

学位論文

Systematic Construction of
Healthy Gravitational Theories
with Higher Derivatives

(高階微分を含んだ健全な重力理論の系統的構築)

平成29年12月博士(理学)申請

東京大学大学院理学系研究科物理学専攻
高橋 一史

Abstract

After the discovery of the late-time cosmic acceleration, there has been a growing interest in modified theories of gravity. Today we have so many models that it is inefficient to handle them separately. This situation necessitates constructing a unified framework of gravitational theories to treat them together. In the context of scalar-tensor theories, there is a class called Horndeski theory, which is the most general scalar-tensor theory that produces second-order Euler-Lagrange equations. This second-order nature of the Horndeski theory ensures that the theory is free from the problem of Ostrogradsky ghost associated with higher-order equations of motion, and thus forms a general class of “healthy” scalar-tensor theories. Indeed, the Horndeski class encompasses many known theories and its cosmological implications have been extensively studied.

It had been believed that the Horndeski class is the most general healthy theory, but the myth was destroyed: Even if higher-order derivatives appear in Euler-Lagrange equations, there are some cases where those higher derivatives can be eliminated by algebraic manipulation to yield a set of second-order differential equations. This is possible only if higher derivative terms are contained in the action in a special combination. Several classes of healthy theories beyond Horndeski have been proposed, but the whole picture of scalar-tensor theories without Ostrogradsky ghost remains hardly understood. Some of these theories can be obtained by disformal transformation of the Horndeski class, where a disformal transformation is a redefinition of the metric that depends on the derivative of the scalar field. As such, a disformal transformation generically maps a scalar-tensor theory to another one with higher derivatives and may play a crucial role in extending the framework of healthy scalar-tensor theories.

Once a new theory without Ostrogradsky ghost is obtained, we should still investigate its theoretical viability from other perspectives, e.g., whether the theory accommodates stable cosmology or not. Regarding this point, it has been shown that all the known healthy theories that cannot be obtained by disformal transformation from the Horndeski class are plagued with gradient instabilities in cosmological perturbations. This implies that, within known healthy scalar-tensor theories, only those connected with the Horndeski theory via disformal transformation admit viable cosmology.

In light of this situation, we propose a new methodology to construct healthy theories by means of transformation of variables. We apply the technique to the case of disformal

transformation on scalar-tensor theories and yield a broad class of healthy theories beyond any existing model. We also investigate the stability of cosmological perturbations in the so-obtained theory. It is shown that perturbations about a flat, homogeneous, and isotropic cosmological background always suffer from ghost/gradient instabilities.

Contents

Notations and conventions	v
1 Introduction	1
2 Scalar-tensor theories	5
2.1 Degenerate higher-order scalar-tensor theories	5
2.1.1 Horndeski theory	6
2.1.2 GLPV theory	8
2.1.3 Quadratic/cubic DHOST theory	9
2.1.4 Extended Galileons	11
2.2 Cosmological perturbations of known theories	11
2.3 Further extensions	12
3 Invertible transformation	15
3.1 Examples	17
3.1.1 Analytical mechanics	17
3.1.2 Scalar-tensor theory	18
3.2 Proof of the Theorem	20
3.2.1 Setup	20
3.2.2 The Theorem	21
3.2.3 Remarks	24
3.3 Applications to scalar-tensor theories	30
3.3.1 Disformal transformation	30
3.3.2 Mixing with derivatives of the metric	32
3.3.3 Possible extensions	33
4 Noninvertible transformation	35
4.1 ADM form of the seed scalar-tensor theory	37
4.1.1 The seed action	37
4.1.2 ADM decomposition	38
4.2 Extended mimetic gravity	39

4.2.1	Hamiltonian analysis	39
4.2.2	Remarks	44
4.3	Cosmological perturbations	46
5	Conclusions	51
	Appendices	57
A	Theorem of Ostrogradsky	59
B	Disformal transformation	63
B.1	Disformal transformation of scalar-tensor theories	63
B.2	Noninvertible disformal transformation	65
B.3	Multi-disformal transformation	67
B.4	Vector disformal transformation	68
C	Unique solvability of DAEs	71
C.1	Conditions for the unique solvability	71
C.2	Standard canonical form of ODAEs	73
C.3	Adjoint ODAE	75
D	Noether identity	77
D.1	Spatial diffeomorphism	79
D.2	Conformal symmetry	80

Notations and conventions

The notations and conventions used in this thesis are summarized.

- A dot stands for differentiation with respect to t : $\dot{\phi} \equiv \partial\phi/\partial t$.
- The symmetrization/antisymmetrization symbols are defined as

$$T_{(\mu_1\mu_2\cdots\mu_n)} \equiv \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} T_{\mu_{\sigma(1)}\mu_{\sigma(2)}\cdots\mu_{\sigma(n)}},$$
$$T_{[\mu_1\mu_2\cdots\mu_n]} \equiv \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma) T_{\mu_{\sigma(1)}\mu_{\sigma(2)}\cdots\mu_{\sigma(n)}},$$

where \mathfrak{S}_n denotes the symmetric group of degree n . Note that any antisymmetrization over indices more than the dimensionality of space(time) is identical to zero.

- The generalized Kronecker delta is defined by

$$\delta_{\mu_1\mu_2\cdots\mu_n}^{\alpha_1\alpha_2\cdots\alpha_n} \equiv n! \delta_{[\mu_1}^{\alpha_1} \delta_{\mu_2}^{\alpha_2} \cdots \delta_{\mu_n]}^{\alpha_n}.$$

We mainly focus on scalar-tensor theories in 4 spacetime dimensions, in which the action is expressed in the language of 4-dimensional (pseudo-)Riemannian geometry. On the other hand, in order to perform a canonical analysis, it is useful to decompose the spacetime into one timelike direction and three orthogonal spacelike ones. Therefore, we employ the following notations depending on the context.

4-dimensional quantities

- For indices of 4-dimensional tensors, we use Greek letters $\alpha, \beta, \cdots, \mu, \nu, \cdots$.
- The “mostly-plus” convention is used for the spacetime metric $g_{\mu\nu}$.
- A covariant derivative associated with $g_{\mu\nu}$ is denoted as ∇_μ .

- The 4-dimensional Riemann tensor $\mathcal{R}^\mu{}_{\nu\lambda\sigma}$ is defined so that

$$2\nabla_{[\lambda}\nabla_{\sigma]}v^\mu = \mathcal{R}^\mu{}_{\nu\lambda\sigma}v^\nu,$$

where v^μ is an arbitrary vector.

- The 4-dimensional Ricci tensor $\mathcal{R}_{\mu\nu}$ is obtained through the contraction of the first and third indices of the Riemann tensor: $\mathcal{R}_{\mu\nu} \equiv \mathcal{R}^\alpha{}_{\mu\alpha\nu}$.
- The 4-dimensional Ricci scalar and Einstein tensor are expressed as \mathcal{R} and $\mathcal{G}_{\mu\nu}$, respectively.
- The double dual of the Riemann tensor is defined by

$$\mathcal{L}^{\mu\nu}{}_{\alpha\beta} \equiv \frac{1}{4}\delta^{\mu\nu\lambda\sigma}\mathcal{R}^{\gamma\delta}{}_{\lambda\sigma} = \mathcal{R}^{\mu\nu}{}_{\alpha\beta} - 4\mathcal{R}^{\mu[\alpha}\delta^{\nu]\beta} + \frac{1}{2}\mathcal{R}\delta^{\mu\nu}{}_{\alpha\beta}.$$

- The first and second covariant derivatives of the scalar field ϕ are often denoted as $\nabla_\mu\phi \equiv \phi_{,\mu}$ and $\nabla_\mu\nabla_\nu\phi \equiv \phi_{\mu\nu}$, respectively.

ADM variables

- For indices of 3-dimensional tensors, we use Latin letters a, b, \dots, i, j, \dots .
- Arnowitt-Deser-Misner (ADM) decomposition for the metric:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -N^2 dt^2 + \gamma_{ij}(dx^i + N^i dt)(dx^j + N^j dt),$$

where N , N^i , and γ_{ij} represent the lapse function, shift vector, and spatial metric, respectively. γ_{ij} is used to raise/lower spatial indices.

- A covariant derivative associated with the spatial metric γ_{ij} is denoted as D_i .
- The unit normal vector to a constant-time hypersurface is denoted as $n_\mu \equiv -N\delta_\mu^0$, and the projection tensor as $h_{\mu\nu} \equiv g_{\mu\nu} + n_\mu n_\nu$.
- The extrinsic curvature is defined by

$$K_{ij} \equiv \frac{1}{2N} (\dot{\gamma}_{ij} - 2D_{(i}N_{j)}),$$

and its trace is denoted as $K \equiv \gamma^{ij}K_{ij}$.

- The 3-dimensional Ricci tensor, Ricci scalar, and Einstein tensor are denoted as R_{ij} , R , and G_{ij} , respectively.

Chapter 1

Introduction

General relativity (GR) is the current “standard model” of gravity, which passes all the Solar System tests performed so far [1]. On the other hand, the late-time cosmic acceleration [2, 3] and extreme gravitational phenomena like gravitational waves originating from merging compact object binaries [4–9] may exhibit some deviation from GR, which motivates us to investigate modified theories of gravity (see Ref. [10] for a recent review). Even if GR is the correct model to describe gravitation, it is still important to study alternative theories of gravity to establish the validity of GR.

A basic strategy for modification of gravity is to add some new dynamical degree of freedom (DOF) to GR. Depending on what kind of DOF is introduced, modified gravity theories are categorized into several classes: Scalar-tensor theories [11–20] are those containing an additional scalar DOF, vector-tensor theories [21–23] are those with an additional vector DOF, and bimetric theories [24] have an additional tensor DOF. Among these, we mainly focus on the (single-field) scalar-tensor class throughout this thesis. Even within this class of theories, so many models have been proposed that it is almost impossible to examine their theoretical and observational consistency on a one-by-one basis. A possible way to overcome the current situation is to construct a framework to treat the theories in a unified manner, namely, to write down the most general action that meets some guiding principles. We first require general covariance, but this does not so much reduce the space of allowed theories: Any higher derivative term like $\phi_t^\mu \phi_\lambda^\nu \phi_\mu^\lambda$, $\mathcal{R}(\square\phi)^2$, $\nabla_\mu \mathcal{R} \nabla^\mu \mathcal{R}$, etc. is allowed at this stage.^{*1}

When one constructs such field theories with higher derivatives, another guiding principle comes from the theorem of Ostrogradsky [26], which states that any theory described by a nondegenerate higher derivative Lagrangian possesses linear momenta in its Hamiltonian on the constraint surface. Here, a Lagrangian L with higher-order derivatives is said to be nondegenerate if its kinetic matrix (i.e., the second derivatives of L with re-

^{*1}In this thesis, we mainly focus on cases where third or higher derivative appears in the action for simplicity. Analytical mechanics models with third derivatives have recently been studied in Ref. [25].

spect to the velocities associated with the higher derivatives) is nondegenerate. The linear momenta render the Hamiltonian unbounded below, which implies that the energy of the system can become arbitrarily negative. Although such a system itself is not pathological, once it is coupled to an ordinary system where the Hamiltonian is bounded below, the system as a whole becomes unstable since negative/positive DOFs will rapidly be excited by interchanging energy between the two subsystems. This instability is called Ostrogradsky ghost.

To illustrate the statement of the Ostrogradsky's theorem, let us consider a nondegenerate Lagrangian in analytical mechanics

$$L = \frac{1}{2}\ddot{x}^2, \quad (1.1)$$

which is rewritten by introducing an auxiliary field y as

$$L' = -\dot{x}\dot{y} - \frac{1}{2}y^2. \quad (1.2)$$

The equivalence between L and L' can be seen as follows: The equation of motion (EOM) for y reads $y = \dot{x}$, which is substituted back into L' to yield the original Lagrangian L . Then, we redefine the coordinates as $X \equiv (x - y)/\sqrt{2}$ and $Y \equiv (x + y)/\sqrt{2}$, which diagonalizes the kinetic part of L' as

$$L' = \frac{1}{2}\dot{X}^2 - \frac{1}{2}\dot{Y}^2 - \frac{1}{4}(X - Y)^2. \quad (1.3)$$

Here, one may notice that the Lagrangian describes a 2-DOF system and the extra DOF (Y in the present case) represents the Ostrogradsky ghost as its kinetic term has a wrong sign. This example captures the essence of the theorem of Ostrogradsky (see Appendix A for a more general proof in the language of Hamiltonian analysis). Therefore, a theory without the problem of Ostrogradsky ghost, often referred to as a *healthy* theory, must have a degenerate Lagrangian. This requirement poses tight constraints on the higher derivative structure of healthy field theories [27, 28].

Within scalar-tensor theories in four dimensions, the Horndeski theory [14] (or its equivalent formulation known as generalized Galileons [15, 16]) provides a basic ground for studying a wide class of such healthy theories having three DOFs, since it is the most general theory that yields second-order Euler-Lagrange (EL) equations. There are further possibilities of healthy theories beyond the Horndeski class, such as Gleyzes-Langlois-Piazza-Vernizzi (GLPV) theories [17] and quadratic/cubic degenerate higher-order scalar-tensor (DHOST) theories [18–20].^{*2} Those quadratic/cubic DHOST theories form the

^{*2}Throughout this thesis, we restrict ourselves to theories that possess general covariance. If one relaxes this requirement to only spatial covariance, yet broader classes of theories can be obtained such as extended Galileons [29, 30]. For details, see §2.1.4.

broadest class of healthy scalar-tensor theories known so far. However, these theories are obtained under the assumption that the Lagrangian depends on up to quadratic/cubic order in $\phi_{\mu\nu}$ (hence the name “quadratic/cubic DHOST”), and thus the very boundary of healthy scalar-tensor theories remains unknown. To go even further, disformal transformations [31] may play a key role:

$$\tilde{g}_{\mu\nu} = A(\phi, X)g_{\mu\nu} + B(\phi, X)\phi_{\mu}\phi_{\nu}, \quad X \equiv g^{\mu\nu}\phi_{\mu}\phi_{\nu}. \quad (1.4)$$

Here, the functions A, B are chosen so that the transformation does not change the metric signature and is consistent with the existence of the inverse matrix of $\tilde{g}_{\mu\nu}$ [31–33]. Note that the transformation (1.4) reduces to a conformal transformation when $B = 0$. Some quadratic/cubic DHOST theories are obtained from the Horndeski theory via disformal transformation [20], which motivates us to think that the quadratic/cubic DHOST class could further be extended in the same manner. Unfortunately, this is not the case because the set of cubic scalar-tensor theories is closed under disformal transformations. Therefore, it seems difficult to generate a new class of healthy theories starting from known healthy theories.

In light of this situation, it is natural to ask whether healthy theories can be generated by performing field transformations on *nondegenerate* theories that contain extra ghost DOFs. Generally speaking, field transformations are classified broadly into *invertible* and *noninvertible* ones. Invertible transformations are such that there is a one-to-one correspondence between the old and new sets of field variables, while noninvertible transformations are those that are not invertible. One may naively think that a nondegenerate theory cannot be mapped to a degenerate theory with fewer DOFs via invertible transformation due to its bijective nature, but at the same time, this may not be obvious when the transformation law contains field derivatives as in the above case of disformal transformations. This is because the resultant theory generically produces EL equations of order higher than the original one. Hence, we need a detailed analysis of the relation between two systems connected by invertible transformation to fill the gap between the two conflicting intuitions. On the other hand, as for noninvertible transformations, there is a growing interest on a framework called *mimetic gravity* [34] (see Ref. [35] for a review). A mimetic gravity model is obtained by performing a noninvertible disformal transformation on some “seed” scalar-tensor theory and exhibits various aspects depending on which theory is chosen as a seed [36–40]. It is notable that even a nondegenerate seed theory could be mapped to a degenerate one [41]. Along this line, it is intriguing to investigate for which seed scalar-tensor theories their mimetic gravity counterparts are degenerate.

It should be noted that, even if we find a theory that circumvents the problem of Ostrogradsky ghost, it could have some other instabilities. In this sense, the “healthiness” is a necessary but not sufficient condition for a theory to be fully viable. One of the possible tests is whether the theory admits stable cosmology or not. If the action is expanded up to quadratic order in perturbations around the flat Friedmann-Lemaître-Robertson-Walker (FLRW) background, one can infer the (in)stability of the theory from

the coefficients of the quadratic action. Interestingly, the authors of Ref. [42] showed that any quadratic/cubic DHOST theory that cannot be mapped to the Horndeski class is plagued by gradient instabilities.

In this thesis, we first address the issue on invertible transformations by demonstrating a theorem on the equivalence between the EL equations obtained in the old frame and those obtained in the new frame. It should be noted that the theorem assumes only the invertibility of the transformation law and is not restricted to disformal transformations on scalar-tensor theories. This allows us to conclude that one cannot generate healthy theories from a nondegenerate theory via invertible transformation. Then, we discuss the case of noninvertible transformations, focusing on the noninvertible conformal transformation. We specify a general class of seed scalar-tensor theories that yields degenerate theories as its mimetic counterpart. This can be regarded as an extension of known mimetic gravity models, and thus we call the resultant class of theories *extended mimetic gravity*. Moreover, these new extended mimetic models also serve as a nontrivial extension of known DHOST theories in the sense that they cannot be obtained by any disformal transformation of known healthy theories. We also examine the stability of perturbations about the flat FLRW background in the extended mimetic gravity. It is shown that either of tensor/scalar perturbations suffers from gradient instabilities, except for cases where scalar perturbations are presumably strongly coupled, or otherwise ghost instabilities arise.

The rest of the thesis is organized as follows. In Chap. 2, we review the healthy scalar-tensor theories proposed so far. Next, in Chap. 3, we prove a theorem on the equivalence between the EL equations before performing an invertible transformation and those after and show that the number of DOFs of a theory is not changed by invertible transformation. This part is based on our published paper [43]. Then in Chap. 4, we introduce the extended mimetic gravity and perform a Hamiltonian analysis to show that this class of theories has at most three DOFs. The stability of cosmological perturbations in the extended mimetic theories is also discussed. This part is based on our published paper [44]. Finally, we draw our conclusions in Chap. 5.

