript S denotes the "symmetric part of ", TT means the "transverse traceless part of", and the characteristics of the harmonics are

$T^{R\ 0,LM}$: pure radial with "even" parity, $c = (-1)^L$,
$T^{T\ 0,LM}$: pure transverse with "even" parity,
$T^{E\ 1,LM}$: mixed radial and transverse with "even" parity

$$
\begin{aligned}
T^{B\ 1,LM} &: \text{mixed radial and transverse with “odd” parity, } c=(-1)^{L+1},\\
T^{E\ 2,LM} &: \text{transverse and traceless with “even” parity},\\
T^{B\ 2,LM} &: \text{transverse and traceless with “odd” parity}.
\end{aligned}
\tag{4.41}
$$

Since we only need the pure transverse part of the tensor spherical harmonics ($T^{T\ 0,LM}$, $T^{E\ 2,LM}$, $T^{B\ 2,LM}$) for the decomposition on the 2-sphere, we can thus write any traverse 2×2 tensor as follows

$$
T_{ab}(t,r,\theta,\phi) = \nabla_a\nabla_b\Psi_1 + \gamma_{ab}\Psi_2 + \frac{1}{2}(E_a^c\nabla_c\nabla_b\Psi_3 + E_b^c\nabla_c\nabla_a\Psi_3), \tag{4.42}
$$

where the transverse traceless part of the first term corresponds to $T^{E\ 2,LM}$, the second term corresponds to $T^{T\ 0,LM}$, and the third term corresponds to $T^{B\ 2,LM}$. Since we have only three components for the symmetric 2×2 metric, the tensor is determined by Ψ_1, Ψ_2, Ψ_3, where the terms with E_{ab} have odd parity and the others have even parity.

Now we are ready to write down the general form of different types of metric perturbations. Using the Regge-Wheeler formalism, the metric perturbation with odd-type parity can be written by

$$
h_{tt} = h_{tr} = h_{rr} = 0,
$$

$$
h_{ta} = \sum_{L,M} h_{0,LM}(t,r)E_{ab}\nabla^b Y^{LM}(\theta,\phi), \tag{4.43}
$$

$$
h_{ra} = \sum_{L,M} h_{1,LM}(t,r)E_{ab}\nabla^b Y^{LM}(\theta,\phi), \tag{4.44}
$$

$$
h_{ab} = \frac{1}{2}\sum_{L,M} h_{2,LM}(t,r)[E_a^c\nabla_c\nabla_b Y^{LM}(\theta,\phi) + E_b^c\nabla_c\nabla_a Y^{LM}(\theta,\phi)]. \tag{4.45}
$$

By using the gauge symmetry in general relativity,[1] we can remove the redundant degrees of freedom by considering the following gauge transformation

$$
\xi_t = \xi_r = 0, \qquad \xi_a = \sum_{LM}\Lambda_{LM}(t,r)E_a^b\nabla_b Y^{LM}. \tag{4.46}
$$

We can use Λ_{LM} to eliminate $h_{2,LM}$ which fixes all of the Λ_{LM}, and there is no remaining gauge degree of freedom. After fixing the gauge, we can further simplify our calculations by realizing that every value of M contributes equally to the Lagrangian after integrating out

[1] I.e. $\exists$ infinitesimal gauge transformation $x^\mu \to x^\mu + \xi^\mu$, represent the same physical system.

θ and ϕ. We can thus focus on $M=0$, with the advantage that ϕ will totally disappear from the calculations.[2] The odd-type metric perturbation can thus be simplified by

$$h_{\mu\nu}^{\text{odd}} = \begin{bmatrix} 0 & 0 & 0 & \sum_L N h_0(t,r)\sin\theta(\partial/\partial\theta)P_L(\cos\theta) \\ 0 & 0 & 0 & \sum_L N h_1(t,r)\sin\theta(\partial/\partial\theta)P_L(\cos\theta) \\ \sum_L N h_1(t,r)\sin\theta(\partial/\partial\theta)P_L(\cos\theta) & 0 & 0 & 0 \\ \sum_L N h_0(t,r)\sin\theta(\partial/\partial\theta)P_L(\cos\theta) & 0 & 0 & 0 \end{bmatrix}, \tag{4.47}$$

where $N=\sqrt{\frac{2L+1}{4\pi}}$ is the normalization constant for spherical harmonics.

Following the same procedure, the even-type metric perturbation can be written as

$$h_{\mu\nu}^{\text{even}} = \begin{bmatrix} -e^{2\alpha(r)}H_0(t,r,\theta) & H_1(t,r,\theta) & 0 & 0 \\ H_1(t,r,\theta) & e^{2\beta(r)}H_2(t,r,\theta) & K(t,r,\theta) & 0 \\ 0 & K(t,r,\theta) & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{4.48}$$

where

$$H_i = \sum_L N H_{iL}(t,r)P_L(\cos\theta), \quad (i=0,1,2) \tag{4.49}$$

$$K = \sum_L N K_L(t,r)\partial_\theta P_L(\cos\theta)(\theta,\phi). \tag{4.50}$$

We have used the gauge transformation to eliminate h_{ta} and h_{ab}. $h_{r\phi}$ vanishes because we choose the preferable $M=0$ to perform the calculations. Similar to odd-type perturbation, the double summation of M in the action will eventually contribute a factor $(2L+1)$. Note that, while Regge-Wheeler-Zerilli chose a slightly different gauge which eliminates h_{ta}, h_{ra}, and $h_{\theta\phi}$, the idea is the same.

[2] Since we perturb the action to the second order, all of the terms in the action are $\mathcal{O}(h^2)$, which means that there is a double summation over index M, M'. We will use one to pick up $\delta_{MM'}$ and the other will give us the factor $(2L+1)$. On the other hand, we can absorb the factor $\sqrt{2L+1}$ in each h and forget about the summation over M.

4.4 Recent Work on the Stability of Spherically Symmetric Static Backgrounds

The perturbation formalism was introduced and demonstrated in Section 4.3 [88–92]. Here, we briefly demonstrate the formalism by studying the model in [91] and leave the whole analysis procedure to the next section, perturbation of the "$\lambda\varphi$-fluid". By using the Regge-Wheeler-Zerilli metric decomposition and perturbation analysis, we study the f(R, $\mathcal{G}$) model, where R is the Ricci scalar and $\mathcal{G}$ the Gauss-Bonnet term. We take this model as a reference and at some limit, our calculations in the next section should recover the result in [91].

The Lagrangian of this model can be written in terms of Lagrange multipliers

$$S = \frac{M_p^2}{2}\int d^4x\sqrt{-g}[FR + \xi\,\mathcal{G} - U(F,\xi)], \tag{4.51}$$

where F and ξ are scalar fields coupled to R and $\mathcal{G}$, and $\mathcal{G} \equiv R^2 - 4R + R_{\mu\nu}R^{\mu\nu} + R_{\mu\nu\sigma\rho}R^{\mu\nu\sigma\rho}$. To study the perturbation around the spherically symmetric static background solution, we again use the most general background metric

$$ds^2 = \bar{g}_{\mu\nu}dx^\mu dx^\nu = -A(r)dt^2 + \frac{dr^2}{B(r)} + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2, \tag{4.52}$$

where we redefine $A(r) \equiv e^{2\alpha(r)}$ and $B(r) \equiv e^{-2\beta(r)}$ in eq. (4.22). The background equations of motion can be derived from the Lagrangian

$$\begin{aligned}
U &= -\frac{4B\xi'A'}{Ar^2} + \frac{12B^2\xi'A'}{Ar^2} - \frac{4BF'}{r} - \frac{2BFA'}{Ar} - \frac{BF'A'}{A} + \frac{2F}{r^2} - \frac{2BF}{r^2},\\
F'' &= -\frac{2B'\xi'}{r^2B} - \frac{F'B'}{2B} + \frac{B\xi''}{4r^2} - \frac{4\xi''}{r^2} - \frac{FB'}{rB} + \frac{2\xi'A'}{Ar^2} + \frac{F'A'}{2A} - \frac{6\xi'BA'}{Ar^2} + \frac{FA'}{Ar} + \frac{6B'\xi'}{r^2},\\
R &= \frac{\partial U}{\partial F},\\
\mathcal{G} &= \frac{\partial U}{\partial \xi},
\end{aligned}$$

where $'$ denotes differentiation with respect to r. The action of the odd-type perturbation can be obtained by substituting the odd-type metric perturbations eq. (4.47) into the action (4.51), expanding it to the second order, performing integration by parts and ignoring all of the boundary terms. After integrating out the θ and ϕ, the action becomes

$$S_{odd} = \frac{M_p^2}{2} \sum_{LM} \int dtdr \left[C_1(\dot{h_1} - h_0')^2 + C_2 h_0 \dot{h}_1 + C_3 h_0^2 - C_4 h_1^2 \right], \tag{4.53}$$

where we omit the L, M suffixes and C_i are functions of r only. The no-ghost condition in the model is

$$C_4 \geq 0, \quad \text{or equivently} \quad AF - 2B\xi' A' \geq 0. \tag{4.54}$$

For the tachyonic stability, considering a solution proportional to $e^{i(wt-kr)}$, at large w and k limit, the radial dispersion relation is

$$w^2 = \frac{B(AF - 2B\xi' A')}{F - 4B\xi'' - 2B'\xi'} k^2, \tag{4.55}$$

and the radial speed of sound is thus

$$c_{odd}^2 = \frac{(AF - 2B\xi' A')}{A(F - 4B\xi'' - 2B'\xi')}. \tag{4.56}$$

If $(F - 4B\xi'' - 2B'\xi') > 0$, there is no tachyonic and ghost instability for the odd modes.

On the other hand, for the action of the even-modes, one needs to substitute even-type metric perturbation eq. (4.48) and scalar fields perturbations ($F = F(r) + \delta F$, $\xi = \xi(r) + \delta\xi$) into the action, there are **six** fields in total (**two** from the scalar perturbation: δF, $\delta\xi$ and **four** from the even-type metric perturbation: H_0, H_1, H_2, K). After performing integration by parts and integrating out θ and ϕ dependence, one can further simplify it by realizing that H_0, H_1 are non-propagation fields (Lagrange multipliers) and do not interact with each other. By using this fact, we can replace H_1 by other fields and use the constraints given by the Lagrange multiplier H_0 to replace H_2 by other fields. Substituting H_2 back into the Lagrangian not only gets rid of H_2, but also automatically eliminates the Lagrange multiplier H_0. The total degree of freedom becomes $6 - 3 = 3$ and we can write down the Lagrangian as

$$\mathcal{L}_{\text{even}} = \sum_{i,j=1}^{3} [K_{ij}(r)\dot{v}_i\dot{v}_j - L_{ij}(r)v_i'v_j' - D_{ij}(r)v_i'v_j - M_{ij}(r)v_iv_j], \tag{4.57}$$

where v_i's are combinations of $\delta F, \delta\xi$, metric perturbation H_2, and K, and one should remember that there are only three independent degrees of freedom in the Lagrangian. All of the matrices in eq. (4.57) can be made symmetric except for D_{ij}, which is anti-symmetric. For the no-ghost condition, we need the matrix K_{ij} to be positive definite so, by using the

Sylvester criterion, the no-ghost condition is translated into

$$det(K_{ij}) > 0, \quad K_{22}K_{33} - K_{23}^2 > 0, \quad K_{33} > 0. \tag{4.58}$$

For the model we studied, the second inequality $K_{22}K_{33} - K_{23}^2 > 0$ can never be satisfied, thus there is always a ghost for each "L".