Chapter 2

Scalar-tensor theories

2.1 Degenerate higher-order scalar-tensor theories

The history of scalar-tensor theories goes back to Jordan-Brans-Dicke theory [11, 12], whose action contains a scalar field nonminimally coupled to gravity:

$$S_{\text{BD}} = \int d^4x \sqrt{-g} \left(\phi \mathcal{R} - \frac{\omega_{\text{BD}}}{\phi} X \right). \quad (2.1)$$

Originally, the parameter ω_{BD} was assumed to be a constant but later it was promoted to a function of ϕ [45]. Another early example is $f(\mathcal{R})$ gravity [13, 46, 47] (see Refs. [48, 49] for reviews). The name comes from an arbitrary function of the Ricci scalar, $f(\mathcal{R})$, which constitutes the action:

$$S_f = \int d^4x \sqrt{-g} f(\mathcal{R}). \quad (2.2)$$

Since we are interested in deviations from GR, it is natural to assume $f''(\mathcal{R}) \neq 0$. Although this theory written in the form (2.2) does not contain scalar field in appearance, it can be recast as follows:

$$S'_f = \int d^4x \sqrt{-g} [f(\phi) - f'(\phi)(\phi - \mathcal{R})]. \quad (2.3)$$

Indeed, one finds $\phi = \mathcal{R}$ by using its EOM and thus can reproduce the action (2.2). Pioneered by these theories, literally hundreds of theories have been developed so far.

The Jordan-Brans-Dicke theory (2.1) and $f(\mathcal{R})$ gravity reformulated as a scalar-tensor theory (2.3) contain at most first derivative of ϕ . Even more complex theories containing second derivatives of the scalar field have recently been investigated, which we will see in subsequent sections. As such, there has been a growing interest in the generalization of gravitational theories. The main motivation to generalize theories of gravity is to provide a useful framework to handle a large number of theories collectively. Along this line, it

is important to specify to what extent we can generalize it. Concerning this point, the theorem of Ostrogradsky offers a possible answer: *Any nondegenerate higher derivative field theories possesses extra DOFs, which exhibit ghost instabilities* (see Appendix A). This theorem poses so tight a constraint on higher derivative scalar-tensor theories that only a tiny fraction of them can avoid the generic instability. Therefore, it serves as a guiding principle to construct “healthy” scalar-tensor theories. Below, we review some known classes of scalar-tensor theories with a degenerate Lagrangian and thus circumvent the problem of Ostrogradsky ghost.

2.1.1 Horndeski theory

Within the class of theories written only by the metric, GR with a cosmological constant is the most general theory in four dimensions that yields second-order EL equations, which is known as Lovelock’s theorem [50]. This second-order nature of EL equations is a sufficient (but not necessary) condition of a theory to be degenerate, which ensures the absence of Ostrogradsky ghost.

The Horndeski class [14–16] is a collection of scalar-tensor theories with such nature, i.e., the most general single-field scalar-tensor theories that produce second-order EL equations both for the metric and the scalar field. This theory was originally invented by Horndeski [14] and then rediscovered in the context of Galileons [15], though the equivalence between them are nontrivial [16]. The action of Horndeski theories is given by

$$S_{\text{H}} = \int d^4x \sqrt{-g} (L_2^{\text{H}} + L_3^{\text{H}} + L_4^{\text{H}} + L_5^{\text{H}}), \quad (2.4)$$

where

$$\begin{aligned} L_2^{\text{H}} &= G_2(\phi, X), \\ L_3^{\text{H}} &= G_3(\phi, X)\square\phi, \\ L_4^{\text{H}} &= G_4(\phi, X)\mathcal{R} - 2G_{4X} [(\square\phi)^2 - \phi_{\mu}^{\nu}\phi_{\nu}^{\mu}], \\ L_5^{\text{H}} &= G_5(\phi, X)\mathcal{G}^{\mu\nu}\phi_{\mu\nu} + \frac{1}{3}G_{5X} [(\square\phi)^3 - 3(\square\phi)\phi_{\mu}^{\nu}\phi_{\nu}^{\mu} + 2\phi_{\mu}^{\nu}\phi_{\nu}^{\lambda}\phi_{\lambda}^{\mu}]. \end{aligned} \quad (2.5)$$

Here, G_i are arbitrary functions of the scalar field ϕ and its kinetic term X . Note that the terms L_4^{H} and L_5^{H} respectively contain quadratic and cubic powers of $\phi_{\mu\nu}$, which produce higher derivative terms when varied with respect to ϕ . The terms with curvature tensors are crucial for canceling out those higher derivative terms. This cancellation can be achieved only if the coefficients of each term are chosen as in Eq. (2.5). Without the tuning of the coefficients, the theory fails to be degenerate. Any theory with second-order EL equations should be recast in the form (2.4) after integration by parts and/or some field redefinition.

The Horndeski class includes many classes of scalar-tensor theories. For example, if we choose the arbitrary functions in Eq. (2.5) as

$$G_2 = -\frac{1}{2}X - V(\phi), \quad G_3 = 0, \quad G_4 = \frac{M_{\text{Pl}}^2}{2}, \quad G_5 = 0, \quad (2.6)$$

we recover GR plus a canonical scalar field with potential $V(\phi)$. A more nontrivial example is $f(\mathcal{R})$ gravity (2.3), which amounts to the choice

$$G_2 = f(\phi) - \phi f'(\phi), \quad G_3 = 0, \quad G_4 = f'(\phi), \quad G_5 = 0. \quad (2.7)$$

By construction, any theory with second-order EL equations should be recast in the form (2.4), even if it contains nontrivial coupling with curvature tensors. For instance, the term containing the double dual Riemann tensor

$$\mathcal{L}^{\mu\nu}{}_{\alpha\beta}\phi^\alpha\phi_\mu\phi_\nu^\beta = \mathcal{R}^{\mu\nu}{}_{\alpha\beta}\phi^\alpha\phi_\mu\phi_\nu^\beta + \mathcal{R}^{\mu\nu}(-\phi_\mu\phi_\nu\Box\phi + 2\phi_\alpha\phi_\mu^\alpha\phi_\nu) + \mathcal{R}(X\Box\phi - 2\phi_\mu\phi_\nu^\mu\phi^\nu) \quad (2.8)$$

is known to have second-order field equations. As it should be, this theory is equivalent to the Horndeski theory with $G_2 = G_3 = G_4 = 0$ and $G_5 = -X/2$ [51]. Yet another example is nonminimal coupling to the Gauss-Bonnet scalar:

$$\xi(\phi)\mathcal{L}^{\mu\nu}{}_{\alpha\beta}\mathcal{R}^{\alpha\beta}{}_{\mu\nu} = \xi(\phi)(\mathcal{R}^2 - 4\mathcal{R}_{\mu\nu}\mathcal{R}^{\mu\nu} + \mathcal{R}_{\mu\nu\lambda\sigma}\mathcal{R}^{\mu\nu\lambda\sigma}), \quad (2.9)$$

which corresponds to the following choice [16]:

$$\begin{aligned} G_2 &= 2\xi^{(4)}X^2(3 - \ln|X|), & G_3 &= 2\xi^{(3)}X(7 - 3\ln|X|), \\ G_4 &= -2\xi^{(2)}X(2 - \ln|X|), & G_5 &= -4\xi^{(1)}\ln|X|, \end{aligned} \quad (2.10)$$

with $\xi^{(n)} \equiv d^n\xi/d\phi^n$. One can confirm these results by comparing the field equations.

In the language of ADM variables and taking the unitary gauge $\phi = t$, the Horndeski action (2.4) is written as follows:

$$\begin{aligned} S_{\text{H}} &= \int dt d^3x N \sqrt{\gamma} [A_2 + A_3 K + A_4(K^2 - K_j^i K_i^j) + B_4 R \\ &\quad + A_5(K^3 - 3K K_j^i K_i^j + 2K_j^i K_k^j K_i^k) + B_5 G^{ij} K_{ij}]. \end{aligned} \quad (2.11)$$

Here, the coefficients A_2, A_3, A_4, A_5, B_4 , and B_5 are functions of (t, N) , but only four of them are independent, which corresponds to the number of arbitrary functions in the Horndeski action (2.4). Indeed, the coefficients in Eq. (2.11) are written in terms of the arbitrary functions in Eq. (2.4) as [52]

$$\begin{aligned} A_2 &= G_2 - X g_{3\phi}, & A_3 &= 2(-X)^{3/2} g_{3X} - 2\sqrt{-X} G_{4\phi}, \\ A_4 &= -G_4 + 2X G_{4X} + \frac{X}{2} G_{5\phi}, & B_4 &= G_4 + \frac{X}{2} (G_{5\phi} - g_{5\phi}), \\ A_5 &= -\frac{(-X)^{3/2}}{3} G_{5X}, & B_5 &= -\sqrt{-X} g_5, \end{aligned} \quad (2.12)$$

where the auxiliary functions g_i are defined so that

$$g_3 + 2Xg_{3X} = G_3, \quad g_{5X} + \frac{g_5}{2X} = G_{5X}. \quad (2.13)$$

From Eqs. (2.12) and (2.13) one finds the following relations [17]:

$$A_4 = -B_4 + 2XB_{4X}, \quad A_5 = -\frac{X}{3}B_{5X}. \quad (2.14)$$

In these expressions, ϕ and X should be read as t and $-1/N^2$, respectively. Note also that the auxiliary functions g_i are determined only up to $c_i(\phi)/\sqrt{-X}$ or $c_i(t)N$, with c_i being arbitrary functions of ϕ ($= t$). However, these ambiguities are irrelevant as they only contribute as total derivative:

$$\begin{aligned} & N\sqrt{\gamma} \left(\frac{\dot{c}_3}{N} + c_3K + \frac{\dot{c}_5}{2N}R - c_5G^{ij}K_{ij} \right) \\ &= \partial_t \left[\sqrt{\gamma} \left(c_3 + \frac{c_5}{2}R \right) \right] + \partial_i \left\{ \sqrt{\gamma} \left[-c_3N^i + \frac{c_5}{2} (2G^{ij}N_j + D_j\dot{\gamma}^{ij} + D^i(\gamma^{kl}\dot{\gamma}_{kl})) \right] \right\}. \end{aligned} \quad (2.15)$$

2.1.2 GLPV theory

It was a common belief that the Horndeski theory is the most general healthy scalar-tensor theory, but the possibility to go beyond Horndeski was suggested in Ref. [53]. The class of GLPV theories [17] was proposed as the first example of healthy theories beyond Horndeski, which contains two more arbitrary functions of (ϕ, X) than the Horndeski class:

$$S_{\text{GLPV}} = S_{\text{H}} + \int d^4x \sqrt{-g} (L_4^{\text{bH}} + L_5^{\text{bH}}), \quad (2.16)$$

where

$$\begin{aligned} L_4^{\text{bH}} &= F_4(\phi, X) \left\{ X \left[(\Box\phi)^2 - \phi_\mu^\nu \phi_\nu^\mu \right] - 2\phi_\mu \phi_\nu^\mu (\phi^\nu \Box\phi - \phi_\lambda^\nu \phi^\lambda) \right\}, \\ L_5^{\text{bH}} &= F_5(\phi, X) \left\{ X \left[(\Box\phi)^3 - 3(\Box\phi) \phi_\mu^\nu \phi_\nu^\mu + 2\phi_\mu^\nu \phi_\nu^\lambda \phi_\lambda^\mu \right] \right. \\ &\quad \left. - 3\phi_\lambda \phi_\sigma^\lambda \phi^\sigma \left[(\Box\phi)^2 - \phi_\mu^\nu \phi_\nu^\mu \right] + 6\phi_\mu \phi_\nu^\mu \phi^\sigma (\phi_\sigma^\nu \Box\phi - \phi_\lambda^\nu \phi_\sigma^\lambda) \right\}. \end{aligned} \quad (2.17)$$

This theory was obtained by detuning the relation (2.14) and then restoring general covariance by use of the Stückelberg trick [54, 55] (see §2.1.4). Indeed, written in terms of ADM variables and taking $\phi = t$, the new terms L_4^{bH} and L_5^{bH} contribute to A_4 and A_5 as

$$A_4 \supset -X^2 F_4, \quad A_5 \supset (-X)^{5/2} F_5, \quad (2.18)$$

which spoils the relation (2.14). Note that this detuning does not modify the result of Hamiltonian analysis under the unitary gauge.

Although the action (2.16) was believed to provide DHOST theories, this turned out to be not the case for generic choices of F_4 and F_5 [56]. The reason is that the higher derivative structure of the theory could be hidden under the unitary gauge. For the theory (2.16) to be degenerate without unitary gauge fixing, the arbitrary functions must satisfy the following constraint [20]:

$$XG_{5X}F_4 = 3F_5(G_4 - 2XG_{4X}). \quad (2.19)$$

which ensures that the action (2.16) is obtained by some disformal transformation of the Horndeski action (2.4) [20, 56].

2.1.3 Quadratic/cubic DHOST theory

The quadratic/cubic DHOST class [18–20] is the collection of all the DHOST theories that contain second derivatives of the scalar field up to cubic order. The action has the form of

$$S_{q/c} = \int d^4x \sqrt{-g} \left(f_2 \mathcal{R} + \sum_{i=1}^5 a_i L_i^{(2)} + f_3 \mathcal{G}^{\mu\nu} \phi_{\mu\nu} + \sum_{j=1}^{10} b_j L_j^{(3)} \right), \quad (2.20)$$

where f_2 , f_3 , a_i , and b_j are arbitrary functions of (ϕ, X) . Here, the building blocks $L_i^{(2)}$ and $L_j^{(3)}$ are defined as

$$\begin{aligned} L_1^{(2)} &= \phi_\mu^\nu \phi_\nu^\mu, & L_2^{(2)} &= (\Box\phi)^2, & L_3^{(2)} &= (\Box\phi) \phi^\mu \phi_\mu^\nu \phi_\nu, \\ L_4^{(2)} &= \phi^\mu \phi_\mu^\nu \phi_\nu^\lambda \phi_\lambda, & L_5^{(2)} &= (\phi^\mu \phi_\mu^\nu \phi_\nu)^2, \end{aligned} \quad (2.21)$$

and

$$\begin{aligned} L_1^{(3)} &= (\Box\phi)^3, & L_2^{(3)} &= (\Box\phi) \phi_\mu^\nu \phi_\nu^\mu, & L_3^{(3)} &= \phi_\mu^\nu \phi_\nu^\lambda \phi_\lambda^\mu, & L_4^{(3)} &= (\Box\phi)^2 \phi^\mu \phi_\mu^\nu \phi_\nu, \\ L_5^{(3)} &= (\Box\phi) \phi^\mu \phi_\mu^\nu \phi_\nu^\lambda \phi_\lambda, & L_6^{(3)} &= (\phi_\mu^\nu \phi_\nu^\mu) (\phi^\mu \phi_\mu^\nu \phi_\nu), & L_7^{(3)} &= \phi^\mu \phi_\mu^\nu \phi_\nu^\lambda \phi_\lambda^\sigma \phi_\sigma, \\ L_8^{(3)} &= (\phi^\mu \phi_\mu^\nu \phi_\nu^\lambda \phi_\lambda) (\phi^\sigma \phi_\sigma^\rho \phi_\rho), & L_9^{(3)} &= (\Box\phi) (\phi^\mu \phi_\mu^\nu \phi_\nu)^2, & L_{10}^{(3)} &= (\phi^\mu \phi_\mu^\nu \phi_\nu)^3. \end{aligned} \quad (2.22)$$

Note that all the possible scalar constructed from ϕ_μ and $\phi_{\mu\nu}$ up to cubic powers are exhausted by Eqs. (2.21) and (2.22). The name ‘‘quadratic/cubic’’ comes from the structure of the action: The first two terms in Eq. (2.20) represent quadratic and the last two stand for cubic DHOST theories as they are respectively quadratic and cubic in velocities when written in the ADM language.^{*1} It should be noted that the action (2.20) with a generic choice of the arbitrary functions does not provide a DHOST theory: Each subclass of the quadratic/cubic DHOST class is obtained by choosing the arbitrary functions $f_2, f_3, a_i,$

^{*1}We discard couplings with curvature tensors other than $f_2 \mathcal{R}$ or $f_3 \mathcal{G}^{\mu\nu} \phi_{\mu\nu}$ because they result in Ostrogradsky ghost [20].

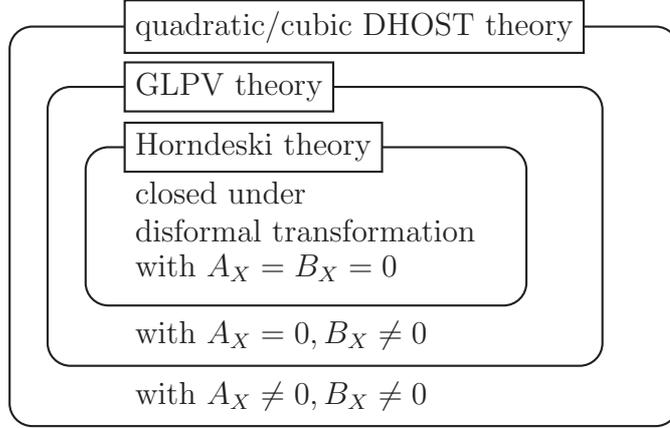


Figure 2.1: The relation between the known DHOST theories. Each class is closed under some specific class of disformal transformations.

and b_j so that the Lagrangian is degenerate. In addition to these arbitrary functions, one can freely add to the action (2.20) terms like $F_0(\phi, X)$ and $F_1(\phi, X)\square\phi$ since it does not affect the (non)degeneracy of a theory. 7 subclasses of quadratic DHOST theories and 9 subclasses of cubic DHOST theories are listed in Ref. [20]. However, a sum of quadratic and cubic DHOST Lagrangians does not necessarily yield a degenerate Lagrangian. For the sum to be degenerate, it is necessary that the null eigenvectors of the two kinetic matrices be parallel. Among the 63 possible combinations, only 7 is freely combined to yield DHOST theories and 18 can be degenerate under certain conditions on the arbitrary functions [20]. It should also be noted that some of these subclasses propagate only one scalar DOF, i.e., no tensor DOFs [42].