For the tachyonic stability, by assuming solutions are proportional to $e^{i(wt-kr)}$, one can find the dispersion relation at a large w and k limit by solving

$$det(w^2K_{ij} - k^2L_{ij}) = 0, \tag{4.59}$$

which is a cubic equation in w^2. We can find three solutions with different radial speeds of propagation

$$\begin{aligned} c_1^2 = c_2^2 &= \frac{(2AB - 2A - rBA')}{(2AB - 2A - rAB')}, \\ c_3^2 &= \frac{(AF - 2A'B\xi')}{(AF - 2AB'\xi' - 4AB\xi'')}. \end{aligned} \tag{4.60}$$

c_1^2 and c_2^2 reduce to unity (speed of light) when $A(r) = B(r)$, and c_3^2, like the odd mode, depends on the behavior of the two new scalar fields.

4.5 $\lambda\varphi$-fluid in a Spherically Symmetric Static Background

In this section, we study the stability of the spherically symmetric static solution of the $\lambda\varphi$-fluid model. First, we demonstrate how to find the equations of motion for the general potential, then, with an ansatz potential $V(\varphi) \propto e^{k\varphi}$, we show that we can find a solution in the spherically symmetric static background metric. With the Regge-Wheeler-Zerilli decomposition, we study the odd and even perturbations around the spherically symmetric static solution and conclude that it is unstable against the even type perturbation.

4.5.1 Background solution with general potential

The action of the $\lambda\varphi$-fluid theory can be written as

$$S = \int d^4x\sqrt{-g}\left[-K(\varphi,X) - \lambda(X + V(\varphi))\right]. \tag{4.61}$$

The spherically symmetric static background solution can be found by assuming that all of the fields and metric perturbations are dependent only on r. The background solution suggests that our kinetic term $X = -\frac{1}{2}\nabla_\mu\varphi\nabla^\mu\varphi$ is spacelike (< 0), and the equations of motion by varying λ and φ are

$$X + V(\varphi) = 0, \tag{4.62}$$

$$K_\varphi + \nabla_\alpha\left[(K_X + \lambda)\nabla^\alpha\varphi\right] + \lambda V_\varphi(\varphi) = 0, \tag{4.63}$$

where eq. (4.62) suggests that $V(\varphi) > 0$. From the action, we can derive the energy-momentum tensor by varying it with respect to metric

$$T_{\alpha\beta} = \frac{2}{\sqrt{-g}}\frac{\delta S}{\delta g^{\alpha\beta}} = (K_X + \lambda)\nabla_\alpha\varphi\nabla_\beta\varphi + [K + \lambda(X + V)]g_{\alpha\beta}. \tag{4.64}$$

In order to find the spherically symmetric static solution, we start from the most general metric ansatz

$$ds^2 = -e^{2\alpha(r)}dt^2 + e^{2\beta(r)}dr^2 + r^2 d\Omega^2, \tag{4.65}$$

and use the background equation of motion to solve $\alpha(r)$ and $\beta(r)$. For simplicity, we assume the function $K = \sigma X$, where $\sigma = \pm 1$. With the constraint eq. (4.62) satisfied, the Lagrangian density $\mathcal{L}$ is equivalent to $-K$, and because λ, φ are only dependent on r, we can thus derive the energy-momentum tensor

$$T^t_t = T^\theta_\theta = T^\phi_\phi = -\mathcal{L} = K = \sigma X = -\sigma V(\varphi), \tag{4.66}$$

where we have used eq. (4.62) to get the last equality. The rr component of the energy-momentum tensor can be obtained by

$$T^r_r = -\sigma V(\varphi) + (\sigma + \lambda)(\partial_r\varphi)^2 e^{-2\beta(r)}, \tag{4.67}$$

with the Ricci tensors for the background metric are

$$R_{tt} = e^{2(\alpha-\beta)}\left(\alpha'' + \alpha'^2 - \alpha'\beta' + \frac{2}{r}\alpha'\right), \tag{4.68}$$

$$R_{rr} = -\left(\alpha'' + \alpha'^2 - \alpha'\beta' - \frac{2}{r}\beta'\right), \tag{4.69}$$

$$R_{\theta\theta} = e^{-2\beta}\left[r(\beta' - \alpha') - 1\right] + 1, \tag{4.70}$$

$$R_{\phi\phi} = R_{\theta\theta} \sin^2\theta, \tag{4.71}$$

where we use $'$ to denote derivative with respect to r.

The Ricci scalar R can be derived as usual

$$R = 2e^{-2\beta}\left[-\alpha'' - \alpha'^2 + \alpha'\beta' + \frac{2}{r}(\beta' - \alpha') - \frac{1}{r^2} + \frac{e^{2\beta}}{r^2}\right], \tag{4.72}$$

and the Einstein tensors are

$$G_{tt} = e^{2(\alpha-\beta)}\left(\frac{2}{r}\beta' - \frac{1}{r^2} + \frac{e^{2\beta}}{r^2}\right), \tag{4.73}$$

$$G_{rr} = \frac{2}{r}\alpha' + \frac{1}{r^2} - \frac{e^{2\beta}}{r^2}, \tag{4.74}$$

$$G_{\theta\theta} = e^{-2\beta}r^2\left[\alpha'' + \alpha'^2 - \alpha'\beta' - \frac{1}{r}(\beta' - \alpha')\right], \tag{4.75}$$

$$G_{\phi\phi} = G_{\theta\theta}\sin^2\theta. \tag{4.76}$$

If we consider the hydrodynamics properties of the energy-momentum tensor, we find $T^{\mu}_{\nu} = diag(-\rho, p_r, p_T, p_T)$.

The "tt" component of the Einstein's equation can be derived by

$$e^{2(\alpha-\beta)}\left[\frac{2}{r}\beta' - \frac{1}{r^2} + \frac{e^{2\beta}}{r^2}\right] = e^{2\alpha}8\pi G\rho,$$

$$\Rightarrow e^{-2\beta} = 1 - \frac{2GM(r)}{r}, \tag{4.77}$$

where M is defined by $M(r) = M_0 + 4\pi\int_{r_0}^{r}\rho x^2 dx$. Similarly, rr equation is

$$\alpha' = \frac{GM + 4\pi G p_r r^3}{r(r-2GM)}, \tag{4.78}$$

while the $\theta\theta$ $(\phi\phi)$ equation can be derived by

$$\frac{dp_r}{dr} = -\left(\frac{2}{r} + \alpha'\right)(\rho + p_r), \tag{4.79}$$

where we have used the fact that $p_T = -\rho$ from eq. (4.66) and the definition of the energy momentum tensor.

By using the constraint equation

$$-\frac{1}{2}e^{-2\beta}\varphi'^2 = X = -V(\varphi), \tag{4.80}$$

we can derive the relation

$$\begin{aligned}
\rho &= -T^t_t = \sigma V(\varphi), && (4.81)\\
p_T &= -\sigma V(\varphi), && (4.82)\\
p_r &= (\sigma+2\lambda)V(\varphi), && (4.83)\\
p'_r &= 2\lambda' V(\varphi) + (\sigma+2\lambda)V_\varphi(\varphi)\varphi'. && (4.84)
\end{aligned}$$

The requirement that the energy density is positive definite suggests that $\sigma = +1$ in this setup, and from now on we will use this convention. Substituting eqs. (4.81), (4.83) and (4.84) into eq. (4.79), we get

$$\lambda' = -\left(\frac{2}{r}+\alpha'\right)(1+\lambda) - \frac{(1+2\lambda)}{2}\frac{V_\varphi(\varphi)\varphi'}{V(\varphi)}. \tag{4.85}$$

By substituting our assumption ($K = X$) into eq. (4.63), the equation of motion can be written as

$$(1+\lambda)e^{-2\beta}\left[\varphi'' + \left(\alpha'+\beta'+\frac{2}{r}\right)\varphi'\right] + \lambda'\varphi' e^{-2\beta} = -\lambda V_\varphi(\varphi). \tag{4.86}$$

The expression of φ'' can be obtained by differentiating both sides of the constraint equation eq. (4.80) with respect to r

$$\begin{aligned}
e^{-2\beta}(-2\beta'\varphi'^2 + 2\varphi'\varphi'') &= 2V_\varphi(\varphi)\varphi',\\
\varphi'' &= V_\varphi(\varphi)e^{2\beta} + \beta'\varphi'. && (4.87)
\end{aligned}$$

In summary, the independent equations are

$$\begin{aligned}
V(\varphi) =& \frac{1}{2}e^{-2\beta}\varphi'^2,\ \ e^{-2\beta} = 1 - \frac{2GM(r)}{r},\ \ \alpha' = \frac{GM+4\pi G p_r r^3}{r(r-2GM)}\\
\lambda' =& -\left(\frac{2}{r}+\alpha'\right)(1+\lambda) - \frac{(1+2\lambda)}{2}\frac{V_\varphi(\varphi)\varphi'}{V(\varphi)}\\
-\lambda V_\varphi(\varphi) =& (1+\lambda)e^{-2\beta}\left[\varphi'' + \left(\alpha'+\beta'+\frac{2}{r}\right)\varphi'\right] + \lambda'\varphi' e^{-2\beta},
\end{aligned}$$

and we can solve $\alpha, \beta, \varphi, \lambda$ either by finding $V(\varphi)$ or $M(r)$ to satisfy the equations. In order to simplify the expression, we can define two dimensionless functions, $c(r) \equiv 8\pi G r^2 e^{2\beta} V(r)$ and $d(r) \equiv e^{2\beta} - 1$, and rewrite the following functions for the general potential $V(r)$ as

$$\begin{aligned}
2r\alpha' &= d(r) + c(r)(1+2\lambda), \\
2r\beta' &= c(r) - d(r), \\
2r(r\alpha'' + \alpha') &= 2r\beta' \left[e^{2\beta} + (1+2\lambda) 8\pi G e^{2\beta} r^2 V \right] - 8\pi G r^2 e^{2\beta} V (2r\alpha')(1+\lambda) - 16\pi G r^2 e^{2\beta} V \\
&= \left[c(r) - d(r) \right] \left[1 + d(r) + (1+2\lambda) c(r) \right] - 2c(r) - c(r)(1+\lambda) \left[d(r) + c(r)(1+2\lambda) \right].
\end{aligned}$$

4.5.2 The Background Solution with Exponential Potential

Motivated by *Dust of Dark Energy* in Section 4.2 [25], we consider our model with an exponential potential ($V(\varphi) \propto e^{k\varphi}$), where k is a constant corresponding to the ansatz $M(r) = \frac{M_* r}{r_*}$, where M_*, r_* are arbitrary constants which can be fixed by a suitable boundary condition. The equations of motion can be obtained by substituting the ansatz into eqs. (4.62), (4.77), (4.78), (4.81), (4.83) and (4.85) and using the definition of $M(r)$

$$M(r) = \frac{M_* r}{r_*} \equiv M_* Y, \qquad Y \equiv \frac{r}{r_*}, \tag{4.88}$$

$$V(r) = \rho = \frac{M_*}{4\pi r_* r^2} = \frac{M_*}{4\pi r_*^3 Y^2}, \tag{4.89}$$

$$e^{-2\beta} = \left(1 - \frac{2GM_*}{r_*} \right) = const > 0, \tag{4.90}$$

$$\varphi = \pm \varphi_* \log\left(\frac{r}{r_*} \right) = \pm \sqrt{\frac{\frac{M_*}{2\pi r_*}}{1 - \frac{2GM_*}{r_*}}} \log(Y) + \varphi_c, \tag{4.91}$$

$$\frac{d\alpha}{dY} = \frac{\frac{2GM_*}{r_*}}{1 - \frac{2GM_*}{r_*}} \frac{1+\lambda}{Y} \equiv c_* \frac{1+\lambda}{Y}, \tag{4.92}$$

$$\frac{d\lambda}{dY} = -\frac{d\alpha}{dY}(1+\lambda) - \frac{1}{Y}, \tag{4.93}$$

where we have chosen that $\sigma = 1$ and define

$$c_* = \frac{\frac{2GM_*}{r_*}}{1 - \frac{2GM_*}{r_*}} \tag{4.94}$$

as the parameter where $G \to 0$ and thus $c_* \to 0$ at the weak field limit. r_* is the boundary condition of the radius of the "star" and M_* is the total mass within the radius r_*. One should notice that our solution is only valid for $r \le r_*$ so, for $r > r_*$, we assume that our solution is asymptotic to the dark energy outside the star and thus the equation of state $w = -1$.[3]