So far we have seen three general classes of degenerate scalar-tensor theories. The relation between these classes is depicted in Figure 2.1. It is notable that each class can be characterized by its transformation property under disformal transformations. The authors of Ref. [33] showed that the Horndeski class is closed under disformal transformations with A, B depending on ϕ only. If one proceeds to X -dependent B , the Horndeski theories are transformed to GLPV theories [17, 57], and GLPV theories themselves are closed under the same class of disformal transformations. Further introduction of X -dependence into A results in quadratic/cubic DHOST theories [20].^{*2} One can verify that the quadratic/cubic DHOST class is stable under general disformal transformations (see Appendix B).

^{*2}Apart from this line of research, the authors of Ref. [58] specified all the theories obtained via disformal transformation from the Horndeski class in the language of differential forms.

2.1.4 Extended Galileons

The above DHOST theories possess general covariance, but we obtain still broader classes if we moderate the requirement. One of such examples is the class of *extended Galileons* (or also known as *spatially covariant gravity*) [29, 30], whose action is written in terms of ADM variables as

$$S_{XG} = \int dt d^3x N \sqrt{\gamma} L(N, \gamma_{ij}, R_{ij}, t; K_{ij}; D_i), \quad (2.23)$$

and thus only has 3-dimensional diffeomorphism invariance. Note that the unitary gauge $\phi = t$ is imposed in the context of extended Galileons. This theory was shown to have 3 DOFs by the authors of Refs. [59, 60].

Even though general covariance is broken apparently, it can always be restored via the *Stückelberg trick* [54, 55], where a scalar field ϕ is introduced so that its gradient is proportional to the unit normal vector to a constant-time hypersurface: $n_\mu = -\phi_\mu / \sqrt{-X}$. Then, we perform replacements such as

$$\begin{aligned} t &\rightarrow \phi, & N &\rightarrow \frac{1}{\sqrt{-X}}, & \gamma_{ij} &\rightarrow h_{\mu\nu} = g_{\mu\nu} - \frac{1}{X} \phi_\mu \phi_\nu, \\ K_{ij} &\rightarrow \mathcal{K}_{\mu\nu} \equiv h_\mu^\lambda \nabla_\lambda n_\nu, & D_i N &\rightarrow h_\mu^\nu \nabla_\nu \left(\frac{1}{\sqrt{-X}} \right), \\ R_{ij} &\rightarrow h_\mu^\alpha h_\nu^\gamma h^{\beta\delta} \mathcal{R}_{\alpha\beta\gamma\delta} - \mathcal{K}_\alpha^\mu \mathcal{K}_{\mu\nu} + \mathcal{K}_\mu^\alpha \mathcal{K}_{\alpha\nu}, \end{aligned} \quad (2.24)$$

which restores 4-dimensional diffeomorphism invariance. However, it turns out that the resultant theories generically yield Ostrogradsky ghosts (see Refs. [18, 54] for examples of such theories). The apparent healthy nature of extended Galileons is a consequence of the unitary gauge chosen for the construction of the theory. This misleading gauge choice could eliminate higher derivatives from a given action and change the number of DOFs [18, 42, 54, 56, 61, 62], though it is a complete gauge fixing [63]. This explains the reason why we keep general covariance and avoid taking the unitary gauge from the beginning.

2.2 Cosmological perturbations of known theories

The authors of Ref. [42] showed that the degeneracy conditions for quadratic/cubic DHOST theories boil down to only two simple conditions under a cosmological background (except for cases where tensor DOFs are nondynamical). The quadratic action for perturbations about the flat FLRW spacetime with $ds^2 = -dt^2 + a^2(t)\delta_{ij}dx^i dx^j$ can

be written in the following form [42]:

$$\begin{aligned}
S^{(2)} = \int dt d^3x a^3 \frac{M^2}{2} & \left[\delta K_{ij} \delta K^{ij} - \left(1 + \frac{2}{3} \alpha_L\right) \delta K^2 + (1 + \alpha_T) \left(R \frac{\delta \sqrt{\gamma}}{a^3} + \delta_2 R \right) \right. \\
& + H^2 \alpha_K \delta N^2 + 4H \alpha_B \delta K \delta N + (1 + \alpha_H) R \delta N \\
& \left. + 4\beta_1 \delta K \delta \dot{N} + \beta_2 \delta \dot{N}^2 + \beta_3 \left(\frac{\partial_i \delta N}{a} \right)^2 \right], \tag{2.25}
\end{aligned}$$

where $H \equiv \dot{a}/a$ is the Hubble parameter and $\delta_2 R$ denotes the second-order perturbation of the 3-dimensional curvature scalar R . Note that R itself is a first-order quantity, and the background field equation has been used in deriving the quadratic action (2.25). The degeneracy conditions are obtained as [42]

$$\begin{aligned}
\mathcal{C}_I : \quad & \alpha_L = 0, \quad \beta_2 = -6\beta_1^2, \quad \beta_3 = -2\beta_1 [2(1 + \alpha_H) + \beta_1(1 + \alpha_T)], \tag{2.26} \\
\mathcal{C}_{II} : \quad & \beta_1 = -(1 + \alpha_L) \frac{1 + \alpha_H}{1 + \alpha_T}, \quad \beta_2 = -6(1 + \alpha_L) \left(\frac{1 + \alpha_H}{1 + \alpha_T} \right)^2, \quad \beta_3 = 2 \frac{(1 + \alpha_H)^2}{1 + \alpha_T}. \tag{2.27}
\end{aligned}$$

The quadratic/cubic DHOST class can be divided into three subclasses: Those satisfying only \mathcal{C}_I , those only \mathcal{C}_{II} , and those both \mathcal{C}_I and \mathcal{C}_{II} (see Figure 2.2). Each of $\mathcal{C}_I/\mathcal{C}_{II}$ has a remarkable feature [42]: All the theories within the class \mathcal{C}_I (but outside \mathcal{C}_{II}) can be mapped from the Horndeski class via disformal transformation. As for the class \mathcal{C}_{II} , any theory that belongs to this class is plagued by gradient instabilities: Either of tensor or scalar perturbations have negative sound speed squared. In contrast to the Jeans instability which is responsible for structure formation of the Universe, the gradient instability is problematic because it results in an anomalous growth of small-scale perturbations. Therefore, a theory that admits viable cosmology should lie in $\mathcal{C}_I \setminus \mathcal{C}_{II}$. In other words, any quadratic/cubic DHOST theory that cannot be mapped to the Horndeski class by disformal transformation suffers from gradient instabilities, or otherwise the tensor modes are nondynamical.

2.3 Further extensions

Within healthy scalar-tensor theories in four dimensions that are at most cubic order in $\phi_{\mu\nu}$, the quadratic/cubic DHOST class is the broadest one. In Chap. 4, we obtain a new class of healthy scalar-tensor theories that depend on quartic or higher order in $\phi_{\mu\nu}$. Besides the case of single-field scalar-tensor theories in four dimensions, several other types of generalization have been explored. One is to include third (or higher) derivatives in the action. This point has recently been addressed in the case of analytical mechanics

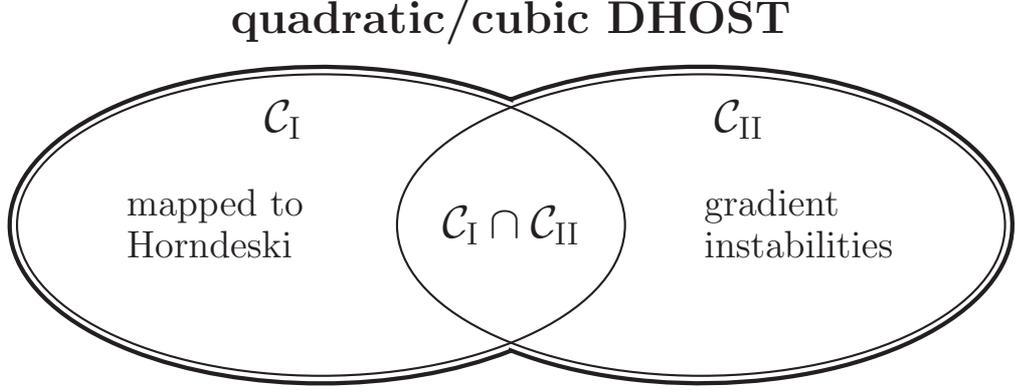


Figure 2.2: The classification of quadratic/cubic DHOST theories in terms of cosmological perturbations. It is fully characterized by the two conditions (2.26) and (2.27), except for those theories with nondynamical tensor modes [42]. The class $\mathcal{C}_I \setminus \mathcal{C}_{II}$ can be mapped to the Horndeski theory by disformal transformation. On the other hand, cosmological perturbations in theories within the class \mathcal{C}_{II} exhibit gradient instabilities.

by the authors of Ref. [25], though no example of scalar-tensor theories has been known.^{*3}

Another generalization is the multi-field extension of scalar-tensor theories. Generalized multi-Galileons [64] provides a broad class of multi-field scalar-tensor theories with second-order EL equations, while it is not the most general one. Employing the similar technique as Ref. [14], the authors of Ref. [65] obtained the most general second-order EL equations of bi-scalar-tensor theory in four dimensions (i.e., *bi-Horndeski theory*). Nevertheless, the Lagrangian that yields this set of EL equations remains unspecified.

Yet another is to consider degenerate scalar-tensor theories in higher dimensions. Although the most general Lagrangian that yields at most second-order EL equations is known up to four dimensions [14], it remains unknown for five or higher dimensions. As was shown in Ref. [16], the framework of generalized Galileons [15], which yields a general class of single-field scalar-tensor theories with second-order EL equations and works in arbitrary dimensions, coincides with the Horndeski theory in four or lower dimensions. One can conjecture that the coincidence holds in any dimensionality, but no rigorous proof for this has been found.

So far we have seen degenerate scalar-tensor theories, but the possibility of degenerate vector-tensor theories has also been explored. The generalized Proca theory [21] provides a broad class of such degenerate vector-tensor theories, which propagates 2 DOFs for tensor and 3 DOFs for vector modes. It is notable that the theory yields EL equations of at most

^{*3}However, one can obtain healthy scalar-tensor theories that contain higher derivatives of the metric by performing a field transformation of the form (3.90) on known healthy theories (see §3.3.2 for details).

Table 2.1: Summary of the known healthy scalar-tensor and vector-tensor theories in four dimensions.

	2nd-order EL eqs.	higher-order EL eqs.
scalar-tensor	Horndeski [14–16]	GLPV [17] quadratic/cubic DHOST [18–20]
vector-tensor	generalized Proca [21]	beyond generalized Proca [22] extended vector-tensor [23]

second order. In this sense, this theory shares the same nature as the Horndeski theory in the case of scalar-tensor theories. Furthermore, there is a stronger relation between these two theories: In the scalar limit $A_\mu \rightarrow \nabla_\mu \phi$, the generalized Proca theory coincides with the shift-symmetric part of the Horndeski theory, i.e., such that invariant under $\phi \rightarrow \phi + \text{constant}$. Motivated by the relation to scalar-tensor theories, the similar extensions as the GLPV and quadratic DHOST theories have been proposed, which are respectively known as beyond generalized Proca [22] and extended vector-tensor theories [23]. The known healthy scalar-tensor and vector-tensor theories are summarized in Table 2.1.

Chapter 3

Invertible transformation

It had been believed that the Horndeski theory [14], also known as the generalized Galileon theory [15, 16], is the most general healthy (single-field) scalar-tensor theory. This was because it forms the broadest class that yields second-order EL equations both for the metric and the scalar field, and thus trivially evades so-called Ostrogradsky ghosts associated with higher-order EOMs [26]. However, it turned out that the second-order nature of the EL equations is just a sufficient condition and not a necessary condition for the absence of Ostrogradsky ghosts [53]. Gleyzes, Langlois, Piazza, and Vernizzi then constructed a healthy theory beyond Horndeski, which is now known as the GLPV theory [17]. The key for healthy theories beyond Horndeski is that the EL equations are *a priori* of higher order, but can be rearranged into a second-order system [61], which is realized in the presence of appropriate degeneracy conditions or a sufficient number of constraints [18, 27, 66]. Thus far many efforts have been made to construct healthy theories beyond Horndeski, which include quadratic/cubic DHOST theories [18, 20].

Along this line, the transformation properties of these theories under the disformal transformation have been investigated in Refs. [17, 19, 20, 30, 33, 53, 57, 67]. As mentioned in Chap. 1, it is a transformation of the metric containing the first derivative of the scalar field:

$$\tilde{g}_{\mu\nu} = A(\phi, X)g_{\mu\nu} + B(\phi, X)\phi_\mu\phi_\nu, \quad \tilde{\phi} = \phi. \quad (3.1)$$

It was shown in Ref. [53] that there exists a unique inverse transformation of Eq. (3.1) if the functions A, B satisfy the condition $A(A - XA_X - X^2B_X) \neq 0$ [31, 43, 53] (see §3.3.1 and Appendix B for detail). One naturally expects that the number of physical DOFs is not changed by such an invertible transformation because there is a one-to-one correspondence between the old and new sets of variables. On the other hand, since the disformal transformation contains derivatives of the scalar field, the EL equations derived from the transformed action contain higher-order derivatives and thus the equivalence between the two frames is not clear. There are some works that addressed this issue. In the special case where the original action is of the Einstein-Hilbert form, the authors of Refs. [53, 68] showed that the EL equations in the new frame containing higher-order

derivatives can be reduced to second-order differential equations by taking their linear combinations. The disformal invariance of cosmological perturbations and their number of DOFs were investigated in Refs. [69–72]. It was clarified in Ref. [73] that the disformal transformation in a cosmological setup amounts to a rescaling of time coordinate and thus leaves physical observables unchanged. The authors of Ref. [37] proved the equivalence between two sets of EL equations for disformally related frames for an arbitrary scalar-tensor theory. The equivalence between the old and new frames has also been confirmed by Hamiltonian analysis in the unitary gauge $\phi = t$ [74], though the similar analysis without gauge fixing remains unaddressed. It was also shown in Ref. [74] that the Hamiltonian structure is unchanged under a broad class of invertible field transformations. However, there exist infinitely many types of invertible transformations that are not covered by their analysis. Also, the equivalence between EL equations for two frames related through general transformations has not been clarified. These facts motivate us to explore the nature of generic invertible transformations that depend on fields and their derivatives.

In light of this situation, in the present chapter, we show the following proposition (hereafter called “Theorem”) on invertible transformations:

If two frames are related by a general invertible transformation, the EL equations in the new frame are completely equivalent to the original-frame EL equations written in terms of the new fields.

In other words, the new-frame EL equations are derived from the original-frame EL equations without any loss/gain of information, and vice versa. For more precise statement and its proof, see §3.2.2. Combining this result with the property of invertible transformations that the fields in the two frames are related by a one-to-one correspondence, it can be concluded that any solution of the EL equations in the original frame is mapped to a solution in the new frame by the invertible transformation. The application of the Theorem is not restricted to scalar-tensor theories, but rather extends to any field theory.

There is a link between this chapter and Chap. 4. As mentioned in §2.1.3, one can obtain healthy theories beyond Horndeski by performing a disformal transformation on the Horndeski theory. However, the known broadest framework of quadratic/cubic DHOST theories cannot further be extended in a similar manner because it is closed under disformal transformations. Therefore, to go beyond the existing broadest framework by using disformal transformations, we have to consider another possibility. A possible way is to perform a disformal transformation on a *nondegenerate* scalar-tensor theory, but the above Theorem implies that the resultant theory is also nondegenerate as long as the transformation is invertible. Hence, in order to generate a healthy scalar-tensor theory from a nondegenerate theory, one has to perform *noninvertible* disformal transformation. Indeed, by doing so we can obtain a new class of degenerate scalar-tensor theories, as we will see in Chap. 4.

The rest of this chapter is organized as follows. In §3.1, we provide two examples

which illustrate the role of (derivative-dependent) invertible transformations. Then in §3.2, we prove the aforementioned Theorem to clarify the relation between invertible transformations and EL equations. Furthermore, we present applications of our Theorem to scalar-tensor theories in §3.3, which include the class of disformal transformations mentioned above.

3.1 Examples

Before proceeding to general arguments in field theories, we give two examples which are useful to get a flavor of the Theorem.