Substituting eq. (4.92) into eq. (4.93), we can solve λ by

$$\lambda = -1 + \frac{1}{\sqrt{c_*}} \tan[-\sqrt{c_*} \log(Y)], \tag{4.95}$$

and the equation of state can be written as

$$w = \frac{p_r}{\rho} = 1 + 2\lambda = -1 + \frac{2}{\sqrt{c_*}} \tan[-\sqrt{c_*} \log(Y)]. \tag{4.96}$$

Substituting eq. (4.95) into eq. (4.92), we can solve α by

$$\alpha = C_1 + \log|\cos[-\sqrt{c_*}(\log Y)]| \tag{4.97}$$

By assuming the weak field limit, where $GM_*/r_* \ll 1$, and therefore $c_* \ll 1$. Since our system admits a star-like solution, we can fix the radius of the star by placing the boundary conditions of the equation of state at the surface. The energy density and pressure are thus determined since, from eq. (4.89), we know that $\rho = V(r) \propto Y^{-2}$. Two possible boundary conditions that we impose are

1. $w = 0$, the radial pressure vanishes at the surface. A star consisting of $\lambda\varphi$-fluid has no radial pressure to expand itself at the boundary, similar to ordinary stars. Note that this boundary condition has non-zero transverse pressure, $p_T = -\rho$.

2. $w = -1$, the radial pressure is the same as the transverse pressure at the surface of the star. A star consisting of $\lambda\varphi$-fluid under this boundary condition is asymptotic to the dark energy.

The radius of the star can be found by

$$Y_{max} = \exp\left[\frac{-1}{\sqrt{c_*}} \tan^{-1}\left(\frac{\sqrt{c_*}}{2}\right)\right] \approx 0.605 \qquad (c_* \to 0) \tag{4.98}$$

for $w(Y_{max}) = 0$, and

$$Y_{max} = 1 \tag{4.99}$$

[3]The functions $c(r), d(r)$ defined in the last subsection reduced to the same constant $c(r) = d(r) = c_*$ under the ansatz $M(r) \propto r$.

for $w(Y_{max}) = -1$.

One can see that the functions λ, α, and w all have the same argument $-\sqrt{c_*}\log(Y) \equiv z$, and the equation of state w is a tangent function of z. If we do not pick a specific interval of z, the equation of state will be discontinuous at some Y, which is unphysical. By picking the interval $z \ni [0, \frac{\pi}{2}]$, which corresponds to the interval $Y \ni [1, 0.03]$ (for $c_* \approx 0.2$), the oscillating part ($\pi/2 < \theta < \infty$) is squeezed into a small region of Y, ($Y < 0.03$). The solution is thus everywhere analytical except for a tiny region within the very small radius, which we can reduce even further by making the parameter c_* smaller.

Since our solution is only valid inside the radius r_*, to make it match the Schwarzschild metric at the boundary smoothly, we need to fix the integration constant C_1 by letting $e^{2\alpha} = e^{-2\beta}$ at $Y = Y_{max}$, so we need

$$e^{2\alpha}(Y_{max}) = e^{2C_1}\cos^2[-\sqrt{c_*}\log(Y_{max})] = \left(1 - \frac{2GM_*}{r_*}\right) = \frac{1}{1+c_*}. \tag{4.100}$$

For the $w = 0$ boundary condition, we need

$$\begin{aligned} e^{2\alpha}(Y_{max}) &= e^{2C_1}\frac{4}{c_*+4} = \frac{1}{1+c_*}, \\ \Rightarrow e^{2C_1} &= \frac{c_*+4}{4c_*+4} \approx 1 \qquad (c_* \to 0), \end{aligned} \tag{4.101}$$

while for the $w = -1$ boundary condition, we need

$$\begin{aligned} e^{2\alpha}(Y_{max} = 1) &= e^{2C_1} = \frac{1}{1+c_*} \\ \Rightarrow e^{2C_1} &= \frac{1}{1+c_*} \approx 1 \qquad (c_* \to 0). \end{aligned} \tag{4.102}$$

Therefore, through adopting suitable boundary conditions, we can find a spherically symmetric static solution for the $\lambda\varphi$-fluid model with the ansatz $M(r) \propto r$. One can also find the exponential potential from eqs. (4.89) and (4.91), where

$$V(\varphi) = \frac{M_*}{4\pi r_*^3} e^{\mp 2\sqrt{\frac{1-\frac{2GM_*}{r_*}}{\frac{M_*}{2\pi r_*}}}(\varphi-\varphi_c)}. \tag{4.103}$$

Figures 4.2 and 4.3 represent the equation of state and metric component g_{tt} as a function of Y for the isotropic pressure case ($w = -1$ at $Y = Y_{max}$), and the non-radial pressure case ($w = 0$ at $Y = Y_{max}$) can be approximately seen in Figures 4.2 and 4.3, if we set $Y_{max} \approx 0.6$.

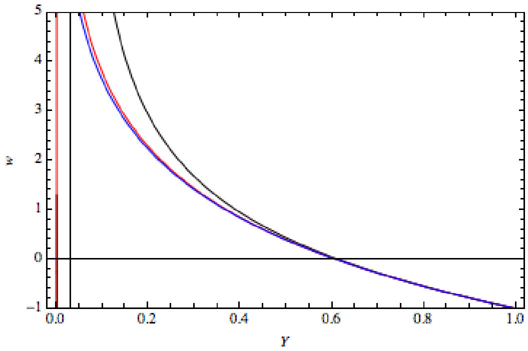

Fig. 4.2 The equation of state w as a function of Y. This is for the isotropic pressure case ($w=-1$ at $Y=Y_{max}$). The black, red, and blue lines correspond to $c_* = 0.2$, $c_* = 0.02$, and $c_* = 0.002$, where c_* is the weak-limit parameter.

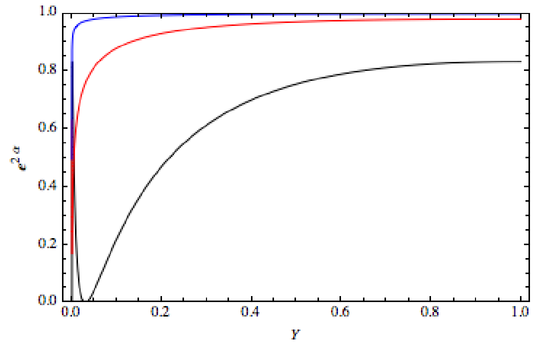

Fig. 4.3 The metric component $g_{tt} = e^{2\alpha}$ as a function of Y. This is also for the isotropic pressure case. The black, red, and blue lines correspond to $c_* = 0.2$, $c_* = 0.02$, and $c_* = 0.002$.

4.5.3 Odd-parity perturbation

The full perturbations include the perturbations of the gravity sector as well as the scalar field sector. By using the Regge-Wheeler-Zerilli decomposition and separating the modes with odd and even type parity, the parameterization of odd metric perturbation is

$$\begin{aligned}
h_{t\phi} &= h_{\phi t} = h_0(t,r)\sin\theta(\partial/\partial\theta)P_L(\cos\theta) \equiv h_0,\\
h_{r\phi} &= h_{\phi r} = h_1(t,r)\sin\theta(\partial/\partial\theta)P_L(\cos\theta) \equiv h_1,\\
h^{t\phi} &= h^{\phi t} = \frac{e^{-2\alpha}h_0}{r^2\sin^2\theta},\\
h^{r\phi} &= h^{\phi r} = -\frac{e^{-2\beta}h_1}{r^2\sin^2\theta}.
\end{aligned} \tag{4.104}$$

We can thus find the perturbed Ricci scalar and substitute it into the total action

$$S = \int d^4x\sqrt{-g}\left(\frac{R}{16\pi G} + \mathcal{L}_M\right), \tag{4.105}$$

and, since the perturbation of the scalar fields are even under the parity transformation, they do not source the odd modes, so the perturbed action is thus

$$\delta S = \int d^4x\left[\sqrt{-g}\left(\frac{\delta R}{16\pi G}\right) + \delta\sqrt{-g}\left(\frac{R}{16\pi G} + \mathcal{L}_M\right)\right]. \tag{4.106}$$

The second order action after using the background equations of motion, performing integration by parts and ignoring boundary terms, can be found to be

$$S_{odd} = \sum_{L,M} \frac{M_p^2}{2} \int dtdrA_1 \left[h_0'(t,r) - \dot{h}_1(t,r)\right]^2 + A_2 h_0(t,r)\dot{h}_1(t,r) + A_3 h_0^2(t,r) - A_4 h_1^2(t,r), \tag{4.107}$$

where we have integrated θ and ϕ, and eliminated a factor of 4π through the renormalization of spherical harmonics. The coefficients are

$$A_1 = \frac{e^{-(\alpha+\beta)}}{2} j^2, \qquad j^2 \equiv L(L+1), \tag{4.108}$$

$$A_2 = \frac{4}{r} A_1, \tag{4.109}$$

$$A_3 = e^{-(\alpha+\beta)} \left\{ \frac{j^2}{r^2}[-c(r) - d(r) - 3c(r)\lambda] + \frac{e^{2\beta}}{2r^2} j^4 \right\}, \tag{4.110}$$

$$A_4 = e^{-(\alpha+\beta)} \left\{ -\frac{e^{2(\alpha-\beta)} j^2}{r^2} (1 + d(r) + 2c(r)\lambda] + \frac{e^{2\alpha}}{2r^2} j^4 \right\}. \tag{4.111}$$

Although there are two fields in the action, one should note that h_0 is a Lagrange multiplier of the action. If we vary the action with respect to h_0 to get the constraint equation and substitute it into the Lagrangian, we should have only one degree of freedom. The constraint equation is

$$[A_1(h_0' - \dot{h}_1)]' = A_3 h_0 + \frac{1}{2} A_2 \dot{h}_1, \tag{4.112}$$

which cannot be easily solved for h_0. Nevertheless, we can use the following steps to eliminate the non-dynamical degree of freedom. First, we can rewrite the Lagrangian by

$$S_{odd} = \sum_{L,M} \frac{M_p^2}{2} \int dtdrA_1 \left(\dot{h}_1 - h_0' + \frac{2h_0}{r}\right)^2 - \frac{2(A_1 + rA_1')}{r^2} h_0^2 + A_3 h_0^2 - A_4 h_1^2, \tag{4.113}$$

where all of the terms containing $\dot{h}_1$ are in the first term. Now, we can introduce an auxiliary field Q to rewrite eq. (4.113) as

$$S_{odd} = \sum_{L,M} \frac{M_p^2}{2} \int dtdrA_1 \left[2Q\left(\dot{h}_1 - h_0' + \frac{2h_0}{r}\right) - Q^2\right] - \frac{2(A_1 + rA_1')}{r^2} h_0^2 + A_3 h_0^2 - A_4 h_1^2, \tag{4.114}$$

and thus use the equation of motion of varying h_0 and h_1 to rewrite them as Q

$$h_0 = \frac{r}{2A_1+2rA_1'-A_3r^2}[(rA_1'+2A_1)Q+rA_1Q'], \tag{4.115}$$

$$h_1 = -\frac{A_1}{A_4}\dot{Q}. \tag{4.116}$$

The auxiliary field Q encodes all of the information about the metric perturbations h_0 and h_1. Substituting eqs. (4.115) and (4.116) into eq. (4.114) and performing integration by parts, we can find a canonical Lagrangian of one degree of freedom

$$\mathcal{L}_{odd} = \frac{A_1^2}{A_4}\dot{Q}^2 - \frac{A_1^2r^2}{A_3r^2-2rA_1'-2A_1}(Q')^2 - \frac{\mu^2(r)}{2}Q^2, \tag{4.117}$$

where $\mu(r)$ is the mass of the auxiliary field Q

$$\mu^2(r) = \frac{2A_1r^2(r^2A_1'A_3'-r^2A_1''A_3+2A_1A_3+4A_1'^2+A_3^2r^2-2A_1A_1''+2rA_1A_3'-4rA_1'A_3)}{(2A_1+2rA_1'-A_3r^2)^2}. \tag{4.118}$$