3.1.1 Analytical mechanics

First, we consider a simple model in analytical mechanics. Let us start from the Lagrangian

$$\tilde{L}(\dot{X}, \dot{Y}) = \frac{1}{2}\dot{X}^2 + \frac{1}{2}\dot{Y}^2. \quad (3.2)$$

As is obvious, the EOMs obtained from this Lagrangian

$$\mathcal{E}_X \equiv -\ddot{X} = 0, \quad \mathcal{E}_Y \equiv -\ddot{Y} = 0, \quad (3.3)$$

are a pair of second-order ordinary differential equations, and thus we need four initial conditions, i.e., the system has two DOFs. Now we perform a derivative-dependent frame transformation with

$$X = x - \dot{y}, \quad Y = y. \quad (3.4)$$

Note that this transformation is invertible: It can be uniquely solved for x, y as

$$x = X + \dot{Y}, \quad y = Y. \quad (3.5)$$

Since X has \dot{y} in its transformation rule, the new Lagrangian contains a higher-order time derivative:

$$L(\dot{x}, \dot{y}, \ddot{y}) = \frac{1}{2}(\dot{x} - \ddot{y})^2 + \frac{1}{2}\dot{y}^2, \quad (3.6)$$

and so do the EOMs:

$$\mathcal{E}_x \equiv -\ddot{x} + \dot{y}^{(3)} = 0, \quad \mathcal{E}_y \equiv -\ddot{y} - x^{(3)} + y^{(4)} = 0. \quad (3.7)$$

At a first glance, this new system of equations seems to require more initial conditions than Eq. (3.3), but this is not true. Indeed, one can eliminate the higher derivative terms by taking linear combinations of the EOMs together with their time derivatives:

$$\begin{aligned} \mathcal{E}_x + \dot{\mathcal{E}}_y - \ddot{\mathcal{E}}_x &= -\ddot{x} = 0, \\ \mathcal{E}_y - \dot{\mathcal{E}}_x &= -\ddot{y} = 0. \end{aligned} \quad (3.8)$$

This system of equations has the same structure as the original one (3.3). Therefore, we need the same number of initial conditions to fix the dynamics of x, y as in Eq. (3.3). The above equivalence between the two frames can also be understood as follows. Written in terms of the original set of variables (X, Y) , the left-hand sides of Eq. (3.7) become

$$\mathcal{E}_x = -\ddot{X}, \quad \mathcal{E}_y = -\ddot{Y} - X^{(3)}. \quad (3.9)$$

Then, they are combined to give \mathcal{E}_X and \mathcal{E}_Y , i.e., the original set of EOMs (3.3), as

$$\mathcal{E}_X = \mathcal{E}_x, \quad \mathcal{E}_Y = \mathcal{E}_y - \dot{\mathcal{E}}_x, \quad (3.10)$$

while $(\mathcal{E}_x, \mathcal{E}_y)$ is expressed in terms of $(\mathcal{E}_X, \mathcal{E}_Y)$ as

$$\mathcal{E}_x = \mathcal{E}_X, \quad \mathcal{E}_y = \mathcal{E}_Y + \dot{\mathcal{E}}_X. \quad (3.11)$$

Equations (3.10) and (3.11) imply that the new-frame EOMs written in terms of the old variables are completely equivalent to the old-frame EOMs. Hence, any solution in the old frame (X, Y) is mapped to a solution in the new frame (x, y) and vice versa, meaning that the two theories (3.2) and (3.6) have a common number of physical DOFs. Note also the similarity between Eqs. (3.4), (3.5) and Eqs. (3.10), (3.11). We shall clarify the origin of the similarity in §3.2.

One may notice that the transformation (3.4) basically captures the essential nature of the disformal transformation (1.4): X and Y loosely correspond to $g_{\mu\nu}$ and ϕ , respectively. The crucial difference between them is that the disformal transformation is more complicated so that it is not always invertible. When it is invertible, the logic is the same as the above discussion. On the other hand, a noninvertible disformal transformation generically maps a scalar-tensor theory to one with a different number of DOFs (see Chap. 4).

3.1.2 Scalar-tensor theory

The second example is the case of scalar-tensor theory. For the Einstein-Hilbert action with a canonical scalar field $\tilde{\phi}$ and some matter fields Ψ^I ,

$$\tilde{S}[\tilde{g}_{\mu\nu}, \tilde{\phi}; \Psi^I] = \int d^4x \sqrt{-\tilde{g}} \left[\frac{M_{\text{Pl}}^2}{2} \tilde{\mathcal{R}} - \frac{1}{2} \tilde{\nabla}_\mu \tilde{\phi} \tilde{\nabla}^\mu \tilde{\phi} - V(\tilde{\phi}) \right] + S_m[\tilde{g}_{\mu\nu}; \Psi^I], \quad (3.12)$$

let us consider the following transformation:

$$\tilde{g}_{\mu\nu} = g_{\mu\nu}, \quad \tilde{\phi} = \phi - f(\mathcal{R}), \quad (3.13)$$

where $f(\mathcal{R})$ is an arbitrary function of the Ricci scalar associated with the metric $g_{\mu\nu}$. The inverse transformation is given by

$$g_{\mu\nu} = \tilde{g}_{\mu\nu}, \quad \phi = \tilde{\phi} + f(\tilde{\mathcal{R}}), \quad (3.14)$$

where now $\tilde{\mathcal{R}}$ is computed from $\tilde{g}_{\mu\nu}$. For this transformation, the original action (3.12) is transformed as

$$S[g_{\mu\nu}, \phi; \Psi^I] = \int d^4x \sqrt{-g} \left[\frac{M_{\text{Pl}}^2}{2} \mathcal{R} - \frac{1}{2} \nabla_\mu (\phi - f(\mathcal{R})) \nabla^\mu (\phi - f(\mathcal{R})) - V(\phi - f(\mathcal{R})) \right] + S_m[g_{\mu\nu}; \Psi^I]. \quad (3.15)$$

Introducing a Lagrange multiplier, one can recast this S into the form of

$$S'[g_{\mu\nu}, \phi, \chi, \lambda; \Psi^I] = \int d^4x \sqrt{-g} \left[\frac{M_{\text{Pl}}^2}{2} \mathcal{R} - \frac{1}{2} \nabla_\mu \varphi \nabla^\mu \varphi - V(\varphi) + \lambda(\chi - \mathcal{R}) \right] + S_m[g_{\mu\nu}; \Psi^I], \quad \varphi \equiv \phi - f(\chi), \quad (3.16)$$

which manifestly yields second-order field equations. The action (3.16), which contains three scalar fields ϕ , χ , and λ , describes a specific model of tensor-multiscalar theory defined in Ref. [75]. Although such a theory has $2 + 3$ DOFs in general, the specific theory defined by S' is expected to have only $2 + 1$ DOFs as it is obtained via the invertible transformation (3.13) from the action (3.12) containing only one scalar field. Actually, we can explicitly show that the EL equations derived from S' are completely equivalent to those derived from the original action \tilde{S} in the following way. The EOMs obtained from S' are

$$E_{\mu\nu} \equiv \frac{1}{\sqrt{-g}} \frac{\delta S'}{\delta g^{\mu\nu}} = 0, \quad E_\Phi \equiv \frac{1}{\sqrt{-g}} \frac{\delta S'}{\delta \Phi} = 0, \quad (\Phi = \phi, \chi, \lambda), \quad (3.17)$$

where

$$E_{\mu\nu} = \frac{M_{\text{Pl}}^2}{2} \mathcal{G}_{\mu\nu} + \frac{1}{2} g_{\mu\nu} \left[\frac{1}{2} \nabla_\sigma \varphi \nabla^\sigma \varphi + V(\varphi) - \lambda(\chi - \mathcal{R}) \right] - \frac{1}{2} \nabla_\mu \varphi \nabla_\nu \varphi + \nabla_\mu \nabla_\nu \lambda - g_{\mu\nu} \square \lambda - \lambda \mathcal{R}_{\mu\nu} - \frac{1}{2} T_{\mu\nu}, \quad (3.18)$$

$$E_\phi = \square \varphi - V'(\varphi), \quad (3.19)$$

$$E_\chi = \lambda - f'(\chi) [\square \varphi - V'(\varphi)], \quad (3.20)$$

$$E_\lambda = \chi - \mathcal{R}, \quad (3.21)$$

with $T_{\mu\nu} \equiv -\frac{2}{\sqrt{-g}} \frac{\delta S_m}{\delta g^{\mu\nu}}$ being the energy-momentum tensor for the matter fields. The EOM for the Lagrange multiplier $E_\lambda = 0$ implies $\chi = \mathcal{R}$. Combining $E_\phi = 0$ and $E_\chi = 0$, one obtains $\lambda = 0$. Thus, the metric EOM $E_{\mu\nu} = 0$ is written as

$$\frac{M_{\text{Pl}}^2}{2} \mathcal{G}_{\mu\nu} + \frac{1}{2} g_{\mu\nu} \left[\frac{1}{2} \nabla_\sigma \varphi \nabla^\sigma \varphi + V(\varphi) \right] - \frac{1}{2} \nabla_\mu \varphi \nabla_\nu \varphi - \frac{1}{2} T_{\mu\nu} = 0. \quad (3.22)$$

This equation and $E_\phi = 0$ are nothing but the Einstein and Klein-Gordon equations derived from the original action (3.12) with the replacement $\phi \rightarrow \varphi$. Hence, for any

solution $(g_{\mu\nu}, \phi) = (g_{\mu\nu}^{(0)}, \phi^{(0)})$ of the original EOMs, the set $(g_{\mu\nu}, \varphi) = (g_{\mu\nu}^{(0)}, \phi^{(0)})$ satisfies the new-frame EOMs. Once $(g_{\mu\nu}, \varphi)$ is fixed, the solution for the new-frame variables $(g_{\mu\nu}, \phi, \chi, \lambda)$ is also fixed as follows:

$$g_{\mu\nu} = g_{\mu\nu}^{(0)}, \quad \phi = \phi^{(0)} + f(\mathcal{R}^{(0)}), \quad \chi = \mathcal{R}^{(0)}, \quad \lambda = 0, \quad (3.23)$$

where $\mathcal{R}^{(0)}$ denotes the Ricci scalar associated with $g_{\mu\nu}^{(0)}$. Therefore, the system of EOMs (3.17) is essentially the system composed of the Einstein equation (3.22) and the Klein-Gordon equation $E_\phi = 0$, and thus has the same number of DOFs as the original system.

What we can learn from these simple examples is that, even if we perform a derivative-dependent transformation to obtain a Lagrangian with higher derivatives, it has the same number of DOFs as the original one as long as the transformation is invertible. In the subsequent section, we prove this statement for general field theories in arbitrary dimensions.

3.2 Proof of the Theorem

3.2.1 Setup

Let us consider a general field theory in D -dimensional spacetime:

$$S = \int d^D x L[\phi], \quad (3.24)$$

$$L[\phi] \equiv L(\phi^i, \partial_\mu \phi^i, \partial_\mu \partial_\nu \phi^i, \dots, \partial_{(n)} \phi^i),$$

where $i = 1, \dots, N$ labels the fields and $\partial_{(k)} \equiv \partial_{\mu_1} \dots \partial_{\mu_k}$. Transforming ϕ^i to a new set of fields ψ^i by^{*1}

$$\phi^i = f^i[\psi] \equiv f^i(\psi^j, \partial_\mu \psi^j, \partial_\mu \partial_\nu \psi^j, \dots, \partial_{(m)} \psi^j), \quad (3.25)$$

we obtain a new theory, symbolically written as

$$L'[\psi] \equiv L[f[\psi]], \quad (3.26)$$

which consists of at most $(m + n)$ th derivatives of ψ^i . It should be noted that the transformation (3.25) depends only on ψ^i and their derivatives evaluated at the same point in the spacetime. A field transformation between ϕ^i and ψ^i is called *invertible* if ψ^i are uniquely determined from ϕ^i and vice versa. As such, Eq. (3.25) can be solved for ψ^i in the form of

$$\psi^i = g^i(\phi^j, \partial_\mu \phi^j, \partial_\mu \partial_\nu \phi^j, \dots, \partial_{(\ell)} \phi^j). \quad (3.27)$$

^{*1}Although in some context it is more natural to begin with the expression of the new variables in terms of the old ones, i.e., in the form of $\psi^i = g^i[\phi]$, we instead start from Eq. (3.25) which is more convenient for the present purpose.

Hereafter we require that the number of ϕ fields be the same as that of ψ fields, because otherwise one cannot define an invertible transformation between ϕ^i and ψ^i . We also require that the dynamics of ϕ^i is restricted within the codomain of f^i .

In general, the transformation law (3.25) could be quite nonlinear. However, as we shall see below, if the invertibility is considered only locally in field space, the invertibility of the transformation can be judged within the language of linear algebra. Let us consider infinitesimal changes $\delta\phi^i, \delta\psi^i$ from configurations of ϕ^i, ψ^i that satisfy the relation (3.25). Then Eq. (3.25) is linearized as

$$\delta\phi^i = \hat{P}_j^i \delta\psi^j, \quad (3.28)$$

where \hat{P}_j^i is a derivative-operator-valued matrix determined from the functional form of f^i :

$$\hat{P}_j^i = \sum_{s=0}^m u_j^{i(s)} \partial_{(s)}, \quad u_j^{i(s)} \equiv \frac{\partial f^i}{\partial (\partial_{(s)} \psi^j)}. \quad (3.29)$$

A system of equations of the form (3.28) is called linear differential-algebraic equations (DAEs), since it consists of coupled linear differential and algebraic equations. The solution to Eq. (3.28) is generically not unique as it may contain integration constants. On the other hand, if the transformation (3.25) is invertible at least locally, then one can uniquely solve the system of DAEs (3.28) for $\delta\psi^i$ in the form of

$$\delta\psi^i = \hat{Q}_j^i \delta\phi^j, \quad (3.30)$$

where \hat{Q}_j^i is a derivative-operator-valued matrix satisfying^{*2}

$$\hat{P}_j^i \hat{Q}_k^j = \hat{Q}_j^i \hat{P}_k^j = \delta_k^i, \quad (3.31)$$

and hence plays the role of the inverse operator of \hat{P}_j^i . In the present chapter, we restrict ourselves to such a special class of field transformations.

3.2.2 The Theorem

In §3.1, we saw that an invertible transformation does not change the number of physical DOFs in two simple models. Below we prove the following Theorem for general field theories:

Theorem. *Suppose two sets of fields ϕ^i and ψ^i are related by an invertible transformation of the form $\phi^i = f^i[\psi]$. If a configuration $\psi_{(0)}^i$ satisfies the EL equations for ψ^i , then its transformation $f^i[\psi_{(0)}]$ satisfies the EL equations for ϕ^i . Conversely, if a configuration $\phi_{(0)}^i$ satisfies the EL equations for ϕ^i , then its inverse transformation $(f^{-1})^i[\phi_{(0)}]$ satisfies the EL equations for ψ^i .*^{*3}

^{*2}If there exists a derivative-operator-valued matrix \hat{Q}_j^i for which $\hat{Q}_j^i \hat{P}_k^j = \delta_k^i$, one can prove $\hat{P}_j^i \hat{Q}_k^j = \delta_k^i$ and the uniqueness of such \hat{Q}_j^i in the same manner as the case of c -number matrices.

^{*3}The symbol $(f^{-1})^i[\phi]$ stands for the configuration of ψ^i that satisfies $\phi^i = f^i[\psi]$.

Proof. Let us consider the variation of the action in two different manners. If we vary the original action written in terms of ϕ^i , then we obtain

$$\delta S = \delta \int d^D x L[\phi] = \int d^D x \mathcal{E}_i^{(\phi)} \delta \phi^i. \quad (3.32)$$

Here, $\mathcal{E}_i^{(\phi)} = 0$ denote the EOMs for ϕ^i . Meanwhile, if we rewrite the action in terms of ψ^i by the relation (3.25), the variation yields

$$\delta S = \delta \int d^D x L'[\psi] = \int d^D x \mathcal{E}_i^{(\psi)} \delta \psi^i, \quad (3.33)$$

where $\mathcal{E}_i^{(\psi)} = 0$ are the EOMs for ψ^i . Note that, in deriving the EL equations, we have imposed independent boundary conditions for ϕ^i and ψ^i : $\partial_{(k)} \phi^i = 0$ ($k = 0, 1, \dots, n-1$) and $\partial_{(k')} \psi^i = 0$ ($k' = 0, 1, \dots, n+m-1$), respectively. This is because we need only the relation between the old- and new-frame EL equations obtained in such manner. Now we impose Eq. (3.28) on $\delta \phi^i$ and reexpress Eq. (3.32) by $\delta \psi^i$:

$$\delta S = \int d^D x \mathcal{E}_i^{(\phi)} \hat{P}_j^i \delta \psi^j = \int d^D x \left(\hat{P}_j^{\dagger i} \mathcal{E}_i^{(\phi)} \right) \delta \psi^j. \quad (3.34)$$

Hereafter ϕ^i and ψ^i are freely replaced with each other via the relation (3.25). Note that, as a result of integration by parts, here we have the adjoint of \hat{P}_j^i which is defined so that

$$\hat{P}_j^{\dagger i} w_i = \sum_{s=0}^m (-1)^s \partial_{(s)} \left(u_j^{i(s)} w_i \right), \quad (3.35)$$

for any vector function w_j . After this, one can compare Eqs. (3.33) and (3.34). Since $\delta \psi^i$ are arbitrary, one obtains the following relation between $\mathcal{E}_i^{(\phi)}$ and $\mathcal{E}_i^{(\psi)}$:

$$\mathcal{E}_i^{(\psi)} = \hat{P}_i^{\dagger j} \mathcal{E}_j^{(\phi)}. \quad (3.36)$$

It should be noted that this relation itself holds even if the transformation is not invertible.

The relation (3.36) can be regarded as the adjoint DAE system to Eq. (3.28). Now the problem is whether the original set of EOMs $\mathcal{E}_i^{(\phi)} = 0$ follows from $\hat{P}_j^{\dagger i} \mathcal{E}_j^{(\phi)} = 0$. To prove this, one can follow the arguments on the unique solvability of adjoint DAEs given in Ref. [63].^{*4} Since the operator matrix \hat{P}_j^i has its inverse \hat{Q}_j^i without integral operator, one can take the adjoint of Eq. (3.31) to obtain

$$\hat{P}_j^{\dagger i} \hat{Q}_k^{\dagger j} = \hat{Q}_j^{\dagger i} \hat{P}_k^{\dagger j} = \delta_k^i, \quad (3.37)$$

^{*4}In Ref. [63], we used the same technique to prove the following theorem on the relation between gauge fixing and EL equations: *In any gauge theory, if a gauge fixing is complete, i.e., the gauge functions are determined uniquely by the gauge conditions, the EL equations derived from the gauge-fixed action are equivalent to those derived from the original action supplemented with the gauge conditions.*

which means that the inverse operator of $\hat{P}_j^{\dagger i}$ is independent of integral operators and given by $\hat{Q}_j^{\dagger i}$. In other words, if a DAE system is uniquely solvable, so is its adjoint DAE system. Therefore, from Eq. (3.36) we obtain

$$\mathcal{E}_i^{(\phi)} = \hat{Q}_i^{\dagger j} \mathcal{E}_j^{(\psi)}, \quad (3.38)$$

which is the adjoint DAE system to Eq. (3.30). Multiplying both sides of $\hat{P}_i^{\dagger j} \mathcal{E}_j^{(\phi)} = 0$ by $\hat{Q}_k^{\dagger i}$ yields $\mathcal{E}_k^{(\phi)} = 0$. Hence, if a configuration of ψ^i that satisfies the new set of EOMs (3.40) is transformed by the relation (3.25), then the resulting configuration of ϕ^i satisfies the original set of EOMs (3.39). Moreover, the opposite direction is also true. This completes the proof of the Theorem. \square

In deriving the relation (3.36), we have not used explicit expressions for the EL equations in each frame. Although technically more complicated, it is also possible to show the relation by a direct comparison between the explicit expression of $\mathcal{E}_i^{(\phi)}$ and that of $\mathcal{E}_i^{(\psi)}$ as follows. The EOMs for ϕ^i are formally written as

$$\mathcal{E}_i^{(\phi)} \equiv \frac{\delta L[\phi]}{\delta \phi^i} = \sum_{q=0}^n (-1)^q \partial_{(q)} v_i^{(q)} = 0, \quad v_i^{(q)} \equiv \frac{\partial L}{\partial (\partial_{(q)} \phi^i)}. \quad (3.39)$$

On the other hand, the EOMs for ψ^i become

$$\mathcal{E}_i^{(\psi)} \equiv \frac{\delta L[f[\psi]]}{\delta \psi^i} = \sum_{p=0}^{m+n} \sum_{q=0}^n (-1)^p \partial_{(p)} \left[v_j^{(q)} \frac{\partial (\partial_{(q)} f^j)}{\partial (\partial_{(p)} \psi^i)} \right] = 0. \quad (3.40)$$

Note that $\mathcal{E}_i^{(\psi)} = 0$ is different from what one obtains by substituting Eq. (3.25) into $\mathcal{E}_i^{(\phi)} = 0$, though these two sets of equations are equivalent by virtue of the relation (3.36). By using the relation^{*5}

$$\frac{\partial (\partial_{(q)} f^j)}{\partial (\partial_{(p)} \psi^i)} = \sum_{\substack{0 \leq k \leq q \\ 0 \leq p-k \leq m}} \binom{q}{k} \partial_{(q-k)} u_i^{j(p-k)}, \quad (3.41)$$

the expression of $\mathcal{E}_i^{(\psi)}$ becomes

$$\mathcal{E}_i^{(\psi)} = \sum_{k=0}^n \sum_{p=k}^{k+m} \sum_{q=k}^n (-1)^p \binom{q}{k} \partial_{(p)} \left[v_j^{(q)} \partial_{(q-k)} u_i^{j(p-k)} \right], \quad (3.42)$$

^{*5}Equation (3.41) can be verified by repeated use of the following identity for $\Phi[\psi] = \partial_{(r)} f^j[\psi]$ ($r = 0, 1, \dots, q-1$):

$$\frac{\partial (\partial_{(1)} \Phi[\psi])}{\partial (\partial_{(p)} \psi^i)} = \partial_{(1)} \frac{\partial \Phi[\psi]}{\partial (\partial_{(p)} \psi^i)} + \frac{\partial \Phi[\psi]}{\partial (\partial_{(p-1)} \psi^i)},$$

which can be checked by expanding the both sides using the chain rule.

where we have interchanged the summations. With the aid of the Leibniz rule,

$$\begin{aligned}\mathcal{E}_i^{(\psi)} &= \sum_{k=0}^n \sum_{p=k}^{k+m} \sum_{q=k}^n \sum_{r=0}^k (-1)^p \binom{q}{k} \binom{k}{r} \partial_{(p-k)} \left[\partial_{(r)} v_j^{(q)} \partial_{(q-r)} u^{j(p-k)} \right] \\ &= \sum_{s=0}^m \sum_{k=0}^n \sum_{q=k}^n \sum_{r=0}^k (-1)^{k+s} \binom{q}{k} \binom{k}{r} \partial_{(s)} \left[\partial_{(r)} v_j^{(q)} \partial_{(q-r)} u^{j(s)} \right],\end{aligned}\quad (3.43)$$

where we have defined $s \equiv p - k$. Interchanging the summations as $\sum_{k=0}^n \sum_{q=k}^n \sum_{r=0}^k = \sum_{q=0}^n \sum_{r=0}^q \sum_{k=r}^q$ and using the formula

$$\sum_{k=r}^q (-1)^k \binom{q}{k} \binom{k}{r} = (-1)^q \delta_{qr},\quad (3.44)$$

we finally obtain

$$\mathcal{E}_i^{(\psi)} = \sum_{s=0}^m \sum_{q=0}^n (-1)^{s+q} \partial_{(s)} \left[u_i^{j(s)} \partial_{(q)} v_j^{(q)} \right] = \hat{P}_i^{\dagger j} \mathcal{E}_j^{(\phi)},\quad (3.45)$$

which is nothing but Eq. (3.36).

According to the Theorem, if one can define an invertible transformation between ϕ^i and ψ^i , then the solution space for ψ^i is mapped to a subspace of the solution space for ϕ^i , and vice versa. Therefore, the two solution spaces have the same number of DOFs.

3.2.3 Remarks

We showed in the previous section that one can recover the original EOMs $\mathcal{E}_i^{(\phi)} = 0$ from the new EOMs $\hat{P}_j^{\dagger i} \mathcal{E}_i^{(\phi)} = 0$ if the field transformation is invertible. Here, $\hat{P}_j^{\dagger i}$ contains derivative operators arising from derivatives in the field transformation, which is the origin of the nontriviality when proving the equivalence between the original- and new-frame EOMs. To circumvent this problem, one may want to reduce the derivative-dependent transformation to a transformation without field derivatives by introducing auxiliary fields and Lagrange multipliers. Naively, such a transformation allows us to obtain the new-frame EOMs in a more concise form and facilitates the proof of the equivalence between $\mathcal{E}_i^{(\phi)} = 0$ and $\hat{P}_j^{\dagger i} \mathcal{E}_i^{(\phi)} = 0$. However, it is actually not the case for the following reasons.

Let us go back to the Lagrangian (3.24) and the derivative-dependent transformation (3.27). To remove derivatives from Eq. (3.27), we introduce auxiliary fields $\chi_{(s)}^i \equiv \chi_{\mu_1 \dots \mu_s}^i$ with Lagrange multipliers $\lambda_i^{(s)} \equiv \lambda_i^{\mu_1 \dots \mu_s}$ and obtain the modified Lagrangian as

$$\tilde{L} \equiv L(\phi^i, \partial_\mu \phi^i, \dots, \partial_{(n)} \phi^i) + \sum_{s=1}^{\ell} \lambda_i^{(s)} (\chi_{(s)}^i - \partial_{(1)} \chi_{(s-1)}^i).\quad (3.46)$$

Here, ℓ denotes the highest order of derivative in the transformation (3.27), and $\chi_{(0)}^i$ is understood as ϕ^i . The EL equations are

$$\tilde{\mathcal{E}}_i^{(\phi)} \equiv \mathcal{E}_i^{(\phi)} + \partial_\mu \lambda_i^\mu = 0, \quad (3.47)$$

$$\tilde{\mathcal{E}}_{\chi_{(s)}^i} \equiv \lambda_i^{(s)} + \partial_{(1)} \lambda_i^{(s+1)} = 0, \quad (3.48)$$

$$\tilde{\mathcal{E}}_{\lambda_i^{(s)}} \equiv \chi_{(s)}^i - \partial_{(1)} \chi_{(s-1)}^i = 0, \quad (3.49)$$

where $\lambda_i^{(\ell+1)} \equiv 0$. Equations (3.48) and (3.49) respectively yield $\lambda_i^{(s)} = 0$ and $\chi_{(s)}^i = \partial_{(s)} \phi^i$, and thus we obtain $\mathcal{E}_i^{(\phi)} = 0$ from Eq. (3.47). Now we formally replace the derivatives contained in the field transformation (3.27) by $\chi_{(s)}^i$, namely,

$$\psi^i = g^i(\phi^j, \chi_\mu^j, \dots, \chi_{(\ell)}^j). \quad (3.50)$$

Assuming that $\chi_{(s)}^i$ and $\lambda_i^{(s)}$ remain unchanged when transformed into the new frame, Eq. (3.50) defines an invertible transformation between extended field sets $(\phi^i, \chi_{(s)}^i, \lambda_i^{(s)})$ and $(\psi^i, \chi_{(s)}^i, \lambda_i^{(s)})$ without field derivatives. This is because the determinant of $J_j^i \equiv \partial g^i / \partial \phi^j$ is nonvanishing due to the invertibility of the field transformation (3.27).^{*6} One may thus expect that

- (i) the relation between the old- and new-frame EOMs becomes clearer than considering the derivative-dependent transformation (3.27), and
- (ii) it would alleviate the proof of the equivalence between $\mathcal{E}_i^{(\phi)} = 0$ and $\hat{P}^{\dagger i} \mathcal{E}_i^{(\phi)} = 0$.

Indeed, (i) is the case. If we perform the transformation (3.50) on the modified Lagrangian (3.46), the variation of the action becomes as

$$\begin{aligned} \delta \tilde{S} &= \int d^D x \delta \tilde{L} = \int d^D x \left(\tilde{\mathcal{E}}_i^{(\phi)} \delta \phi^i + \tilde{\mathcal{E}}_{\chi_{(s)}^i} \delta \chi_{(s)}^i + \tilde{\mathcal{E}}_{\lambda_i^{(s)}} \delta \lambda_i^{(s)} \right) \\ &= \int d^D x \left[\tilde{\mathcal{E}}_i^{(\phi)} (J^{-1})_j^i \delta \psi^j + \left(\tilde{\mathcal{E}}_{\chi_{(s)}^k} + \tilde{\mathcal{E}}_i^{(\phi)} (J^{-1})_j^i \frac{\partial g^j}{\partial \chi_{(s)}^k} \right) \delta \chi_{(s)}^k + \tilde{\mathcal{E}}_{\lambda_i^{(s)}} \delta \lambda_i^{(s)} \right]. \end{aligned} \quad (3.51)$$

Then, the resulting EL equations are

$$(J^{-1})_j^i \tilde{\mathcal{E}}_i^{(\phi)} = 0, \quad \tilde{\mathcal{E}}_{\chi_{(s)}^k} + (J^{-1})_j^i \frac{\partial g^j}{\partial \chi_{(s)}^k} \tilde{\mathcal{E}}_i^{(\phi)} = 0, \quad \tilde{\mathcal{E}}_{\lambda_i^{(s)}} = 0, \quad (3.52)$$

^{*6}The condition $\det J_j^i \neq 0$ is only a necessary and not a sufficient condition for the field transformation to be invertible. For details, see Appendix C.

which are obviously equivalent to Eqs. (3.47)-(3.49), and thus the original EOMs $\mathcal{E}_i^{(\phi)} = 0$. However, (ii) is not the case since Eq. (3.52) does not address the equivalence between $\mathcal{E}_i^{(\phi)} = 0$ and $\hat{P}_j^{\dagger i} \mathcal{E}_i^{(\phi)} = 0$. Therefore, the idea of removing field derivatives from the field transformation does not lead to a simpler proof.

There are several other remarks on the Theorem.

- If a given transformation law $\phi^i = f^i[\psi]$ can be solved for ψ^i without integration constant but with branches of solutions, one can still apply the Theorem by choosing any one of the branches. For instance, the transformation $\phi = \psi^2$ has two inverse transformations, $\psi = \pm\sqrt{\phi}$. In this case, $\hat{P} = 2\psi$ and $\hat{Q} = \pm 1/(2\sqrt{\phi})$. After choosing either of the branches of \hat{Q} , one can apply the Theorem.
- As mentioned in §3.2.2, the unique solvability of the DAE system (3.28) plays a key role in the proof of the Theorem. Unfortunately, in general, there is no systematic way to judge whether a given DAE system is uniquely solvable or not. However, for $N = 1$ or $D = 1$, there exists an algorithm to do this. For the detailed arguments, see Appendix C.
- The original- and new-frame EL equations ($\mathcal{E}_i^{(\phi)} = 0$ and $\mathcal{E}_i^{(\psi)} = 0$) are at most $(2n)$ th- and $(2n + 2m)$ th-order differential equations, respectively. We have seen in Eq. (3.38) that the latter can be reduced to the equivalent $(2n + m)$ th-order differential equations by operating the regular matrix $\hat{Q}_j^{\dagger i}$. Since the Theorem states that the DOFs of the reduced equations are the same as the ones of the original-frame EL equations, it is natural to expect that the reduced equations can be further reduced to the manifestly equivalent $(2n)$ th-order differential equations by some manipulations. While the total number of required initial conditions is the same in both frames, it does not mean EL equations in the new frame are reducible to the set of equations each of which has the same orders of derivatives as those for EL equations in the old frame. To see this, let us consider the following Lagrangian in analytical mechanics depending on (X, Y) :

$$\tilde{L} = \frac{1}{2}\dot{X}^2 + \frac{1}{2}\dot{Y}^2 + XY. \quad (3.53)$$

Clearly, the EOMs for (X, Y) ,

$$\mathcal{E}_X \equiv -\ddot{X} + Y = 0, \quad \mathcal{E}_Y \equiv -\ddot{Y} + X = 0, \quad (3.54)$$

are two second-order differential equations, which require four initial conditions. Now we perform a transformation of variables $(X, Y) \rightarrow (x, y)$ defined by

$$x = \mathcal{E}_Y = X - \ddot{Y}, \quad y = Y, \quad (3.55)$$

which has the inverse transformation

$$X = x + \ddot{y}, \quad Y = y. \quad (3.56)$$

Note that $x = 0$ and is nondynamical by definition. The Lagrangian is then transformed as

$$L = \frac{1}{2}(\dot{x} + y^{(3)})^2 - \frac{1}{2}\dot{y}^2 + xy, \quad (3.57)$$

where we have performed integration by parts. The EOMs for (x, y) are given by

$$\mathcal{E}_x \equiv -\ddot{x} - y^{(4)} + y = 0, \quad \mathcal{E}_y \equiv -x^{(4)} - y^{(6)} + \ddot{y} + x = 0. \quad (3.58)$$

Due to our Theorem, the EOMs in the new frame are equivalent to those in the old frame through the relation (3.38) with the replacement (3.56). Indeed,

$$\begin{aligned} \mathcal{E}_X &= \mathcal{E}_x = -\ddot{x} - y^{(4)} + y = 0, \\ \mathcal{E}_Y &= \mathcal{E}_y - \ddot{\mathcal{E}}_x = x = 0. \end{aligned} \quad (3.59)$$

Hence, the EOMs for (x, y) are $x = 0$, which is consistent with Eq. (3.55), and a fourth-order differential equation $-y^{(4)} + y = 0$ obtained by substituting $x = 0$ into $\mathcal{E}_X = 0$. As such, the two second-order equations (3.54) are transformed into one zeroth-order equation and one fourth-order equation. Obviously, the new EOMs are not reducible to two second-order differential equations. Nevertheless, the two sets are still related through Eq. (3.36) or Eq. (3.38), and have the same number of DOFs. In this case, the EOMs in both frames indeed require four initial conditions. This example demonstrates that in general the old and new sets of EOMs have different derivative structures.^{*7} To clarify their structures, one has to investigate on a case-by-case basis.

- Another thing to note is that if the original theory has gauge symmetries, then Eq. (3.36) is not the only way to express $\mathcal{E}_i^{(\psi)}$ in terms of $\mathcal{E}_i^{(\phi)}$. This is because there exist identities among the EL equations corresponding to the gauge symmetries, i.e., Noether identities (see Appendix D for details). If the original theory is invariant under an infinitesimal gauge transformation in the form of

$$\Delta_\epsilon \phi^i = \hat{G}_I^i \epsilon^I, \quad (3.60)$$

where $I = 1, \dots, M$ labels the gauge symmetries, then the Noether identities are written as

$$\hat{G}_I^{\dagger i} \mathcal{E}_i^{(\phi)} = 0, \quad (3.61)$$

^{*7}This type of situation happens whenever one defines an invertible transformation in such a way that a part of the new fields becomes nondynamical. The above example has a problem that the Hamiltonian obtained from the Lagrangian (3.53) is not bounded below. Such a model was chosen just for simplicity in calculation.

which reduces the dimensionality of the old-frame EOM space by M . Note here that the number of the gauge symmetries M is smaller than that of the fields N . Correspondingly, the new system also has gauge symmetries under the infinitesimal transformation $\psi^i \rightarrow \psi^i + \Delta_\epsilon \psi^i$, where $\Delta_\epsilon \psi^i = \hat{Q}_j^i \hat{G}_I^j \epsilon^I$. Therefore, the new-frame EOMs satisfy the corresponding Noether identities $\hat{G}_I^{\dagger i} \hat{Q}_i^{\dagger j} \mathcal{E}_j^{(\psi)} = 0$, which means that the new-frame EOM space also has dimension $N - M$. Even in this case, the proof of the Theorem still holds since it relies only on the invertibility of $\hat{P}^{\dagger j}$. On the other hand, combining Eqs. (3.36) and (3.61), we obtain

$$\mathcal{E}_i^{(\psi)} = (\hat{P}_i^{\dagger j} + \hat{F}_i^I \hat{G}_I^{\dagger j}) \mathcal{E}_j^{(\phi)} \equiv \hat{R}_i^j \mathcal{E}_j^{(\phi)}, \quad (3.62)$$

with \hat{F}_i^I being an arbitrary derivative-operator-valued matrix. Note that this arbitrariness of the relation between the EOMs does not spoil the proof of the Theorem. For some choice of \hat{F}_i^I , the matrix \hat{R}_i^j may become singular, in which case \hat{R}_i^j is a projection operator onto the $(N - M)$ -dimensional constrained surface in the N -dimensional EOM space defined by the Noether identity (3.61). Nevertheless, the singularity is not problematic since it is only this constrained surface that is physically relevant.

Moreover, even if a given transformation is noninvertible and hence $\hat{P}_i^{\dagger j}$ is singular, there may exist a matrix \hat{F}_i^I such that \hat{R}_i^j is regular. If this is the case, one can prove the equivalence between $\mathcal{E}_i^{(\phi)} = 0$ and $\mathcal{E}_i^{(\psi)} = 0$, and thus the number of DOFs of the theory remains unchanged by such special noninvertible transformation.