By using this auxiliary field Lagrangian, we can thus get the condition for no-ghost and stable perturbations. The no-ghost condition can be easily seen from eq. (4.117)

$$\begin{aligned} & A_4 \geq 0 \\ \Rightarrow & \frac{j^2}{2} \geq e^{-2\beta}(2-e^{2\beta}+3r\alpha'-3r\beta'+r^2\alpha'^2-r^2\alpha'\beta'+r^2\alpha''+8\pi Gr^2e^{2\beta}V) \\ \Rightarrow & \frac{j^2}{2} \geq e^{-2\beta}[1+d(r)+2c(r)\lambda], \end{aligned} \tag{4.119}$$

where we have simplified the expression by using the dimensionless functions $c(r) \equiv 8\pi Gr^2e^{2\beta}V(r)$ and $d(r) \equiv e^{2\beta}-1$, and the condition of the tachyonic stability is

$$\begin{aligned} A_3r^2-2rA_1'-2A_1 \geq 0, \\ \Rightarrow r^2e^{2(\beta-\alpha)}A_4 \geq 0, \end{aligned} \tag{4.120}$$

which gives us the same condition as the no-ghost condition eq. (4.119). In the exponential potential with ansatz $M(r) \propto r$ case, the condition becomes

$$\left(\frac{j^2}{2}-1\right) \geq \frac{2c_*}{1+c_*}\lambda, \tag{4.121}$$

which can be easily satisfied once $c_* \ll 1$ for $L \neq 0, 1$.

4.5.4 Even Perturbation

In order to study the even type perturbation, we first parameterize the metric as

$$g_{\mu\nu} = \begin{bmatrix} -e^{2\alpha(r)}(1+H_0(t,r,\theta)) & H_1(t,r,\theta) & 0 & 0 \\ H_1(t,r,\theta) & e^{2\beta(r)}(1+H_2(t,r,\theta)) & K(t,r,\theta) & 0 \\ 0 & K(t,r,\theta) & r^2 & 0 \\ 0 & 0 & 0 & r^2\sin^2\theta \end{bmatrix}, \tag{4.122}$$

with the inverse metric

$$g^{\mu\nu} = \begin{bmatrix} -e^{-2\alpha(r)}(1-H_0(t,r,\theta)) & e^{-2(\alpha(r)+\beta(r))}H_1(t,r,\theta) & 0 & 0 \\ e^{-2(\alpha(r)+\beta(r))}H_1(t,r,\theta) & e^{-2\beta(r)}(1-H_2(t,r,\theta)) & -\frac{e^{-2\beta(r)}K(t,r,\theta)}{r^2} & 0 \\ 0 & -\frac{e^{-2\beta(r)}K(t,r,\theta)}{r^2} & \frac{1}{r^2} & 0 \\ 0 & 0 & 0 & \frac{1}{r^2\sin^2\theta} \end{bmatrix}, \tag{4.123}$$

where the functions are defined by

$$H_i(t,r,\theta) = \sum_{L,M} H_{i,LM}(t,r)Y_{LM}(\theta,\phi) \quad (i=0,1,2), \tag{4.124}$$

$$K(t,r,\theta) = \sum_{L,M} K_{LM}(t,r)\partial_\theta Y_{LM}(\theta,\phi), \tag{4.125}$$

$$\varphi(t,r,\theta) = \varphi_0(r) + \varphi_1(t,r,\theta) = \varphi_0(r) + \sum_{L,M} \varphi(t,r)Y_{LM}(\theta,\phi), \tag{4.126}$$

$$\lambda(t,r,\theta) = \lambda_0(r) + \lambda_1(t,r,\theta) = \lambda_0(r) + \sum_{L,M} \lambda(t,r)Y_{LM}(\theta,\phi). \tag{4.127}$$

The first order perturbation of the action vanishes due to the background equations of motion and the second order perturbation of the action is (with L, M, indices suppressed)

$$S_{even} = \frac{M_{pl}^2}{2}\sum_{L,M}\int dtdr\mathcal{L}_{even}, \tag{4.128}$$

where

$$\mathcal{L}_{even} = H_0'(a_1\varphi_1 + a_2H_2 + j^2a_3K) + a_4H_0^2 + H_0(a_5H_2 + j^2a_6H_2 + j^2a_7K) \\ + (b_1 + j^2b_2)H_1^2 + H_1(b_3\dot{H}_2 + j^2b_4\dot{K} + b_5\dot{\varphi}_1) + c_1H_2^2$$

$$
\begin{aligned}
&+H_2(j^2c_2K+c_3\lambda_1+c_4\varphi_1+c_5\varphi_1')+j^2(d_1\dot{K}^2+d_2K^2)\\
&+j^2K(d_3\varphi_1)+\lambda_1(e_1\varphi_1+e_2\varphi_1')+f_1\dot{\varphi}_1{}^2+f_2\varphi_1'^2+j^2f_3\varphi_1^2,
\end{aligned}
\tag{4.129}
$$

and the coefficients a_i, b_i, c_i, d_i, e_i, f_i are functions of r only; their expressions are given in Appendix D. Since there is no time derivation of fields H_0, H_1 and λ_1, we can vary these Lagrange multipliers to get the constraint equations

$$
H_0 = \frac{1}{2a_4}\left[(a_1\varphi_1+a_2H_2+j^2a_3K)'-(a_5H_2+a_6j^2H_2+j^2a_7K)\right], \tag{4.130}
$$

$$
H_1 = -\frac{1}{2(b_1+j^2b_2)}(b_3\dot{H}_2+j^2b_4\dot{K}+b_5\dot{\varphi}_1), \tag{4.131}
$$

$$
H_2 = -\frac{e_1}{c_3}\varphi_1-\frac{e_2}{c_3}\varphi_1'. \tag{4.132}
$$

After substituting H_0, H_1 into the Lagrangian, it can be reduced to

$$
\begin{aligned}
\mathcal{L}_{even} = &-\frac{1}{4a_4}[(a_1\varphi_1+a_2H_2+j^2a_3K)'-(a_5H_2+j^2a_6H_2+j^2a_7K)]^2\\
&-\frac{1}{4(b_1+j^2b_2)}(b_3\dot{H}_2+j^2b_4\dot{K}+b_5\dot{\varphi}_1)^2+c_1H_2^2\\
&+H_2(j^2c_2K+c_3\lambda_1+c_4\varphi_1+c_5\varphi_1')+j^2(d_1\dot{K}^2+d_2K^2)\\
&+j^2K(d_3\varphi_1)+\lambda_1(e_1\varphi_1+e_2\varphi_1')+f_1\dot{\varphi}_1{}^2+f_2\varphi_1'^2+j^2f_3\varphi_1^2.
\end{aligned}
\tag{4.133}
$$

After substituting H_2 into the Lagrangian, H_2 will be replaced by other fields while λ_1 is eliminated by the substitution. The Lagrangian can thus be expressed by $\varphi_1''^2$ and $\dot{\varphi}'^2$ and it can schematically be written as

$$
\begin{aligned}
\mathcal{L}_{even} = &\,g_1(r)\dot{\varphi}_1{}'^2+g_2(r)\dot{\varphi}_1{}^2+g_3(r)\dot{K}^2+g_4(r)\dot{\varphi}_1{}'\dot{K}+g_5(r)\dot{\varphi}_1\dot{K}\\
&+g_6(r)\varphi_1''^2+g_7(r)\varphi_1'^2+g_8(r)\varphi_1^2+g_9(r)\varphi_1''K'+g_{10}(r)\varphi_1'K'\\
&+g_{11}(r)\varphi_1K'+g_{12}(r)\varphi_1'K+g_{13}(r)\varphi_1K+g_{14}(r)K'^2+g_{15}K^2,
\end{aligned}
\tag{4.134}
$$

while the full expansion of the reduced Lagrangian is given in Appendix D.

Effectively, there are two degrees of freedom (φ, K) in eq. (4.134) and in general we can study the stability issue of the even type perturbation with the coefficients. However, as we can see in eq. (4.134), there is interaction between the two degrees of freedom while there is $\dot{\varphi}_1'^2$ in the Lagrangian which might be related to the stability issue. The best we can do is to check the stability issue at some limit. If we focus on the canonical kinetic term of φ_1 at low frequency limit, one can check from Appendix D

$$g_2 = -\frac{1}{[4c(r)\lambda_0(r) + j^2]} \frac{4r^2 e^{-(\alpha+\beta)}}{V(r)}, \tag{4.135}$$

and because of our background solution $M(r) \propto r$, $c(r) = c^* \ll 1$, g_2 is always negative for $L > 0$. We can thus conclude that there is always a ghost in the even-type perturbation because the negative sign of the kinetic term of φ_1.

4.6 Conclusion

In conclusion, we have found a solution for the $\lambda\varphi$-fluid model in the spherically symmetric static background with an exponential potential $V(\varphi) \propto e^{k\varphi}$, and show that there is no ghost or tachyon instability in odd-type perturbations. For the even-type perturbation, there is a technical problem associated with writing the Lagrangian in a canonical way with two degrees of freedom. We managed to write down the Lagrangian and show that the canonical kinetic term of one of the degrees of freedom always has a negative sign at low frequency limit and thus can be viewed as a ghost. Since we have only worked in linear theory, we do not know what the non-linear counterpart of this theory would be. It is totally possible that the ghost is removed at the non-linear level and the theory thus admits a stable spherically symmetric static background solution. We will leave this issue for future work.

Chapter 5

Conclusion

In this thesis, we have discussed aspects of higher derivative theories, constrained dynamics and the stability issue. In this chapter, we summarize the main results of this thesis.

In Chapter 2, we showed that all of the non-degenerate higher derivative theories suffer from the Ostrogradski instability in the context of the one-dimensional point particle where the position of the particle is a function only dependent on time. We proved that linear instability could be removed by the addition of constraints if the dimensionality of the phase space is reduced.

In Chapter 3, we further generalize our formalism to exorcise the Ostrogradski ghost in the higher derivative gravity models with quadratic curvature invariant R^2, $R^{\mu\nu}R_{\mu\nu}$ up to the quadratic level of metric fluctuation. We show that, at the linear level, the instability in the helicity-0, 1, and 2 sectors can be removed with suitable constraints if the dimensionality of the original theory's phase space is reduced. Using the formalism we introduced, the desirable renormalization property is retained at the price of breaking the Lorentz invariance, similar to Hořava-Lifshitz gravity.

In Chapter 4, we find a spherically symmetric static solution in the $\lambda\varphi$-fluid model with the potential $V(\varphi) \propto e^{k\varphi}$. The perturbation around the solution can be cast into odd and even types, dependent on its symmetry. We find that the odd type perturbation is a healthy degree of freedom while the even type perturbation contains one degree of freedom which always has a negative kinetic term and thus is a ghost.