- Before closing this section, let us remark that not all the variables relevant to an invertible transformation have to be dynamical, in which case however the transformed theory acquires redundant DOFs in general. We consider the following Lagrangian as an example:

$$\tilde{L}(\dot{X}) = \frac{1}{2} \dot{X}^2, \quad (3.63)$$

with the invertible field transformation of the same form as Eq. (3.4). Note that Y does not appear in the original Lagrangian (3.63). In this case, the new Lagrangian takes the form

$$L(\dot{x}, \dot{y}) = \frac{1}{2} (\dot{x} - \dot{y})^2, \quad (3.64)$$

which has a gauge symmetry under

$$x \rightarrow x + \dot{\xi}, \quad y \rightarrow y + \xi, \quad (3.65)$$

with ξ being an arbitrary function of time. Once the gauge is completely fixed by setting $y = 0$, we recover the original Lagrangian. In other words, introducing a gauge DOF y to the original theory defined by \tilde{L} is an invertible transformation,

whose inverse is fixing the gauge completely by setting $y = 0$ in the resultant new theory described by L , and vice versa.

This example is related to the Stückelberg formalism for a massive vector field. We start from the Proca Lagrangian

$$\tilde{L}_{\text{Proca}}(\tilde{A}_\mu, \partial_\lambda \tilde{A}_\mu) = -\frac{1}{4} \tilde{F}_{\mu\nu} \tilde{F}^{\mu\nu} + m^2 \tilde{A}_\mu \tilde{A}^\mu, \quad \tilde{F}_{\mu\nu} \equiv \partial_\mu \tilde{A}_\nu - \partial_\nu \tilde{A}_\mu. \quad (3.66)$$

One can restore $U(1)$ gauge symmetry via introducing a Stückelberg scalar ϕ by promoting

$$\tilde{A}_\mu \rightarrow A_\mu - \partial_\mu \phi, \quad (3.67)$$

and assuming the following gauge transformation law

$$A_\mu \rightarrow A_\mu + \partial_\mu \Lambda, \quad \phi \rightarrow \phi + \Lambda. \quad (3.68)$$

Indeed, the new Lagrangian

$$L_{\text{Proca}}(A_\mu, \phi, \partial_\lambda A_\mu, \partial_\lambda \phi) = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + m^2 (A_\mu - \partial_\mu \phi)(A^\mu - \partial^\mu \phi), \quad (3.69)$$

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu \quad (3.70)$$

is invariant under the transformation (3.68). In this case, the replacement (3.67) can be regarded as an invertible transformation by identifying it as the following field redefinition:

$$\tilde{A}_\mu = A_\mu - \partial_\mu \phi, \quad \tilde{\phi} = \phi. \quad (3.71)$$

The inverse transformation is given by

$$A_\mu = \tilde{A}_\mu + \partial_\mu \tilde{\phi}, \quad \phi = \tilde{\phi}. \quad (3.72)$$

Hence, as it should be, the Stückelberg formalism just introduces a redundant DOF and it does not change the number of physical DOFs. On the other hand, imposing a complete gauge fixing $\phi = 0$ in L_{Proca} can be identified as performing an invertible transformation (3.72) on L_{Proca} .^{*8}

Similarly, in the context of scalar-tensor theories, any additional scalar/vector/tensor fields can be introduced without changing the number of DOFs. This may be related to the work [58], which suggested a connection between tensor-multiscalar theories [75], generalized Proca theories [21], and bimetric theories [24].

^{*8}The Stückelberg trick for general covariance [54, 55] is not so simple as the case of $U(1)$ gauge symmetry (see §2.1.4). To the best of our knowledge, we cannot express it in the form of invertible transformation.

3.3 Applications to scalar-tensor theories

In the previous section, we have shown that the new-frame EL equations can be made equivalent to the original-frame EL equations by using the regular matrix $\hat{P}^{\dagger i}_j$. In this section, we consider two types of invertible transformations in the context of scalar-tensor theories and present explicit forms of the matrix $\hat{P}^{\dagger i}_j$.

3.3.1 Disformal transformation

Let us consider the disformal transformation

$$\tilde{g}_{\mu\nu} = A(\phi, X)g_{\mu\nu} + B(\phi, X)\nabla_{\mu}\phi\nabla_{\nu}\phi, \quad \tilde{\phi} = \phi. \quad (3.73)$$

One can define the inverse matrix of $\tilde{g}_{\mu\nu}$ as long as $A(A + XB) \neq 0$. Note that any composition of disformal transformations is again a disformal transformation. The necessary and sufficient condition for the invertibility of the disformal transformation is given by [53]

$$A(A - XA_X - X^2B_X) \neq 0, \quad (3.74)$$

which ensures the Jacobian determinant for the metric transformation is nonvanishing.^{*9} It is notable that the condition $A(A + XB) \neq 0$, which guarantees the existence of the inverse matrix $\tilde{g}^{\mu\nu}$, automatically follows from Eq. (3.74). In Chap. 4, we study noninvertible disformal transformations with $A - XA_X - X^2B_X = 0$ and $A(A + XB) \neq 0$.

If Eq. (3.74) is the case, the inverse disformal transformation is written as

$$g_{\mu\nu} = \tilde{A}(\tilde{\phi}, \tilde{X})\tilde{g}_{\mu\nu} + \tilde{B}(\tilde{\phi}, \tilde{X})\tilde{\nabla}_{\mu}\tilde{\phi}\tilde{\nabla}_{\nu}\tilde{\phi}, \quad \phi = \tilde{\phi}, \quad (3.75)$$

where $\tilde{\nabla}_{\mu}$ denotes a covariant derivative with respect to $\tilde{g}_{\mu\nu}$, and the canonical kinetic term of the scalar field in the original frame is related to the new variables by

$$\tilde{X} \equiv \tilde{g}^{\mu\nu}\tilde{\nabla}_{\mu}\tilde{\phi}\tilde{\nabla}_{\nu}\tilde{\phi} = \frac{X}{A + XB}. \quad (3.76)$$

The functional forms of \tilde{A}, \tilde{B} are given by the following relation:

$$\tilde{A}(\tilde{\phi}, \tilde{X}) = \frac{1}{A(\tilde{\phi}, X)}, \quad \tilde{B}(\tilde{\phi}, \tilde{X}) = -\frac{B(\tilde{\phi}, X)}{A(\tilde{\phi}, X)}, \quad (3.77)$$

where X should be written in terms of $(\tilde{\phi}, \tilde{X})$ by solving Eq. (3.76). As it should be, the solvability of Eq. (3.76) for X is guaranteed by the condition (3.74) as

$$\frac{\partial \tilde{X}}{\partial X} = \frac{A - XA_X - X^2B_X}{(A + XB)^2} \neq 0. \quad (3.78)$$

^{*9}The Jacobian for the metric transformation is nothing but $a_{\mu\nu}^{\alpha\beta}$ in Eq. (3.79) (see also Appendix B).

As stated in §2.1.3, the known classes of scalar-tensor theories are closed under disformal transformations. This fact explicitly shows that an invertible transformation does not change the number of physical DOFs. In what follows, we show this for an arbitrary scalar-tensor theory as an application of our Theorem. For the disformal transformation (3.73), the linearization yields

$$\begin{bmatrix} \delta \tilde{g}_{\mu\nu} \\ \delta \tilde{\phi} \end{bmatrix} = \hat{P} \begin{bmatrix} \delta g_{\alpha\beta} \\ \delta \phi \end{bmatrix}, \quad \hat{P} = \begin{bmatrix} a_{\mu\nu}^{\alpha\beta} & \hat{b}_{\mu\nu} \\ 0 & 1 \end{bmatrix}, \quad (3.79)$$

where

$$a_{\mu\nu}^{\alpha\beta} \equiv -(A_X g_{\mu\nu} + B_X \nabla_\mu \phi \nabla_\nu \phi) \nabla^\alpha \phi \nabla^\beta \phi + A \delta_{(\mu}^\alpha \delta_{\nu)}^\beta, \quad (3.80)$$

$$\hat{b}_{\mu\nu} \equiv (A_\phi g_{\mu\nu} + B_\phi \nabla_\mu \phi \nabla_\nu \phi) + 2 [B \delta_{(\mu}^\sigma \nabla_{\nu)} \phi + (A_X g_{\mu\nu} + B_X \nabla_\mu \phi \nabla_\nu \phi) \nabla^\sigma \phi] \nabla_\sigma. \quad (3.81)$$

Similarly, the inverse disformal transformation (3.75) is linearized in the form of

$$\begin{bmatrix} \delta g_{\mu\nu} \\ \delta \phi \end{bmatrix} = \hat{Q} \begin{bmatrix} \delta \tilde{g}_{\alpha\beta} \\ \delta \tilde{\phi} \end{bmatrix}, \quad \hat{Q} = \begin{bmatrix} c_{\mu\nu}^{\alpha\beta} & \hat{d}_{\mu\nu} \\ 0 & 1 \end{bmatrix}. \quad (3.82)$$

Here, the matrix elements $c_{\mu\nu}^{\alpha\beta}$ and $\hat{d}_{\mu\nu}$ can be written in terms of $(g_{\mu\nu}, \phi)$ as

$$c_{\mu\nu}^{\alpha\beta} \equiv \frac{1}{A(A - X A_X - X^2 B_X)} (A_X g_{\mu\nu} + B_X \nabla_\mu \phi \nabla_\nu \phi) \nabla^\alpha \phi \nabla^\beta \phi + \frac{1}{A} \delta_{(\mu}^\alpha \delta_{\nu)}^\beta, \quad (3.83)$$

$$\hat{d}_{\mu\nu} \equiv -c_{\mu\nu}^{\alpha\beta} \hat{b}_{\alpha\beta}. \quad (3.84)$$

As it should be, this \hat{Q} defines the inverse matrix of \hat{P} : One can check that

$$\begin{bmatrix} a_{\mu\nu}^{\rho\sigma} & \hat{b}_{\mu\nu} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} c_{\rho\sigma}^{\alpha\beta} & \hat{d}_{\rho\sigma} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} c_{\mu\nu}^{\rho\sigma} & \hat{d}_{\mu\nu} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{\rho\sigma}^{\alpha\beta} & \hat{b}_{\rho\sigma} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \delta_{(\mu}^\alpha \delta_{\nu)}^\beta & 0 \\ 0 & 1 \end{bmatrix}. \quad (3.85)$$

Now we confirm the equivalence between the old- and new-frame EOMs using the relation (3.36). Starting from a generic action $\tilde{S}[\tilde{g}_{\mu\nu}, \tilde{\phi}]$, the EOMs for the metric and the scalar field are derived as

$$\tilde{\mathcal{E}}^{\mu\nu} \equiv \frac{\delta \tilde{S}}{\delta \tilde{g}_{\mu\nu}} = 0, \quad \tilde{\mathcal{E}}_\phi \equiv \frac{\delta \tilde{S}}{\delta \tilde{\phi}} = 0. \quad (3.86)$$

On the other hand, if the action is written in terms of the new variables as $S[g_{\mu\nu}, \phi]$, the resulting EOMs are

$$\mathcal{E}^{\mu\nu} \equiv \frac{\delta S}{\delta g_{\mu\nu}} = 0, \quad \mathcal{E}_\phi \equiv \frac{\delta S}{\delta \phi} = 0. \quad (3.87)$$

Then the relation (3.36) reads

$$\begin{bmatrix} \mathcal{E}^{\alpha\beta} \\ \mathcal{E}_\phi \end{bmatrix} = \hat{P}^\dagger \begin{bmatrix} \tilde{\mathcal{E}}^{\mu\nu} \\ \tilde{\mathcal{E}}_\phi \end{bmatrix}, \quad \hat{P}^\dagger = \begin{bmatrix} a_{\mu\nu}^{\alpha\beta} & 0 \\ \hat{b}_{\mu\nu}^\dagger & 1 \end{bmatrix}, \quad (3.88)$$

where the scalar equation \mathcal{E}_ϕ acquires higher derivative terms due to the contribution $\hat{b}_{\mu\nu}^\dagger \tilde{\mathcal{E}}^{\mu\nu}$. However, Eq. (3.88) can be solved for the old-frame EOMs as

$$\begin{bmatrix} \tilde{\mathcal{E}}^{\alpha\beta} \\ \tilde{\mathcal{E}}_\phi \end{bmatrix} = \hat{Q}^\dagger \begin{bmatrix} \mathcal{E}^{\mu\nu} \\ \mathcal{E}_\phi \end{bmatrix}, \quad \hat{Q}^\dagger = \begin{bmatrix} c_{\mu\nu}^{\alpha\beta} & 0 \\ \hat{d}_{\mu\nu}^\dagger & 1 \end{bmatrix}, \quad (3.89)$$

which means that the lower-order EOMs (3.86) in the old frame can be recovered from the higher-order EOMs (3.87) in the new frame. The authors of Ref. [37] gave the same result based on a heuristic approach, but our method has an advantage that the equivalence between the old- and new-frame EOMs can be verified in a systematic manner.

3.3.2 Mixing with derivatives of the metric

Contrary to the case of disformal transformations where only the metric is nontrivially transformed, here we consider a nontrivial transformation of the scalar field, namely,

$$\tilde{g}_{\mu\nu} = g_{\mu\nu}, \quad \tilde{\phi} = F(\phi; g_{\mu\nu}, \partial_\lambda g_{\mu\nu}, \partial_\lambda \partial_\sigma g_{\mu\nu}, \dots), \quad (3.90)$$

where F is an arbitrary scalar quantity constructed without derivatives of ϕ . Note that the transformation (3.90) generalizes the transformation (3.13) in §3.1.2 and it introduces higher-order derivatives of the metric on a given scalar-tensor theory. If $\partial F / \partial \phi \neq 0$, one can solve $F(\phi) = \tilde{\phi}$ for ϕ in the form of

$$\phi = \tilde{F}(\tilde{\phi}; g_{\mu\nu}, \partial_\lambda g_{\mu\nu}, \partial_\lambda \partial_\sigma g_{\mu\nu}, \dots), \quad (3.91)$$

which defines the inverse transformation as

$$g_{\mu\nu} = \tilde{g}_{\mu\nu}, \quad \phi = \tilde{F}(\tilde{\phi}; \tilde{g}_{\mu\nu}, \partial_\lambda \tilde{g}_{\mu\nu}, \partial_\lambda \partial_\sigma \tilde{g}_{\mu\nu}, \dots). \quad (3.92)$$

As we did in the previous section, we check the recoverability of the original-frame EOMs. Following the prescription, Eq. (3.90) is linearized as

$$\begin{bmatrix} \delta \tilde{g}_{\mu\nu} \\ \delta \tilde{\phi} \end{bmatrix} = \hat{P} \begin{bmatrix} \delta g_{\alpha\beta} \\ \delta \phi \end{bmatrix}, \quad \hat{P} = \begin{bmatrix} \delta_{(\mu}^{\alpha} \delta_{\nu)}^{\beta} & 0 \\ \hat{p}^{\alpha\beta} & F_\phi \end{bmatrix}, \quad \hat{p}^{\alpha\beta} \equiv \sum_s \frac{\partial F}{\partial (\partial_{(s)} g_{\alpha\beta})} \partial_{(s)}, \quad (3.93)$$

and its inverse transformation (3.92) as

$$\begin{bmatrix} \delta g_{\mu\nu} \\ \delta \phi \end{bmatrix} = \hat{Q} \begin{bmatrix} \delta \tilde{g}_{\alpha\beta} \\ \delta \tilde{\phi} \end{bmatrix}, \quad \hat{Q} = \begin{bmatrix} \delta_{(\mu}^{\alpha} \delta_{\nu)}^{\beta} & 0 \\ -\frac{1}{F_\phi} \hat{p}^{\alpha\beta} & \frac{1}{F_\phi} \end{bmatrix} = \hat{P}^{-1}. \quad (3.94)$$

Now we find the relation between the old- and new-frame EOMs in the same manner as in the previous section:

$$\begin{bmatrix} \mathcal{E}^{\alpha\beta} \\ \mathcal{E}_\phi \end{bmatrix} = \hat{P}^\dagger \begin{bmatrix} \tilde{\mathcal{E}}^{\mu\nu} \\ \tilde{\mathcal{E}}_\phi \end{bmatrix}, \quad \hat{P}^\dagger = \begin{bmatrix} \delta_{(\mu}^\alpha \delta_{\nu)}^\beta & \hat{p}^{\dagger\alpha\beta} \\ 0 & F_\phi \end{bmatrix}. \quad (3.95)$$

In this case, the metric equation $\mathcal{E}^{\alpha\beta}$ becomes of higher order due to the contribution $\hat{p}^{\dagger\alpha\beta} \tilde{\mathcal{E}}_\phi$. Nevertheless, the new system has the same DOFs as the original one because the EOMs in the old frame can be recovered as

$$\begin{bmatrix} \tilde{\mathcal{E}}^{\alpha\beta} \\ \tilde{\mathcal{E}}_\phi \end{bmatrix} = \hat{Q}^\dagger \begin{bmatrix} \mathcal{E}^{\mu\nu} \\ \mathcal{E}_\phi \end{bmatrix}, \quad \hat{Q}^\dagger = \begin{bmatrix} \delta_{(\mu}^\alpha \delta_{\nu)}^\beta & -\hat{p}^{\dagger\alpha\beta} \frac{1}{F_\phi} \\ 0 & \frac{1}{F_\phi} \end{bmatrix}. \quad (3.96)$$

3.3.3 Possible extensions

We have discussed two types of transformations on scalar-tensor theories separately, but one can further consider their compositions. Such a composition of transformations generically takes highly nontrivial form in which the metric and the scalar field are mixed with each other. For example, the following transformation

$$\tilde{g}_{\mu\nu} = e^{2\phi} g_{\mu\nu}, \quad \tilde{\phi} = \phi + e^{-2\phi} (\mathcal{R} - 6X - 6\Box\phi) \quad (3.97)$$

has its inverse transformation and it is given by

$$g_{\mu\nu} = e^{-2(\tilde{\phi} - \tilde{\mathcal{R}})} \tilde{g}_{\mu\nu}, \quad \phi = \tilde{\phi} - \tilde{\mathcal{R}}. \quad (3.98)$$

As such, the space of invertible transformations on scalar-tensor theories has quite a rich structure, and one can obtain healthy theories with arbitrarily higher-order derivatives from a given healthy scalar-tensor theory via invertible transformation.