References

[1] M. Ostrogradski. Memoires sur les equations differentielles relatives au probleme des isoperimetres. *Mem. Ac. St. Petersbourg*, VI: 385, 1850.

[2] John Archibald Wheeler and Richard Phillips Feynman. Interaction with the absorber as the mechanism of radiation. *Rev.Mod.Phys.*, 17:157–181, 1945.

[3] Alexei A. Starobinsky. A New Type of Isotropic Cosmological Models Without Singularity. *Phys.Lett.*, B91:99–102, 1980.

[4] Antonio De Felice and Shinji Tsujikawa. f(R) theories. *Living Rev.Rel.*, 13:3, 2010.

[5] Thomas P. Sotiriou and Valerio Faraoni. f(R) Theories Of Gravity. *Rev.Mod.Phys.*, 82:451–497, 2010.

[6] Paul A. M. Dirac. *Lectures on quantum mechanics*. Belfer graduate school of science, Yeshiva University, 1964.

[7] B. Bertotti, L. Iess, and P. Tortora. A test of general relativity using radio links with the Cassini spacecraft. *Nature*, 425:374, 2003.

[8] S.S. Shapiro, J.L. Davis, D.E. Lebach, and J.S. Gregory. Measurement of the Solar Gravitational Deflection of Radio Waves using Geodetic Very-Long-Baseline Interferometry Data, 1979-1999. *Phys.Rev.Lett.*, 92:121101, 2004.

[9] I. Ciufolini and E.C. Pavlis. A confirmation of the general relativistic prediction of the Lense-Thirring effect. *Nature*, 431:958–960, 2004.

[10] H. Weyl. Gravitation and electricity. *Sitzungsber.Preuss.Akad.Wiss.Berlin (Math.Phys.)*, 1918:465, 1918.

[11] H. Weyl. A New Extension of Relativity Theory. (In German). *Annalen Phys.*, 59:101–133, 1919.

[12] H. Weyl and J. Ehlers. Space, time, matter: Lectures on general relativity. (In German). 1993.

[13] A.D. Sakharov. Vacuum quantum fluctuations in curved space and the theory of gravitation. *Sov.Phys.Dokl.*, 12:1040–1041, 1968.

[14] K.S. Stelle. Renorization of Higher Derivative Quantum Gravity. *Phys.Rev.*, D16:953–969, 1977.

[15] K.S. Stelle. Classical Gravity with Higher Derivatives. *Gen.Rel.Grav.*, 9:353–371, 1978.

[16] P.A.R. Ade et al. Planck 2013 results. I. Overview of products and scientific results. *Gen.Rel.Grav.*, 2013.

[17] Timothy Clifton, Pedro G. Ferreira, Antonio Padilla, and Constantinos Skordis. Modified Gravity and Cosmology. *Phys.Rept.*, 513:1–189, 2012.

[18] P. Teyssandier and Ph. Tourrenc. The Cauchy problem for the R+R**2 theories of gravity without torsion. *J.Math.Phys.*, 24:2793, 1983.

[19] Sean M. Carroll, Vikram Duvvuri, Mark Trodden, and Michael S. Turner. Is cosmic speed - up due to new gravitational physics? *Phys.Rev.*, D70:043528, 2004.

[20] Ignacy Sawicki and Wayne Hu. Stability of Cosmological Solution in f(R) Models of Gravity. *Phys.Rev.*, D75:127502, 2007.

[21] Alexei A. Starobinsky. Disappearing cosmological constant in f(R) gravity. *JETP Lett.*, 86:157–163, 2007.

[22] Wayne Hu and Ignacy Sawicki. Models of f(R) Cosmic Acceleration that Evade Solar-System Tests. *Phys.Rev.*, D76:064004, 2007.

[23] Marc Henneaux and Claudio Teitelboim. *Quantization of Gauge Systems*. Princeton University Press, 1992.

[24] David Tong. *Lectures on Quantum Field Theory*. University of Cambridge, unpublished, 2006.

[25] Eugene A. Lim, Ignacy Sawicki, and Alexander Vikman. Dust of Dark Energy. *JCAP*, 1005:012, 2010.

[26] Renata Kallosh, Jin U. Kang, Andrei D. Linde, and Viatcheslav Mukhanov. The New ekpyrotic ghost. *JCAP*, 0804:018, 2008.

[27] D. A. Eliezer and R. P. Woodard. The Problem of Nonlocality in String Theory. *Nucl.Phys.*, B325:389, 1989.

[28] Robert C. Myers. HIGHER DERIVATIVE GRAVITY, SURFACE TERMS AND STRING THEORY. *Phys.Rev.*, D36:392, 1987.

[29] C. Charmousis, B. Gouteraux, and E. Kiritsis. Higher-derivative scalar-vector-tensor theories: black holes, Galileons, singularity cloaking and holography. *JHEP*, page 011, 2012.

[30] Jacob Moldenhauer, Mustapha Ishak, John Thompson, and Damien A. Easson. Supernova, baryon acoustic oscillations, and CMB surface distance constraints on f (G) higher order gravity models. *Phys.Rev.*, D81:063514, 2010.

[31] Hans A. Buchdahl. Non-linear Lagrangians and cosmological theory. *Mon.Not.Roy.Astron.Soc.*, 150:1, 1970.

[32] Gianluca Calcagni, Shinji Tsujikawa, and M Sami. Dark energy and cosmological solutions in second-order string gravity. *Class.Quant.Grav.*, 22:3977–4006, 2005.

[33] Ahmed Hindawi, Burt A. Ovrut, and Daniel Waldram. Consistent spin two coupling and quadratic gravitation. *Phys.Rev.*, D53:5583–5596, 1996.

[34] Takeshi Chiba. Generalized gravity and ghost. *JCAP*, 0503:008, 2005.

[35] Alvaro Nunez and Slava Solganik. Ghost constraints on modified gravity. *Phys.Lett.*, B608:189–193, 2005.

[36] N.H. Barth and S.M. Christensen. Quantizing Fourth Order Gravity Theories. 1. The Functional Integral. *Phys.Rev.*, D28:1876, 1983.

[37] Philip D. Mannheim. Making the Case for Conformal Gravity. *Found.Phys.*, 42:388–420, 2012.

[38] Andrew Strominger. Positive Energy Theorem for $R+R^2$ Gravity. *Phys.Rev.*, D30:2257, 1984.

[39] Guido Magnano and Leszek M. Sokolowski. On physical equivalence between non-linear gravity theories and a general relativistic selfgravitating scalar field. *Phys.Rev.*, D50:5039–5059, 1994.

[40] Tirthabir Biswas, Anupam Mazumdar, and Warren Siegel. Bouncing universes in string-inspired gravity. *JCAP*, 0603:009, 2006.

[41] Richard P. Woodard. Avoiding dark energy with 1/r modifications of gravity. *Lect.Notes Phys.*, 720:403–433, 2007.

[42] D. Lovelock. The Einstein tensor and its generalizations. *J.Math.Phys.*, 12:498–501, 1971.

[43] Alberto Nicolis, Riccardo Rattazzi, and Enrico Trincherini. The Galileon as a local modification of gravity. *Phys.Rev.*, D79:064036, 2009.

[44] C. Deffayet, Gilles Esposito-Farese, and A. Vikman. Covariant Galileon. *Phys.Rev.*, D79:084003, 2009.

[45] Cedric Deffayet, Oriol Pujolas, Ignacy Sawicki, and Alexander Vikman. Imperfect Dark Energy from Kinetic Gravity Braiding. *JCAP*, 1010:026, 2010.

[46] C. Deffayet, S. Deser, and G. Esposito-Farese. Generalized Galileons: All scalar models whose curved background extensions maintain second-order field equations and stress-tenxrs. *Phys.Rev.*, D80:064015, 2009.

[47] Alberto Nicolis, Riccardo Rattazzi, and Enrico Trincherini. Energy's and amplitudes' positivity. *JHEP*, 1005:095, 2010.

[48] Jonathan Z. Simon. Higher derivative Lagrangians, nonlocality, problems and solutions. *Phys.Rev.*, D41:3720, 1990.

[49] Brian Whitt. Fourth Order Gravity as General Relativity Plus Matter. *Phys.Lett.*, B145:176, 1984.

[50] Ahmed Hindawi, Burt A. Ovrut, and Daniel Waldram. Nontrivial vacua in higher derivative gravitation. *Phys.Rev.*, D53:5597–5608, 1996.

[51] Sean M. Carroll, Antonio De Felice, Vikram Duvvuri, Damien A. Easson, Mark Trodden, et al. The Cosmology of generalized modified gravity models. *Phys.Rev.*, D71:063513, 2005.

[52] Ignacio Navarro and Karel Van Acoleyen. Consistent long distance modification of gravity from inverse powers of the curvature. *JCAP*, 0603:008, 2006.

[53] Antonio De Felice, Mark Hindmarsh, and Mark Trodden. Ghosts, Instabilities, and Superluminal Propagation in Modified Gravity Models. *JCAP*, 0608:005, 2006.

[54] Juan Maldacena. Einstein Gravity from Conformal Gravity. 2011.

[55] H. Lu, Y. Pang, C. N. Pope, and Justin F. Vazquez-Poritz. AdS and Lifshitz Black Holes in Conformal and Einstein-Weyl Gravities. *Phys.Rev.*, D86:044011, 2012.

[56] Rodrigo Olea. Mass, angular momentum and thermodynamics in four-dimensional Kerr-AdS black holes. *JHEP*, 0506:023, 2005.

[57] X. Jaen, J. Llosa, and A. Molina. A Reduction of order two for infinite order lagrangians. *Phys.Rev.*, D34:2302, 1986.

[58] F.J. de Urries and J. Julve. Ostrogradski formalism for higher derivative scalar field theories. *J.Phys.A*, A31:6949–6964, 1998.

[59] I. Ya. Aref'eva and I.V. Volovich. On the null energy condition and cosmology. *Theor.Math.Phys.*, 155:503–511, 2008.

[60] Kurt Hinterbichler. Theoretical Aspects of Massive Gravity. *Rev.Mod.Phys.*, 84:671–710, 2012.

[61] Claudia de Rham. Galileons in the Sky. *Comptes Rendus Physique*, 13:666–681, 2012.

[62] V.I. Tkach. Ghost-Free Higher Derivative Quantum Gravity, the Hierarchy and the Cosmological Constant Problems. *Mod.Phys.Lett.*, A24:1193–1197, 2009.

[63] Neil Barnaby, Tirthabir Biswas, and James M. Cline. p-adic Inflation. *JHEP*, 0704:056, 2007.

[64] Tirthabir Biswas, Erik Gerwick, Tomi Koivisto, and Anupam Mazumdar. Towards singularity and ghost free theories of gravity. *Phys.Rev.Lett.*, 108:031101, 2012.

[65] Minjoon Park and Lorenzo Sorbo. Massive Gravity from Higher Derivative Gravity with Boundary Conditions. *JHEP*, 1301:043, 2013.

[66] Tai-jun Chen, Matteo Fasiello, Eugene A. Lim, and Andrew J. Tolley. Higher derivative theories with constraints: Exorcising Ostrogradski's Ghost. *JCAP*, 1302:042, 2013.

[67] Matej Pavsic. Stable Self-Interacting Pais-Uhlenbeck Oscillator. *Mod.Phys.Lett.*, A28:1350165, 2013.

[68] Carl M. Bender and Philip D. Mannheim. No-ghost theorem for the fourth-order derivative Pais-Uhlenbeck oscillator model. *Phys.Rev.Lett.*, 100:110402, 2008.

[69] Philip D. Mannheim and Aharon Davidson. Dirac quantization of the Pais-Uhlenbeck fourth order oscillator. *Phys.Rev.*, A71:042110, 2005.

[70] Ruth Durrer and Roy Maartens. Dark Energy and Modified Gravity. pages 48 – 91, 2008.

[71] A. Pais and G.E. Uhlenbeck. On Field theories with nonlocalized action. *Phys.Rev.*, 79:145–165, 1950.

[72] D. M. Gitman and I. V. Tyutin. *Quantization of fields with constraints*. Springer-Verlag, 1990.

[73] Lev V. Prokhorov and Sergei V. Shabanov. *Hamiltonian mechanics of gauge systems*. Cambridge University Press, 2011.

[74] Jan Govaerts. *Hamiltonian quantisation and constrained dynamics*. Leuven University Press, 1989.

[75] Petr Horava. Membranes at Quantum Criticality. *JHEP*, 0903:020, 2009.

[76] Petr Horava. Quantum Gravity at a Lifshitz Point. *Phys.Rev.*, D79:084008, 2009.

[77] Petr Horava. Spectral Dimension of the Universe in Quantum Gravity at a Lifshitz Point. *Phys.Rev.Lett.*, 102:161301, 2009.

[78] D. Blas, O. Pujolas, and S. Sibiryakov. Consistent Extension of Horava Gravity. *Phys.Rev.Lett.*, 104:181302, 2010.

[79] Shinji Mukohyama. Horava-Lifshitz Cosmology: A Review. *Class.Quant.Grav.*, 27:223101, 2010.

[80] Thomas P. Sotiriou. Horava-Lifshitz gravity: a status report. *J.Phys.Conf.Ser.*, 283:012034, 2011.

[81] Ibrahim Gullu, Tahsin Cagri Sisman, and Bayram Tekin. Unitarity analysis of general Born-Infeld gravity theories. *Phys.Rev.*, D82:124023, 2010.

[82] Stanley Deser and Bayram Tekin. Energy in generic higher curvature gravity theories. *Phys.Rev.*, D67:084009, 2003.

[83] Yong-Seon Song, Wayne Hu, and Ignacy Sawicki. The Large Scale Structure of f(R) Gravity. *Phys.Rev.*, D75:044004, 2007.

[84] D.G. Boulware and Stanley Deser. Can gravitation have a finite range? *Phys.Rev.*, D6:3368–3382, 1972.

[85] C. Armendariz-Picon and Eugene A. Lim. Haloes of k-essence. *JCAP*, 0508:007, 2005.

[86] Garrett L. Goon, Kurt Hinterbichler, and Mark Trodden. Stability and superluminality of spherical DBI galileon solutions. *Phys.Rev.*, D83:085015, 2011.

[87] Vladimir Dzhunushaliev, Vladimir Folomeev, and Douglas Singleton. Chameleon stars. *Phys.Rev.*, D84:084025, 2011.

[88] Tullio Regge and John A. Wheeler. Stability of a Schwarzschild singularity. *Phys.Rev.*, 108:1063–1069, 1957.

[89] K.S. Thorne. Multipole Expansions of Gravitational Radiation. *Rev.Mod.Phys.*, 52:299–339, 1980.

[90] Jon Mathews. Gravitational Multipole Radiation. *j-J-SIAM*, 10(4):768–780, December 1962.

[91] Antonio De Felice, Teruaki Suyama, and Takahiro Tanaka. Stability of Schwarzschild-like solutions in f(R,G) gravity models. *Phys.Rev.*, D83:104035, 2011.

[92] Frank J. Zerilli. Effective potential for even parity Regge-Wheeler gravitational perturbation equations. *Phys.Rev.Lett.*, 24:737–738, 1970.

Appendix A

The Difference between Ghost and Tachyonic Instability

In this appendix, we aim to demonstrate the difference between ghost and tachyonic instability. In the literature, ghosts usually appear due to a “wrong sign" in the kinetic term in the Lagrangian, i.e. the coefficient of $\dot{\phi}^2$, and the Hamiltonian is thus unbounded from below (but bounded above) and will generate a negative energy spectrum or negative norm of state. Here, we want to emphasize that the ghost on its own is not a problem, since it shares the same equation of motion with the normal fields and one can flip the sign of the Lagrangian to make the Hamiltonian bounded from below, as with the normal fields. The ghost becomes severe when it interacts with the normal fields, since the interaction with the normal modes with positive energy spectra discriminate the ghost from the normal fields and the interaction between the normal modes and the ghost with the negative energy spectrum[1] guarantees that the Hamiltonian of the system would be unbounded from below in part of the phase space.

On the other hand, tachyonic instability is commonly viewed as the existence of exponential growing modes, and the states thus cannot be normalized as normal modes. Here, we want to discuss the difference between the ghost and tachyonic instability. It will be possible to discriminate between these two types of instability after reading this appendix. We will use free field theory as an example to discuss these issues.

[1]Here we choose the one particle state whose norm are positive definite but whose energy eigenvalue is negative.

c_1	c_2	mode	problem
+	+	Normal mode	Well defined
+	-	Tachyonic mode	Imaginary energy spectrum with vanishing norm/ Exponential growing mode
-	+	Exponential ghost mode	Imaginary energy spectrum with vanishing norm/ Exponential growing mode
-	-	Oscillating ghost mode	Negative norm/spectrum

Table A.1

A.1 General calculation

We start by considering the Lagrangian of free scalar field theory with arbitrary coefficients

$$\mathcal{L} = -\frac{c_1 \partial_\mu \phi \partial^\mu \phi}{2} - \frac{c_2 m^2 \phi^2}{2}, \tag{A.1}$$

where $\phi = \phi(\mathbf{x}, t)$ and c_1, c_2 are ± 1. The different choices of c_1, c_2 will correspond to the normal, tachyonic, and ghost modes, as shown in Table A.1.

The Euler-Lagrange equation of the Lagrangian is

$$\ddot{\phi} - \nabla^2 \phi = -\frac{c_2}{c_1} m^2 \phi, \tag{A.2}$$

which can be solved by using the Fourier transformation $\tilde{\phi}(\mathbf{p}, t)$

$$\tilde{\phi}(\mathbf{p}, t) = D_1(\mathbf{p}) e^{-iw_p t} + D_2(\mathbf{p}) e^{iw_p t}, \tag{A.3}$$

where $w_p = \sqrt{|p|^2 + \frac{c_2}{c_1} m^2}$. The conjugate momentum and Hamiltonian density are

$$\pi(\mathbf{x}, t) \equiv c_1 \dot{\phi}, \tag{A.4}$$

$$\mathcal{H} = \frac{\pi^2}{2c_1} + \frac{c_1 (\nabla \phi)^2}{2} + \frac{c_2 m^2 \phi^2}{2}. \tag{A.5}$$

Using the ladder operators, the canonical field and its conjugate momentum can be written as

$$\phi(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^{3/2}} \frac{1}{\sqrt{2w_p}} [D_1(\mathbf{p}) + D_2(-\mathbf{p})] e^{i\mathbf{p}\cdot\mathbf{x}}, \tag{A.6}$$

$$\pi(\mathbf{x}) = \int \frac{d^3\mathbf{p}}{(2\pi)^{3/2}} (-ic_1) \sqrt{\frac{w_p}{2}} [D_1(\mathbf{p}) - D_2(-\mathbf{p})] e^{i\mathbf{p}\cdot\mathbf{x}}. \tag{A.7}$$

The non-vanishing equal time commutation relations can be calculated

$$[\phi(\mathbf{x}), \pi(\mathbf{y})] = i\delta(\mathbf{x}-\mathbf{y}), \tag{A.8}$$

$$[D_1(\mathbf{p}), D_2(\mathbf{q})] = \frac{1}{c_1}\delta(\mathbf{p}-\mathbf{q}). \tag{A.9}$$

and the Hamiltonian is

$$H = \int d^3\mathbf{p}\; c_1\sqrt{w_p^2}D_2(\mathbf{p})D_1(\mathbf{p}) + \frac{\sqrt{w_p^2}}{2}\delta^3(0). \tag{A.10}$$

A.2 Spectrum and Norm of the Theory

The difference between the normal, ghost and tachyonic modes arises due to their different energy spectra and norms of states. Here we want to discuss the problems of the unstable modes by going through the full quantization process.

The choice $c_1 = c_2 = 1$ corresponds to the normal free scalar field. In this theory, the w_p is always real and $D_2(\mathbf{p}) = D_1^\dagger(\mathbf{p})$ from the reality requirement of the scalar field $\phi(\mathbf{x})$. The spectrum is bounded below if we define the vacuum state $|0\rangle$ as $D_1(\mathbf{p})|0\rangle = 0$ where, after we shift the vacuum energy, the energy $E_0 = 0$. All of the many-particle states $|p_1, p_2, \dots\rangle$ are defined as creation operators acting on the vacuum state $D_2(\mathbf{p_1})D_2(\mathbf{p_2})\dots|0\rangle$ with the total energy $E = w_{p_1} + w_{p_2} + \cdots$. The norm of many-particle state is positive definite because of $D_2(\mathbf{p}) = D_1^\dagger(\mathbf{p})$ and the commutation relation eq. (A.9).

By generalizing the quantization procedure to the theory with other choices of c_1, c_2, we can discuss the unstable modes.

A.2.1 The Oscillating Ghost

The oscillating ghost corresponds to $c_1 = c_2 = -1$, where w_p^2 is positive definite. The commutation relations between the Hamiltonian and ladder operators are:

$$[H, D_1(\mathbf{p})] = -w_p D_1(\mathbf{p}), \tag{A.11}$$

$$[H, D_2(\mathbf{p})] = w_p D_2(\mathbf{p}). \tag{A.12}$$

If we want our energy to be bounded from below, the vacuum state should be defined as $D_1(\mathbf{p})\,|0\rangle = 0$, and the one particle state is $D_2(\mathbf{p})\,|0\rangle$, with $E = w_p$. In this case, the norm of the one particle state is negative definite since the commutation relation eq. (A.9) has a sign change on the R.H.S.. On the other hand, we can choose to have a positive definite norm of the one particle state if we define the vacuum state as $D_2(\mathbf{p})\,|0\rangle = 0$, and the one particle state is $D_1(\mathbf{p})\,|0\rangle$, with $E = -w_p$. This oscillating ghost is the common ghost in the literature, which either has a negative norm of states or the energy of the system is unbounded from below.

A.2.2 Tachyonic Mode

For $c_1 = 1$, $c_2 = -1$, the frequency $w_p^2 = |p|^2 - m^2$ and there are two branches of this class of model. If $w_p^2 > 0$, the modes are similar to the normal modes and the analysis is exactly the same as the usual free scalar field theory. If $w_p^2 < 0$, the reality requirement of $\phi(\mathbf{x})$ and $\pi(\mathbf{x})$ gives us $D_1^\dagger(\mathbf{p}) = -iD_1(-\mathbf{p})$ and $D_2^\dagger(\mathbf{p}) = -iD_2(-\mathbf{p})$. The commutation relation between the Hamiltonian and the ladder operators becomes

$$[H, D_1(\mathbf{p})] = -i|w_p|D_1(\mathbf{p}), \tag{A.13}$$

$$[H, D_2(\mathbf{p})] = i|w_p|D_2(\mathbf{p}). \tag{A.14}$$

The time-dependent field and its conjugate momentum in Heisenberg's picture are

$$\phi(\mathbf{x},t) = \int \frac{d^3\mathbf{p}}{(2\pi)^{3/2}} \frac{1}{\sqrt{2w_p}} [D_1(\mathbf{p})e^{|w_p|t} + D_2(-\mathbf{p})e^{-|w_p|t}]e^{i\mathbf{p}\cdot\mathbf{x}}, \tag{A.15}$$

$$\pi(\mathbf{x},t) = \int \frac{d^3\mathbf{p}}{(2\pi)^{3/2}} (-i)\sqrt{\frac{w_p}{2}} [D_1(\mathbf{p})e^{|w_p|t} - D_2(-\mathbf{p})e^{-|w_p|t}]e^{i\mathbf{p}\cdot\mathbf{x}}. \tag{A.16}$$

The common misconception is that the exponential modes always exist in tachyonic field theory, but they can, in fact, be removed by choosing an appropriate vacuum state. If we do not want these modes which are proportional to $e^{|w_p|t}$, we can choose our vacuum state such that $D_1(\mathbf{p})\,|0\rangle = 0$. The one particle state in the theory is then $D_2(\mathbf{p})\,|0\rangle$, with energy $E = i|w_p|$ and the norm of the state $D_2(\mathbf{p})\,|0\rangle$ is

$$\langle 0|\, D_2^\dagger(\mathbf{p}) D_2(\mathbf{p})\,|0\rangle = -i\,\langle 0|\, D_2(-\mathbf{p}) D_2(\mathbf{p})\,|0\rangle = 0. \tag{A.17}$$

We can see that this theory has an imaginary energy spectrum and vanishing norm of states, which is problematic, but this has nothing to do with the exponential growing wave function of the theory.