Moreover, it may be possible to extend the framework of disformal transformation by including the second derivative of ϕ , as was suggested in Ref. [53]. Namely, we can consider a transformation of the metric of the form

$$\tilde{g}_{\mu\nu} = A g_{\mu\nu} + B \phi_\mu \phi_\nu + C \phi_{\mu\nu} + D \phi_\mu^\lambda \phi_{\lambda\nu} + \dots, \quad (3.99)$$

where A, B, C, \dots are arbitrary scalar functions constructed from ϕ, ϕ_μ , and $\phi_{\mu\nu}$. However, to the best of our knowledge, we do not know any invertible transformation that has nontrivial dependence on $\phi_{\mu\nu}$. Therefore, we expect that the number of DOFs would be changed by such transformations as Eq. (3.99).

So far we have considered only the case of single-field scalar-tensor theories, but multi-field extension of disformal transformations was introduced in Ref. [71]. For theories with \mathcal{N} scalar fields ϕ^I , one can consider the following multi-disformal transformation:

$$\tilde{g}_{\mu\nu} = A(\phi^I, X^{IJ}) g_{\mu\nu} + B_{KL}(\phi^I, X^{IJ}) \phi_\mu^K \phi_\nu^L, \quad X^{IJ} \equiv g^{\mu\nu} \phi_\mu^I \phi_\nu^J, \quad (3.100)$$

where the functions A and B_{IJ} are chosen so that

$$A \neq 0 \quad \text{and} \quad \det D_J^I \neq 0, \quad D_J^I \equiv A\delta_J^I + X^{IK}B_{KJ}, \quad (3.101)$$

which guarantees the existence of the inverse metric $\tilde{g}^{\mu\nu}$. As shown in Appendix B.3, the transformation (3.100) has an inverse transformation if and only if the determinant of the $\frac{\mathcal{N}(\mathcal{N}+1)}{2} \times \frac{\mathcal{N}(\mathcal{N}+1)}{2}$ matrix

$$\mathcal{J}^{IJ}{}_{KL} \equiv D_N^I \frac{\partial}{\partial X^{KL}} \left[(D^{-1})_M^N X^{MJ} \right] \quad (3.102)$$

is nonvanishing. Then, it follows from the Theorem that a multi-disformal transformation with this property does not change the DOFs of a given multi-field scalar-tensor theory.

A similar transformation has been discussed even in the context of vector-tensor theories [23], which is of the form

$$\tilde{g}_{\mu\nu} = \Omega(Y)g_{\mu\nu} + \Gamma(Y)A_\mu A_\nu, \quad \tilde{A}_\mu = \Upsilon(Y)A_\mu, \quad (3.103)$$

where Y is the vector mass term: $Y \equiv g^{\mu\nu}A_\mu A_\nu$. It should be noted that the metric and the vector field are transformed simultaneously. One can verify that there exists an inverse transformation of Eq. (3.103) if the functions Ω , Γ , and Υ satisfy

$$\Omega\Upsilon\Xi \neq 0 \quad \text{and} \quad \frac{\Upsilon^2}{\Xi} \neq \text{constant}, \quad \Xi \equiv \Gamma + \frac{\Omega}{Y}, \quad (3.104)$$

as shown in Appendix B.4. Hence, a vector disformal transformation that meets the condition (3.104) does not change the number of DOFs of vector-tensor theories.

Chapter 4

Noninvertible transformation

Among healthy single-field scalar-tensor theories in four dimensions, the known broadest framework is the quadratic/cubic DHOST theories [18–20], whose action is at most cubic order in the second derivative of the scalar field. Disformal transformations are known as a useful tool to obtain a scalar-tensor theory from another one, but it does not help discovering unknown theories since a disformal transformation of quadratic/cubic DHOST theories belongs to the same class. In this sense, it is impossible to generate new healthy theories from known ones via disformal transformation.

In light of this situation, it is intriguing to study whether a new class of healthy scalar-tensor theories can be generated from *nondegenerate* theories that contain extra DOFs. This cannot be achieved by invertible transformation since it does not change the number of physical DOFs, as our Theorem in the previous chapter suggests. However, the possibility is still open for *noninvertible* transformations. In this context, an interesting theory is *mimetic gravity* [34] (see Ref. [35] for a review). This theory is obtained from the Einstein-Hilbert action of GR by performing the following noninvertible conformal transformation:^{*1}

$$\tilde{g}_{\mu\nu} = -X g_{\mu\nu}, \quad (4.1)$$

where $\tilde{g}_{\mu\nu}$ and $g_{\mu\nu}$ denote respectively the metrics of the original frame (namely, the GR frame) and the new frame. This theory can mimic the behavior of pressureless dust in GR [34] and thus is one of the candidates for dark matter. The above formulation of mimetic gravity can be straightforwardly extended by generalizing the “seed” action of the original frame to that of a scalar-tensor theory such as the Horndeski theory [37]. The noninvertibility of the transformation (4.1) is manifested by the fact that the right-hand side of Eq. (4.1) is invariant under the conformal transformation $g_{\mu\nu} \rightarrow \Omega^2 g_{\mu\nu}$, with Ω being an arbitrary function of spacetime. Although the resultant theory written in terms of $(g_{\mu\nu}, \phi)$ has higher-order derivatives of ϕ , the authors of Ref. [76] performed

^{*1}Although the use of the word “singular” would be more appropriate than “noninvertible” for the transformation (4.1), we use the latter in connection with Chap. 3.

a Hamiltonian analysis to show that the theory has only three DOFs due to the local conformal symmetry introduced by the transformation (4.1). Remarkably, theories with 3 DOFs could be obtained even if one starts from a large class of nondegenerate higher-order scalar-tensor theories instead of GR or healthy scalar-tensor theories with 3 DOFs. Indeed, the author of Ref. [41] performed a Hamiltonian analysis of the mimetic theory resulting from

$$S = \int d^4x \sqrt{-g} \left[\frac{M_{\text{Pl}}^2}{2} \mathcal{R} + f(\square\phi) \right], \quad (4.2)$$

with f being an arbitrary scalar function, and proved that the model has at most three DOFs.^{*2} (Note here that it might be more appropriate to use the notation like $\tilde{\mathcal{R}}$ etc. to write the seed action, but we do not do so in order to avoid too many tildes. Therefore, to generate a mimetic theory one should replace $g_{\mu\nu}$ in the seed action by $-Xg_{\mu\nu}$.) In the fluid description, such higher derivative terms typically introduce imperfectness [38] and the scalar DOF acquires a nonzero sound speed [39], which may solve some of the small-scale problems like the missing-satellite problem and the core-cusp problem [40].

In the present chapter, we consider generic scalar-tensor theories that could possess an unwanted extra DOF as a generalization of Eq. (4.2), and perform the noninvertible transformation (4.1) on them, as suggested in Ref. [78]. We show explicitly that the extended mimetic gravity models obtained thus have at most three DOFs based on a Hamiltonian analysis. This turns out to be true also for models obtained via noninvertible disformal transformation that is more general than (4.1) (see Appendix B.2). Due to the diversity of the original theories, many of the mimetic theories lie outside the quadratic/cubic DHOST class and they cannot be obtained by disformal transformation of any known class of healthy scalar-tensor theories. Nevertheless, such mimetic theories have a problem with cosmological perturbations: It was demonstrated in Refs. [79, 80] that the mimetic model obtained from the action of the form (4.2) has ghost/gradient instabilities in cosmological perturbations. The simplest version of mimetic gravity is closely related to the low-energy limit of Hořava-Lifshitz gravity [81, 82], and the same instability was also pointed out in the latter context in Ref. [83]. Concerning this point, the authors of Ref. [84] developed an effective theory of cosmological perturbations in mimetic theories and showed that the instability can be cured by introducing nonminimal derivative couplings to gravity. In this chapter, we also study the linear stability of cosmological perturbations in our extended mimetic gravity and show that the models obtained in the aforementioned manner generically exhibit the very same problem of ghost/gradient instabilities (except for the special case in which scalar perturbations appear to be strongly coupled).

The rest of this chapter is organized as follows. In §4.1, we begin with presenting the general seed action which we use to generate a variety of mimetic gravity theories and then perform the ADM decomposition of the seed action. Then, in §4.2 we transform

^{*2}Strictly speaking, the author of Ref. [41] used an alternative formulation of mimetic gravity proposed in Ref. [77] (see §4.2.2).

the seed action to its mimetic counterpart via the noninvertible conformal transformation (4.1), and analyze the resultant theory in the language of the Hamiltonian formalism. Cosmological perturbations in the extended mimetic gravity models are discussed in §4.3.

4.1 ADM form of the seed scalar-tensor theory

4.1.1 The seed action

Generalizing the action (4.2), we start from the following seed theory:

$$S = \int d^4x \sqrt{-g} [f_2 \mathcal{R} + f_3 \mathcal{G}^{\mu\nu} \nabla_\mu \nabla_\nu \phi + F(g_{\mu\nu}, \phi, \nabla_\mu \phi, \nabla_\mu \nabla_\nu \phi)], \quad (4.3)$$

where f_2 and f_3 are arbitrary functions of (ϕ, X) , and F denotes any scalar quantity constructed from the metric, the scalar field, and its derivatives up to second order. Note that the couplings to the curvature tensors are the same as those found in the Horndeski theory [see Eq. (2.5)].

Using the action of the form (4.3) as a seed, we perform the noninvertible conformal transformation

$$g_{\mu\nu} \rightarrow \tilde{g}_{\mu\nu} = -X g_{\mu\nu}, \quad (4.4)$$

where $\tilde{g}_{\mu\nu}$ is identified as the metric in the original frame (4.3), while $g_{\mu\nu}$ is now the metric of the new theory. It should be noted that the resultant action is also of the form (4.3) (see Appendix B). The transformation (4.4) is noninvertible as the right-hand side is invariant under conformal transformation of $g_{\mu\nu}$. As a result, the new theory acquires a local conformal symmetry.^{*3} See the recent paper by Horndeski [85] for conformally invariant scalar-tensor theories that are *flat space compatible*, i.e., such that one can take the limit $g_{\mu\nu} \rightarrow \eta_{\mu\nu}$ and $\phi \rightarrow \text{constant}$. In general, one could consider noninvertible disformal transformations other than Eq. (4.4). However, as far as our purpose is concerned, we lose no generality by restricting ourselves to the transformation (4.4) since any noninvertible disformal transformation can be recast in this form. This point is addressed in Appendix B.

In Eq. (4.3) we consider only particular couplings between the scalar field and the curvature tensors, i.e., $f_2(\phi, X)\mathcal{R}$ and $f_3(\phi, X)\mathcal{G}^{\mu\nu}\nabla_\mu\nabla_\nu\phi$. Other types of couplings such as $\mathcal{R}(\square\phi)^2$, $f(\phi, X)\square\mathcal{R}$, etc. would give rise to unwanted extra DOFs in the resultant mimetic theory. This point will become clear in the Hamiltonian analysis below: If one considers the couplings other than the first two terms in Eq. (4.3), then one would be forced to introduce some new velocities, leading to the extra DOFs (see §4.2.2). For

^{*3}In this sense, the noninvertible conformal transformation resembles the Stückelberg trick mentioned in §3.2.3, which restores $U(1)$ gauge symmetry of a vector field. A crucial difference is the (non)invertibility of the transformation law.

the same reason, we do not include third or higher derivatives of ϕ in Eq. (4.3). Thus, Eq. (4.3) would be a minimal seed theory whose mimetic gravity counterpart has at most 3 DOFs.

The class of theories defined by Eq. (4.3) includes all the known healthy scalar-tensor theories that possess general covariance, such as the Horndeski theory [14], GLPV theories [17], and quadratic/cubic DHOST theories [18–20]. Such healthy theories correspond to specific choices of the functions f_2 , f_3 , and F . However, for generic functions, the seed theory (4.3) in its original frame would have Ostrogradsky ghosts. Nevertheless, as we will show, the noninvertible transformation (4.4) makes the resultant theory degenerate, leaving only 3 DOFs.

4.1.2 ADM decomposition

To proceed to a Hamiltonian analysis, we first express the seed action (4.3) in the ADM form and then perform the conformal transformation (4.4) written in terms of the ADM variables. In this subsection we present some technical detail for recasting Eq. (4.3) into the ADM form.

In addition to the usual ADM variables, we also introduce the following variables associated with time derivatives of the scalar field as in Ref. [62]:

$$A_* \equiv n^\mu \nabla_\mu \phi = \frac{\dot{\phi} - N^i D_i \phi}{N}, \quad (4.5)$$

$$V_* \equiv n^\mu n^\nu \nabla_\mu \nabla_\nu \phi = \frac{\dot{A}_* - D^i \phi D_i N - N^i D_i A_*}{N}. \quad (4.6)$$

Below, we will introduce a Lagrange multiplier and regard A_* as an auxiliary variable which satisfies Eq. (4.5) dynamically so that second-order time derivatives do not appear explicitly in the action. Then, V_* plays the role of the velocity of A_* . Using A_* , the scalar kinetic term X is written as

$$X = -A_*^2 + D^i \phi D_i \phi. \quad (4.7)$$

It is worth emphasizing here that the unitary gauge $\phi = t$ is not imposed from the beginning since it would be misleading in some cases [18, 62]. Nevertheless, in the case of mimetic gravity, it turns out that the number of physical DOFs is not changed by the unitary gauge fixing, as we will see in §4.2.2.

With these notations, one can decompose $\nabla_\mu \phi$ and $\nabla_\mu \nabla_\nu \phi$ as

$$\nabla_\mu \phi = h_\mu^i D_i \phi - n_\mu A_*, \quad (4.8)$$

$$\nabla_\mu \nabla_\nu \phi = h_{(\mu}^i h_{\nu)}^j (D_i D_j \phi - A_* K_{ij}) - 2h_{(\mu}^i n_{\nu)} (D_i A_* - K_{ij} D^j \phi) + n_\mu n_\nu V_*. \quad (4.9)$$

These decompositions allow us to recast the third term of the action (4.3) in the form

$$\begin{aligned} & \int d^4x \sqrt{-g} F(g_{\mu\nu}, \phi, \nabla_\mu \phi, \nabla_\mu \nabla_\nu \phi) \\ &= \int dt d^3x \left[N \sqrt{\gamma} L_F(\gamma_{ij}, \phi, A_*; K_{ij}, V_*; D_i) + \Lambda (N A_* + N^i D_i \phi - \dot{\phi}) \right], \end{aligned} \quad (4.10)$$

where the concrete form of L_F depends on that of F , and the last term with a Lagrange multiplier Λ fixes A_* so that it satisfies Eq. (4.5).

The first two terms in Eq. (4.3) involving the curvature tensors can be written in the ADM form by using the Gauss/Codazzi/Ricci equations. To simplify the manipulation, we first perform integration by parts to move one of the derivative operators acting on ϕ to f_3 . Then, the result is given as follows:

$$\begin{aligned} & \int d^4x \sqrt{-g} (f_2 \mathcal{R} + f_3 \mathcal{G}^{\mu\nu} \nabla_\mu \nabla_\nu \phi) = \int d^4x \sqrt{-g} (f_2 \mathcal{R} - \mathcal{G}^{\mu\nu} \nabla_\mu \phi \nabla_\nu f_3) \\ &= \int dt d^3x N \sqrt{\gamma} \left\{ f_2 (R + K_{ij}^2 - K^2) - 2K f_{2\perp} - 2D^i D_i f_2 \right. \\ &\quad - \frac{1}{2} (R - K_{ij}^2 + K^2) A_* f_{3\perp} - \left[R_{ij} - \frac{1}{2} (R + K_{kl}^2 - K^2) \gamma_{ij} \right] D^i \phi D^j f_3 \\ &\quad + (K \gamma^{ij} - K^{ij}) (2K_i^k D_k \phi D_j f_3 + f_{3\perp} D_i D_j \phi + A_* D_i D_j f_3) \\ &\quad \left. + D_i D_j (D^i \phi D^j f_3) - D_i D^i (D_j \phi D^j f_3) + \Lambda (N A_* + N^i D_i \phi - \dot{\phi}) \right\}. \end{aligned} \quad (4.11)$$

Here, for a scalar function $f(\phi, X)$ we have defined

$$\begin{aligned} f_\perp(\gamma_{ij}, \phi, A_*; K_{ij}, V_*; D_i) &\equiv n^\mu \nabla_\mu f \\ &= f_\phi A_* - 2f_X (K_{ij} D^i \phi D^j \phi + A_* V_* - D^i \phi D_i A_*). \end{aligned} \quad (4.12)$$

Putting Eqs. (4.10) and (4.11) together, one finds that the total action (4.3) can be written in the form

$$S = \int dt d^3x \left[N \sqrt{\gamma} L_0(\gamma_{ij}, R_{ij}, \phi, A_*; K_{ij}, V_*; D_i) + \Lambda (N A_* + N^i D_i \phi - \dot{\phi}) \right], \quad (4.13)$$

where the dependence of L_0 on N, N^i is encapsulated in K_{ij} and V_* . Note that in the actual expression of L_0 the spatial derivative D_i does not act on K_{ij} or V_* .

4.2 Extended mimetic gravity

4.2.1 Hamiltonian analysis

In the previous section, we have written the seed action (4.3) in terms of the ADM variables to obtain Eq. (4.13). Now we move from the original frame (4.13) to another by

performing the noninvertible conformal transformation (4.4), and thereby generate new mimetic gravity actions.