A.2.3 Exponential Ghost

The exponential ghost theory is similar to the tachyonic theory. For $c_1 = -1$, $c_2 = 1$, the frequency $w_p^2 = |p|^2 - m^2$ and the two branches of this class of model correspond to $w_p^2 > 0$, which is an oscillating ghost and the analysis is exactly the same as that shown in A.2.1. If $w_p^2 < 0$, which is similar to the tachyonic mode, with the commutation relation between the Hamiltonian and the ladder operators becoming

$$[H, D_1(\mathbf{p})] = -i|w_p|D_1(\mathbf{p}), \tag{A.18}$$

$$[H, D_2(\mathbf{p})] = i|w_p|D_2(\mathbf{p}). \tag{A.19}$$

The time-dependent field and its conjugate momentum in Heisenberg's picture are

$$\phi(\mathbf{x},t) = \int \frac{d^3\mathbf{p}}{(2\pi)^{3/2}} \frac{1}{\sqrt{2w_p}} [D_1(\mathbf{p})e^{|w_p|t} + D_2(-\mathbf{p})e^{-|w_p|t}] e^{i\mathbf{p}\cdot\mathbf{x}}, \tag{A.20}$$

$$\pi(\mathbf{x},t) = \int \frac{d^3\mathbf{p}}{(2\pi)^{3/2}} i\sqrt{\frac{w_p}{2}} [D_1(\mathbf{p})e^{|w_p|t} - D_2(-\mathbf{p})e^{-|w_p|t}] e^{i\mathbf{p}\cdot\mathbf{x}}, \tag{A.21}$$

while the reality requirement of $\phi(\mathbf{x},t)$ gives us $D_1^\dagger(\mathbf{p}) = -iD_1(-\mathbf{p})$ and $D_2^\dagger(\mathbf{p}) = -iD_2(-\mathbf{p})$. If we do not want the exponential growing mode, which is proportional to $e^{|w_p|t}$, we can define our vacuum state as $D_1(\mathbf{p})|0\rangle = 0$. The one particle state in the theory is defined as $D_2(\mathbf{p})|0\rangle$ with energy $E = i|w_p|$. We thus have an imaginary spectrum even when there is no exponential growing wave function in the theory. The norm of the state $D_2(\mathbf{p})|0\rangle$ is

$$\langle 0| D_2^\dagger(\mathbf{p}) D_2(\mathbf{p}) |0\rangle = -i \langle 0| D_2(-\mathbf{p}) D_2(\mathbf{p}) |0\rangle = 0, \tag{A.22}$$

and we thus have a vanishing norm of states for the exponential ghost model. The difference between the tachyonic and the exponential ghost modes are the sign of the Hamiltonian (c_1 in eq. (A.10)).

Appendix B

Quantization of higher derivative theory

In this appendix, we use a higher derivative scalar field theory to demonstrate the subtleties of the quantum higher derivative theory. We begin with

$$S = \int d^4x \left[\frac{1}{2}\phi\Box\phi + \frac{\sigma}{2M^2}(\Box\phi)^2 - \frac{m^2\phi^2}{2} \right], \tag{B.1}$$

where $\Box$ is the d'Alembert operator, M, m are constants with mass dimension 1 and $\sigma = \pm 1$. The Euler-Lagrange equation is

$$\Box\phi + \frac{\sigma}{M^2}\Box\Box\phi - m^2\phi = 0, \tag{B.2}$$

by Fourier transform, the solution is a set of harmonic oscillators with frequency

$$w_p^2 - p^2 = \frac{-M^2 \pm M^2\sqrt{1 + \frac{4\sigma m^2}{M^2}}}{2\sigma}. \tag{B.3}$$

We can see that there are two frequencies correspond to each p, which means that this theory has two degrees of freedom. To simplify the calculation, we can take $m = 0$ and one of the degrees of freedom thus becomes massless. We can also take $\sigma = -1$, which makes the other degree of freedom ghostlike ($\sigma = 1$ would instead make it an exponential ghost). We can thus denote the frequencies by

$$\begin{aligned} w_p^2 &= p^2 \\ v_p^2 &= p^2 + M^2. \end{aligned} \tag{B.4}$$

To describe the theory in the Hamiltonian picture, we need first to define the canonical variables

$$q_1 = \phi \Leftrightarrow p_1 = \frac{\delta S}{\delta \dot{\phi}}$$
$$q_2 = \dot{\phi} \Leftrightarrow p_2 = -\frac{\ddot{\phi}}{M^2}. \tag{B.5}$$

Since there is no constraint in the theory, the Hamiltonian is

$$H = \int d^3x \left[p_1 q_2 - \frac{M^2 p_2^2}{2} + q_1(-\frac{1}{2}\nabla^2 + \frac{1}{2M^2}\nabla^2\nabla^2)q_1 + q_2(-\frac{1}{2} + \frac{1}{M^2}\nabla^2)q_2 \right]. \tag{B.6}$$

To quantize the theory, we write q_1, q_2, p_1, p_2 as linear combinations of the two pairs of creation and annihilation operators $(a_p^\dagger, a_p)$, $(b_p^\dagger, b_p)$

$$q_1 = \int \frac{d^3p}{(2\pi)^3} \left[\frac{1}{\sqrt{2w_p}}(a_p e^{ip\cdot x} + a_p^\dagger e^{-ip\cdot x}) + \frac{i}{\sqrt{2v_p}}(b_p^\dagger e^{ip\cdot x} - b_p e^{-ip\cdot x}) \right]$$
$$q_2 = \int \frac{d^3p}{(2\pi)^3} \left[(-i)\sqrt{\frac{w_p}{2}}(a_p e^{ip\cdot x} - a_p^\dagger e^{-ip\cdot x}) + \sqrt{\frac{v_p}{2}}(b_p^\dagger e^{ip\cdot x} + b_p e^{-ip\cdot x}) \right]$$
$$p_1 = \int \frac{d^3p}{(2\pi)^3} \left[\frac{(-i)}{M^2}\sqrt{\frac{w_p v_p^4}{2}}(a_p e^{ip\cdot x} - a_p^\dagger e^{-ip\cdot x}) + \frac{1}{M^2}\sqrt{\frac{v_p w_p^4}{2}}(b_p^\dagger e^{ip\cdot x} + b_p e^{-ip\cdot x}) \right]$$
$$p_2 = \int \frac{d^3p}{(2\pi)^3} \left[\frac{1}{M^2}\sqrt{\frac{w_p^3}{2}}(a_p e^{ip\cdot x} + a_p^\dagger e^{-ip\cdot x}) + \frac{i}{M^2}\sqrt{\frac{v_p^3}{2}}(b_p^\dagger e^{ip\cdot x} - b_p e^{-ip\cdot x}) \right]. \tag{B.7}$$

The coefficients of creation and annihilation operators are chosen in such a way that the commutators

$$[q_1(x), p_1(y)] = [q_2(x), p_2(y)] = i\delta^{(3)}(x-y) \tag{B.8}$$

are consistent with the usual commutator relation

$$[a_p, a_k^\dagger] = [b_p, b_k^\dagger] = (2\pi)^3 \delta^{(3)}(p-k), \tag{B.9}$$

with all other possible commutators vanishing. It should also be noted that each canonical variable is a combination of two degrees of freedom, as the two degrees of freedom vibrate at different frequencies, i.e. in the Heisenberg picture, $a_p \to a_p e^{-iw_p t}$, $b_p \to b_p e^{iv_p t}$. Equipped with all of this information, we can substitute eq. (B.7) into the Hamiltonian (B.6). After some work, we find the Hamiltonian to be

$$H = \int \frac{d^3p}{(2\pi)^3}\, w_p \left[a_p^\dagger a_p + \frac{1}{2}(2\pi)^3\delta^{(3)}(0)\right] - v_p \left[b_p^\dagger b_p + \frac{1}{2}(2\pi)^3\delta^{(3)}(0)\right], \tag{B.10}$$

where $w_p = \sqrt{p^2}$ and $v_p = \sqrt{p^2 + M^2}$. One can see that, while $a_p^\dagger$ creates a massless particle with positive energy, $b_p^\dagger$ creates a massive particle with negative energy, and thus the theory has a massive ghost. One can always redefine $b_p \equiv b_p^\dagger$, and the new $b_p^\dagger$ will create a massive particle with positive energy, but saddled with a negative norm.

Appendix C

Equivalence of Ostrogradski's formalism and the auxiliary field method

In this appendix, we will use eq. (B.1) with $\sigma = -1$ and $m = 0$, as a toy-model to show the equivalence between Ostrogradski's formalism of higher derivative theory and the auxiliary field method used in the literature (e.g. [14]). In the auxiliary field method, the action with one higher derivative scalar field

$$S = \int d^4x \, \left[\frac{1}{2}\phi\Box\phi - \frac{1}{2M^2}(\Box\phi)^2\right], \tag{C.1}$$

is equivalent to the action with two standard scalar fields

$$S = \int d^4x \, \left\{\frac{1}{2}\phi\Box\phi - \frac{1}{2M^2}(\Box\phi)^2 + \frac{1}{2M^2}\left[\Box\phi + \frac{M^2(\lambda-\phi)}{2}\right]^2\right\}. \tag{C.2}$$

The action can be reduced to

$$S = \int d^4x \, \left[\frac{1}{2}\lambda\Box\phi + \frac{M^2}{8}(\lambda-\phi)^2\right], \tag{C.3}$$

and diagonalized as

$$S = \int d^4x \, \left(\frac{1}{2}\Phi\Box\Phi - \frac{1}{2}\Psi\Box\Psi + \frac{M^2}{2}\Psi^2\right), \tag{C.4}$$

where $\phi = \Phi - \Psi$ and $\lambda = \Phi + \Psi$. The action (C.4) describes a healthy massless scalar field with a massive ghostlike scalar field. The conjugate momenta and Hamiltonian of the system can easily be written as

$$p_\Phi = \dot{\Phi}$$
$$p_\Psi = -\dot{\Psi}$$
$$H = \int d^3x \left\{ \frac{p_\Phi^2}{2} + \frac{(\nabla\Phi)^2}{2} - \left[\frac{p_\Psi^2}{2} + \frac{(\nabla\Psi)^2}{2} + \frac{M^2}{2}\Psi^2 \right] \right\}. \tag{C.5}$$

On the other hand, Ostrogradski's formalism leads to Hamiltonian (B.6), which is linearly dependent on p_1

$$\begin{aligned} H = \int d^3x \, \Bigg[& p_1 q_2 + q_1 \left(-\frac{1}{2}\nabla^2 + \frac{1}{2M^2}\nabla^2\nabla^2 \right) q_1 \\ & - \frac{M^2 p_2^2}{2} + q_2 \left(-\frac{1}{2} + \frac{1}{M^2}\nabla^2 \right) q_2 \Bigg], \end{aligned} \tag{C.6}$$

which can be diagonalized by the following canonical transformation

$$\begin{aligned} q_1 &= \Phi + \Psi \\ q_2 &= p_\Phi - p_\Psi \\ p_1 &= p_\Phi - \frac{\nabla^2}{M^2}(p_\Phi - p_\Psi) \\ p_2 &= \Psi - \frac{\nabla^2}{M^2}(\Phi + \Psi). \end{aligned} \tag{C.7}$$

The final Hamiltonian becomes

$$H = \int d^3x \left\{ \frac{p_\Phi^2}{2} + \frac{(\nabla\Phi)^2}{2} - \left[\frac{p_\Psi^2}{2} + \frac{(\nabla\Psi)^2}{2} + \frac{M^2}{2}\Psi^2 \right] \right\}, \tag{C.8}$$

which is the same as eq. (C.5). Hence we have shown that Ostrogradski's formalism is equivalent to the auxiliary field method up to some canonical transformation.

Appendix D

Full expression of the second order action of even-type perturbation

The coefficients of the second order action of even-type perturbation eq. (4.133) and the useful coefficients in eq. (4.134) are defined as follows. Here "$\approx$" means that we have used the ansatz such that the background solution satisfies $M(r) \propto r$ and thus $c(r) = d(r) = c_*$. We also present the full expansion of the Lagrangian at the end of this appendix.

$$
\begin{aligned}
a_1 &= \frac{2r^2 V(r) e^{(\alpha+\beta)}}{M_{pl}^2 \varphi_0'(r)} [1 + \lambda_0(r)] = \frac{r^2 e^{(\alpha-\beta)}}{M_{pl}^2} [1 + \lambda_0(r)] \varphi_0'(r) \\
a_2 &= -\, r e^{(\alpha-\beta)} \\
a_3 &= -\, e^{(\alpha-\beta)} \\
a_4 &= \frac{e^{(\alpha-\beta)}}{4} \left(8\pi G r^2 e^{2\beta} V(r) + 1 - e^{2\beta} + 8r\alpha' + 4r^2\alpha'^2 - 4r^2\alpha'\beta' + 4r^2\alpha'' - 2r\beta' \right) \\
&\approx e^{(\alpha-\beta)} c_* \lambda_0(r) \\
a_5 &= 0 \\
a_6 &= \frac{e^{(\alpha+\beta)}}{2} \\
a_7 &= e^{(\alpha-\beta)} \left(\frac{1}{r} - \alpha' \right) \\
b_1 &= e^{-(\alpha+3\beta)} \left[-2r^2\alpha'' - 4r\alpha' - 2r^2\alpha'^2 + 2r\beta' + 2r^2\alpha'\beta' - 1 + e^{2\beta} - 8\pi G r^2 e^{2\beta} V(r) \right] \\
&\approx \, -2e^{-(\alpha+3\beta)} c_* \lambda_0(r) \\
b_2 &= \frac{e^{-(\alpha+\beta)}}{2}
\end{aligned}
$$

$$
\begin{aligned}
b_3 =& -2re^{-(\alpha+\beta)} \\
b_4 =& -e^{-(\alpha+\beta)} \\
b_5 =& \frac{2r^2e^{-(\alpha+\beta)}}{M_{pl}^2}\left[1+\lambda_0(r)\right]\varphi_0' \\
c_1 =& \frac{e^{(\alpha-\beta)}}{4}\left(3-e^{2\beta}+4r^2\alpha'^2-8r\beta'+6r\alpha'-4r^2\alpha'\beta'+4r^2\alpha''\right)+\frac{r^2V(r)e^{(\alpha+\beta)}}{4M_{Pl}^2}[5+4\lambda_0(r)] \\
=& \frac{e^{(\alpha-\beta)}}{2}\left[1+d(r)+3c(r)\lambda_0(r)\right] \\
\approx& \frac{e^{(\alpha-\beta)}}{2}\left\{1+c_*[1+3\lambda_0(r)]\right\} \\
c_2 =& e^{(\alpha-\beta)}\left(\alpha'+\frac{1}{r}\right) \\
c_3 =& \frac{2r^2V(r)e^{(\alpha+\beta)}}{M_{pl}^2} \\
c_4 =& \frac{r^2V_\varphi(r)e^{(\alpha+\beta)}}{M_{pl}^2}\lambda_0(r) \\
c_5 =& \frac{r^2e^{(\alpha-\beta)}}{M_{pl}^2}[1+\lambda_0(r)]\varphi_0'(r) \\
d_1 =& \frac{e^{-(\alpha+\beta)}}{2} \\
d_2 =& \frac{e^{(\alpha-3\beta)}}{r^2}\left[2-e^{2\beta}+r^2\alpha'^2-3r\beta'+3r\alpha'-r^2\alpha'\beta'+r^2\alpha''+8\pi Ge^{2\beta}r^2V(r)\right] \\
=& \frac{e^{(\alpha-3\beta)}}{r^2}\left[1+d(r)+2c(r)\lambda_0(r)\right] \\
d_3 =& -\frac{2e^{(\alpha-\beta)}}{M_{pl}^2}\left[1+\lambda_0(r)\right]\varphi_0'(r) \\
e_1 =& \frac{2r^2V_\varphi(r)e^{(\alpha+\beta)}}{M_{pl}^2} \\
e_2 =& -\frac{2r^2e^{(\alpha-\beta)}}{M_{pl}^2}\varphi_0'(r) \\
f_1 =& \frac{r^2e^{-(\alpha-\beta)}}{M_{pl}^2}\left[1+\lambda_0(r)\right] \\
f_2 =& -\frac{r^2e^{(\alpha-\beta)}}{M_{pl}^2}\left[1+\lambda_0(r)\right]
\end{aligned}
$$

$$f_3 = -\frac{e^{(\alpha+\beta)}}{M_{pl}^2}\,[1+\lambda_0(r)]$$
$$g_2 = -\frac{4r^2 e^{-(\alpha+\beta)}}{[4c(r)\lambda_0(r)+j^2]\,V(r)}$$
$$g_3 = -\frac{4c(r)\lambda_0(r)\,j^2 e^{-(\alpha+\beta)}}{[-4c(r)\lambda_0(r)+j^2]}$$
$$g_6 = -\frac{r^2 e^{(\alpha-3\beta)}}{c(r)\lambda_0(r)V(r)}$$

$$\begin{aligned}
\mathcal{L}_{even} = & -\frac{1}{4(b_1+j^2b_2)}\left(\frac{b_3e_2}{c_3}\right)^2\dot{\varphi}_1{}'^2 - \frac{1}{4(b_1+j^2b_2)}[j^4b_4^2-4j^2(b_1+j^2b_2)d_1]\dot{K}^2 \\
& -\frac{1}{4}\left\{\left[\frac{\left(b_5-\frac{b_3e_1}{c_3}\right)}{(b_1+j^2b_2)}\left(\frac{b_3e_2}{c_3}\right)\right]' + \frac{\left[\left(b_5-\frac{b_3e_1}{c_3}\right)^2-4(b_1+j^2b_2)f_1\right]}{(b_1+j^2b_2)}\right\}\dot{\varphi}_1{}^2 \\
& +\frac{1}{2(b_1+j^2b_2)}\left(\frac{b_3e_2}{c_3}\right)j^2b_4\dot{\varphi}_1{}'\dot{K} - \frac{1}{(b_1+j^2b_2)}\left(b_5-\frac{b_3e_1}{c_3}\right)j^2b_4\dot{\varphi}_1\dot{K} \\
& -\frac{1}{4a_4}\left(\frac{a_2e_2}{c_3}\right)^2\varphi_1''^2 - \left\{\frac{1}{4a_4}\left(\frac{a_2e_2}{c_3}\right)\left[a_1-\frac{a_2e_1}{c_3}-\left(\frac{a_2e_2}{c_3}\right)'+\frac{j^2a_6e_2}{c_3}\right]\right\}'\varphi_1'^2 \\
& -\frac{1}{2a_4}\left(\frac{a_2e_2}{c_3}\right)\left[a_1'-\left(\frac{a_2e_1}{c_3}\right)'+\frac{j^2a_6e_1}{c_3}\right]\varphi_1'^2 \\
& -\left\{\frac{1}{4a_4}\left[a_1-\frac{a_2e_1}{c_3}-\left(\frac{a_2e_2}{c_3}\right)'+\frac{j^2a_6e_2}{c_3}\right]^2-\left(f_2+\frac{c_1e_2^2}{c_3^2}-\frac{c_5e_2}{c_3}\right)\right\}\varphi_1'^2 \\
& -\left\{\frac{1}{4a_4}\left[a_1'-\left(\frac{a_2e_1}{c_3}\right)'+\frac{j^2a_6e_1}{c_3}\right]^2-\left(j^2f_3-\frac{c_4e_1}{c_3}+\frac{c_1e_1^2}{c_3^2}\right)\right\}\varphi_1^2 \\
& +\left(\left\{\frac{1}{4a_4}\left(\frac{a_2e_2}{c_3}\right)\left[a_1'-\left(\frac{a_2e_1}{c_3}\right)'+\frac{j^2a_6e_1}{c_3}\right]\right\}''-\left(\frac{2c_1e_1e_2}{c_3^2}-\frac{c_5e_1}{c_3}-\frac{c_4e_2}{c_3}\right)'\right)\varphi_1^2 \\
& +\left\{\frac{1}{4a_4}\left[a_1-\frac{a_2e_1}{c_3}-\left(\frac{a_2e_2}{c_3}\right)'+\frac{j^2a_6e_2}{c_3}\right]\left[a_1'-\left(\frac{a_2e_1}{c_3}\right)'+\frac{j^2a_6e_1}{c_3}\right]\right\}'\varphi_1^2 \\
& +\frac{1}{2a_4}\left(\frac{a_2e_2}{c_3}\right)j^2a_3\varphi_1''K' - \frac{1}{2a_4}\left[a_1-\frac{a_2e_1}{c_3}-\left(\frac{a_2e_2}{c_3}\right)'+\frac{j^2a_6e_2}{c_3}\right]j^2a_3\varphi_1'K' \\
& -\frac{1}{2a_4}\left[a_1'-\left(\frac{a_2e_1}{c_3}\right)'+\frac{j^2a_6e_1}{c_3}\right]j^2a_3\varphi_1K' + \frac{1}{2a_4}\left(\frac{a_2e_2}{c_3}\right)j^2(a_3'+a_7)\varphi_1''K \\
& -\left\{\frac{1}{2a_4}j^2\left(a_3'+a_7\right)\left[a_1-\frac{a_2e_1}{c_3}-\left(\frac{a_2e_2}{c_3}\right)'+\frac{j^2a_6e_2}{c_3}\right]+j^2\frac{e_2c_2}{c_3}\right\}\varphi_1'K
\end{aligned}$$

$$
\begin{aligned}
&-\left\{\frac{1}{2a_4}j^2\left(a_3'+a_7\right)\left[a_1'-\left(\frac{a_2e_1}{c_3}\right)'+\frac{j^2a_6e_1}{c_3}\right]+j^2\left(\frac{e_1c_2}{c_3}-d_3\right)\right\}\varphi_1K\\
&-\left[\frac{1}{4a_4}j^4\left(a_3'-a_7\right)^2+j^2d_2\right]K^2-\frac{1}{4a_4}j^4a_3^2K'^2-\frac{1}{2a_4}j^4a_3\left(a_3'+a_7\right)K'K.
\end{aligned}
$$