Under the transformation (4.4), the 3-dimensional quantities are transformed as follows:

$$\tilde{N} = \sqrt{-X}N, \quad \tilde{N}^i = N^i, \quad \tilde{\gamma}_{ij} = -X\gamma_{ij}, \quad \tilde{A}_* = \frac{1}{\sqrt{-X}}A_*, \quad (4.14)$$

$$\tilde{R}_{ij} = R_{ij} + \frac{3}{4X^2}D_iX D_jX - \frac{1}{2X}D_iD_jX + \gamma_{ij} \left(\frac{1}{4X^2}D_kX D^kX - \frac{1}{2X}D_kD^kX \right), \quad (4.15)$$

$$\tilde{K}_{ij} = \sqrt{-X} \left[K_{ij} - \frac{1}{X}\gamma_{ij} (K_{kl}D^k\phi D^l\phi + A_*V_* - D^k\phi D_kA_*) \right], \quad (4.16)$$

$$\tilde{V}_* = -\frac{1}{X^2} (A_*K_{ij}D^i\phi D^j\phi + V_*D^i\phi D_i\phi - D^i\phi D^j\phi D_iD_j\phi), \quad (4.17)$$

where original-frame variables are now denoted with tildes. Note that the original-frame scalar kinetic term \tilde{X} is mapped to a constant:

$$\tilde{X} = \tilde{g}^{\mu\nu}\tilde{\nabla}_\mu\phi\tilde{\nabla}_\nu\phi = -\frac{1}{X}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi = -1, \quad (4.18)$$

which immediately leads to $\tilde{\nabla}_\lambda\tilde{X} = 2\tilde{g}^{\mu\nu}\tilde{\nabla}_\mu\phi\tilde{\nabla}_\nu\tilde{\nabla}_\lambda\phi = 0$. This means that any scalar quantity that contains a contraction of $\tilde{\nabla}_\mu\phi$ and $\tilde{\nabla}_\mu\tilde{\nabla}_\nu\phi$ vanishes in the new frame.

It should be noted that in Eqs. (4.16) and (4.17) the velocity dependence appears only through the particular combination

$$V_{ij} \equiv K_{ij} + \frac{V_*}{A_*}\gamma_{ij}, \quad (4.19)$$

which is a consequence of the conformal symmetry introduced by performing the noninvertible conformal transformation (4.4). In terms of this V_{ij} , Eqs. (4.16) and (4.17) can be written as

$$\tilde{K}_{ij} = \sqrt{-X} \left[\left(\delta_i^k\delta_j^l - \frac{D^k\phi D^l\phi}{X}\gamma_{ij} \right) V_{kl} + \frac{D^k\phi D_kA_*}{X}\gamma_{ij} \right], \quad (4.20)$$

$$\tilde{V}_* = -\frac{1}{X^2}D^i\phi D^j\phi (A_*V_{ij} - D_iD_j\phi). \quad (4.21)$$

As a side remark, the following identity holds,

$$\tilde{V}_{ij} = \sqrt{-X}V_{ij} - \left(\frac{D^k\phi}{A_*}D_k\sqrt{-X} \right) \gamma_{ij}, \quad (4.22)$$

which will be used later. Substituting Eqs. (4.14), (4.15), (4.20), and (4.21) to the seed action (4.13), we finally arrive at the action in the new frame:

$$S = \int dt d^3x \left[N\sqrt{\gamma}L_M(\gamma_{ij}, R_{ij}, \phi, A_*; V_{ij}; D_i) + \Lambda(NA_* + N^iD_i\phi - \dot{\phi}) \right]. \quad (4.23)$$

For a technical purpose, we introduce auxiliary variables B_{ij} with Lagrange multipliers λ^{ij} and rewrite the action as

$$S = S_M[N, \gamma_{ij}, \phi, A_*, B_{ij}] + \int dt d^3x \left[\Lambda(NA_* + N^i D_i \phi - \dot{\phi}) + N\lambda^{ij}(B_{ij} - V_{ij}) \right], \quad (4.24)$$

$$S_M[N, \gamma_{ij}, \phi, A_*, B_{ij}] \equiv \int dt d^3x N \sqrt{\gamma} L_M(\gamma_{ij}, R_{ij}, \phi, A_*; B_{ij}; D_i). \quad (4.25)$$

Here, the lapse function N in front of λ^{ij} was introduced so that the resultant Hamiltonian is linear in N and N^i . In the following, we perform a Hamiltonian analysis of the theory (4.24). We define the canonical pairs as follows:

$$\left(N, N^i, \gamma_{ij}, \phi, A_*, B_{ij}, \Lambda, \lambda^{ij} \right) \\ \left(\pi_N, \pi_i, \pi^{ij}, p_\phi, p_*, p^{ij}, P, P_{ij} \right). \quad (4.26)$$

These variables form a 50-dimensional phase space.

The canonical momenta are calculated from the action (4.24) in the standard manner. Since the action does not contain the velocities of N, N^i, B_{ij}, Λ , and λ^{ij} , the corresponding canonical momenta vanish:

$$\pi_N = \pi_i = p^{ij} = P = P_{ij} = 0, \quad (4.27)$$

which provides primary constraints. The canonical momenta for γ_{ij}, ϕ , and A_* are given by

$$\pi^{ij} = \frac{\delta S}{\delta \dot{\gamma}_{ij}} = -\frac{1}{2} \lambda^{ij}, \quad (4.28)$$

$$p_\phi = \frac{\delta S}{\delta \dot{\phi}} = -\Lambda, \quad (4.29)$$

$$p_* = \frac{\delta S}{\delta \dot{A}_*} = -\frac{1}{A_*} \gamma_{ij} \lambda^{ij}. \quad (4.30)$$

These expressions for the canonical momenta yield further primary constraints as

$$\bar{\pi}^{ij} \equiv \pi^{ij} + \frac{1}{2} \lambda^{ij} \approx 0, \quad (4.31)$$

$$\bar{p}_\phi \equiv p_\phi + \Lambda \approx 0, \quad (4.32)$$

$$\mathcal{C} \equiv A_* p_* - 2\gamma_{ij} \pi^{ij} \approx 0. \quad (4.33)$$

It should be noted that \mathcal{C} is the generator of conformal transformation. This relation between π^{ij} and p_* comes from Eq. (4.19) with the identity

$$A_* \frac{\partial V_{ij}}{\partial \dot{A}_*} = 2\gamma_{kl} \frac{\partial V_{ij}}{\partial \dot{\gamma}_{kl}}. \quad (4.34)$$

To see the first-class nature of \mathcal{C} , we construct a linear combination with P_{ij} so that the resultant constraint weakly Poisson commutes with all the other primary constraints:

$$\bar{\mathcal{C}} \equiv \mathcal{C} + 2\lambda^{ij}P_{ij} = A_*p_* - 2\gamma_{ij}\pi^{ij} + 2\lambda^{ij}P_{ij}. \quad (4.35)$$

Note that the discussion so far does not depend on whether the original seed theory (4.13) is degenerate or not. It might be possible that the resultant mimetic theory (4.24) possesses additional primary constraints which have not been specified above, but it has nothing to do with the (non)degeneracy of the original theory.

Now the total Hamiltonian is obtained as

$$H_T = \int d^3x \left(N\mathcal{H} + N^i\mathcal{H}_i + \mu_N\pi_N + \mu^i\pi_i + \mu_{ij}\bar{\pi}^{ij} + u_\phi\bar{p}_\phi + u_*\bar{\mathcal{C}} + u_{ij}p^{ij} + UP + U^{ij}P_{ij} \right), \quad (4.36)$$

where

$$\mathcal{H} \equiv -\sqrt{\gamma}L_M(\gamma_{ij}, R_{ij}, \phi, A_*; B_{ij}; D_i) + 2\pi^{ij}B_{ij} + p_\phi A_* - \sqrt{\gamma}D_i \left(\frac{p_*}{\sqrt{\gamma}} D^i \phi \right), \quad (4.37)$$

$$\mathcal{H}_i \equiv -2\sqrt{\gamma}D^j \left(\frac{\pi_{ij}}{\sqrt{\gamma}} \right) + p_\phi D_i \phi + p_* D_i A_* + p^{jk} D_i B_{jk} - 2\sqrt{\gamma}D_j \left(\frac{p^{jk}}{\sqrt{\gamma}} B_{ik} \right). \quad (4.38)$$

Note that the last two terms in \mathcal{H}_i are proportional to p^{ij} , which vanishes on the constraint surface. Nevertheless, we keep these terms because they generate an infinitesimal spatial diffeomorphism of B_{ij} .

Let us calculate the time evolution of the primary constraints. It is easy to see that the time evolution of $\bar{\pi}^{ij}$, \bar{p}_ϕ , P , and P_{ij} fixes the Lagrange multipliers U^{ij} , U , u_ϕ , and μ_{ij} , respectively. The time evolution of π_N , π_i , p^{ij} leads to

$$\dot{\pi}_N = \{\pi_N, H_T\}_P \approx -\mathcal{H}, \quad (4.39)$$

$$\dot{\pi}_i = \{\pi_i, H_T\}_P \approx -\mathcal{H}_i, \quad (4.40)$$

$$\dot{p}^{ij} = \{p^{ij}, H_T\}_P \approx N \left(\sqrt{\gamma} \frac{\partial L_M}{\partial B_{ij}} - 2\pi^{ij} \right), \quad (4.41)$$

where we have used the fact that the derivatives of B_{ij} do not appear in L_M . Therefore, we find secondary constraints as

$$\mathcal{H} \approx 0, \quad \mathcal{H}_i \approx 0, \quad \varphi^{ij} \equiv \sqrt{\gamma} \frac{\partial L_M}{\partial B_{ij}} - 2\pi^{ij} \approx 0. \quad (4.42)$$

Note that the time evolution of $\bar{\mathcal{C}}$ does not yield a new constraint because its Poisson bracket with the total Hamiltonian is written only by the constraints:

$$\begin{aligned} \dot{\bar{\mathcal{C}}} &= \{\bar{\mathcal{C}}, H_T\}_P \\ &= -N\mathcal{H} + NB_{ij}\varphi^{ij} + \partial_i \left[N \frac{D^i \phi}{A_*} (\gamma_{jk}\varphi^{jk} - \mathcal{C}) + N^i \mathcal{C} \right] - 2\mu_{ij}\bar{\pi}^{ij} + 2U^{ij}P_{ij} \approx 0, \end{aligned} \quad (4.43)$$

where we have used the Noether identity with respect to the conformal symmetry (D.20).

Having obtained all the secondary constraints, let us consider their time evolution. Assuming that L_M depends at least quadratically on B_{ij} , the time evolution of φ^{ij} fixes the Lagrange multiplier u_{ij} . To discuss the evolution of \mathcal{H} and \mathcal{H}_i , we first note that $\mathcal{H} \approx 0$ and $\mathcal{H}_i \approx 0$ are expected to correspond to the Hamiltonian and momentum constraints, respectively. Let us define smeared quantities as

$$H_L[\mathcal{N}] \equiv \int d^3x \mathcal{N} \mathcal{H}, \quad H_S[\mathcal{N}^i] \equiv \int d^3x \mathcal{N}^i \mathcal{H}_i, \quad (4.44)$$

where \mathcal{N} and \mathcal{N}^i are arbitrary test functions (and not the lapse function and the shift vector). One can check that all the primary constraints Poisson commute with $H_L[\mathcal{N}]$ and $H_S[\mathcal{N}^i]$. With some manipulation, the Poisson brackets between $H_L[\mathcal{N}]$ and $H_S[\mathcal{N}^i]$ are found to be identical to those in GR,

$$\{H_S[\mathcal{N}^i], H_S[\mathcal{M}^i]\}_P = H_S[\mathcal{N}^j D_j \mathcal{M}^i - \mathcal{M}^j D_j \mathcal{N}^i], \quad (4.45)$$

$$\{H_S[\mathcal{N}^i], H_L[\mathcal{M}]\}_P = H_L[\mathcal{N}^i D_i \mathcal{M}], \quad (4.46)$$

where we have used the Noether identity associated with 3-dimensional diffeomorphism (D.17).^{*4} The explicit calculation of $\{H_L[\mathcal{N}], H_L[\mathcal{M}]\}_P$ is lengthy and tedious, and therefore we skip it. Nevertheless, it is reasonable to conclude that $\mathcal{H} \approx 0$ and $\mathcal{H}_i \approx 0$ are first-class constraints corresponding to general covariance and their time evolution does not yield any new constraint. See Ref. [62] for the related discussion on this point.

To sum up, we have obtained the set of constraints as follows:

$$\begin{aligned} 9 \text{ first-class} &: \pi_N, \pi_i, \bar{\mathcal{C}}, \mathcal{H}, \mathcal{H}_i, \\ 26 \text{ second-class} &: \bar{\pi}^{ij}, \bar{p}_\phi, p^{ij}, P, P_{ij}, \varphi^{ij}. \end{aligned} \quad (4.47)$$

These constraints reduce the phase-space dimension and

$$\frac{1}{2} (50 - 9 \times 2 - 26) = 3 \text{ DOFs} \quad (4.48)$$

are left, which means that there is no unwanted extra DOF. Note that this is the maximum possible number of physical DOFs that the theory (4.24) has. Even if the evolution of φ^{ij} yields some additional constraint as opposed to the above argument, it never increases the number of DOFs. Note also that the original theory (4.3) could have a different number of physical DOFs. The above result holds irrespective of whether we start from GR with two DOFs (as in original mimetic gravity [34]) or generic nondegenerate higher-order scalar-tensor theories with four DOFs.

^{*4}When we apply the Noether identity (D.17), we may replace the lapse function N with a scalar test function \mathcal{M} in $H_L[\mathcal{M}]$.

4.2.2 Remarks

Several remarks are in order. First, we restricted ourselves to the case where the curvature tensors appear only in the form of $f_2(\phi, X)\mathcal{R}$ and $f_3(\phi, X)\mathcal{G}^{\mu\nu}\nabla_\mu\nabla_\nu\phi$, since otherwise additional velocities must be introduced. For example, in the case of mimetic $f(\mathcal{R})$ gravity [36], one has to introduce a new velocity $V_K \equiv n^\mu\nabla_\mu K$ in addition to K_{ij} and V_* . This results in an undesired extra DOF, which cannot be killed by the constraint corresponding to the conformal symmetry [84]. The situation is similar if we include higher-order derivatives of ϕ in the action.

Second, let us comment on the relation between our extended mimetic gravity models and quadratic/cubic DHOST theories. For a generic choice of the function F in the seed action (4.3), the resultant mimetic action has terms of quartic or higher order in $\nabla_\mu\nabla_\nu\phi$, which cannot be reached from the quadratic/cubic DHOST class via any disformal transformation. On the other hand, as it should be, if F is of at most cubic order in $\nabla_\mu\nabla_\nu\phi$, the corresponding mimetic model falls into the quadratic/cubic DHOST class [44]. Now we perform the transformation (4.4) on the action (2.20). It should be noted that the old-frame coefficients, which are originally functions of (ϕ, \tilde{X}) , are now interpreted as functions only of ϕ because $\tilde{X} = -1$ [see Eq. (4.18)]. Moreover, with the aid of this constraint on \tilde{X} , the term with the Einstein tensor in the action (4.3) can be integrated by parts to give

$$\begin{aligned} & \int d^4x \sqrt{-\tilde{g}} \tilde{f}_3 \tilde{\mathcal{G}}^{\mu\nu} \tilde{\nabla}_\mu \tilde{\nabla}_\nu \phi \\ &= \int d^4x \sqrt{-\tilde{g}} \left\{ -\tilde{f}_{3\phi} \left[\frac{1}{2} \tilde{\mathcal{R}} + (\tilde{\square}\phi)^2 - (\tilde{\nabla}_\mu \tilde{\nabla}_\nu \phi)^2 \right] + \tilde{f}_{3\phi\phi\phi} \right\}, \end{aligned} \quad (4.49)$$

from which we see that the contribution of the \tilde{f}_3 term can be absorbed into $\tilde{f}_2, \tilde{a}_1, \tilde{a}_2$, and \tilde{F}_0 . Thus, we set $\tilde{f}_3 = 0$ from the beginning without loss of generality. The new-frame action also belongs to the same class of theories as the original one, and the functions f_2, a_i, b_j are written in terms the old-frame functions $\tilde{f}_2, \tilde{a}_i, \tilde{b}_j$ as follows:

$$f_2 = -X \tilde{f}_2, \quad (4.50)$$

$$\begin{aligned} a_1 &= \tilde{a}_1, & a_2 &= \tilde{a}_2, & a_3 &= \frac{2}{X} (\tilde{a}_1 + 2\tilde{a}_2), \\ a_4 &= -\frac{2}{X} (\tilde{a}_1 + 3\tilde{f}_2), & a_5 &= \frac{2}{X^2} (\tilde{a}_1 + 2\tilde{a}_2), \end{aligned} \quad (4.51)$$

$$\begin{aligned} b_1 &= -\frac{1}{X} \tilde{b}_1, & b_2 &= -\frac{1}{X} \tilde{b}_2, & b_3 &= -\frac{1}{X} \tilde{b}_3, & b_4 &= -\frac{2}{X^2} (3\tilde{b}_1 + \tilde{b}_2), \\ b_5 &= \frac{2}{X^2} \tilde{b}_2, & b_6 &= -\frac{1}{X^2} (2\tilde{b}_2 + 3\tilde{b}_3), & b_7 &= \frac{3}{X^2} \tilde{b}_3, & b_8 &= \frac{1}{X^3} (4\tilde{b}_2 + 3\tilde{b}_3), \\ b_9 &= -\frac{3}{X^3} (4\tilde{b}_1 + 2\tilde{b}_2 + \tilde{b}_3), & b_{10} &= -\frac{2}{X^4} (4\tilde{b}_1 + 2\tilde{b}_2 + \tilde{b}_3). \end{aligned} \quad (4.52)$$

Note that there is no contribution of $\tilde{a}_3, \tilde{a}_4, \tilde{a}_5$, or $\tilde{b}_4, \dots, \tilde{b}_{10}$ because the corresponding building blocks contain $\tilde{\nabla}_\mu \phi \tilde{\nabla}^\mu \tilde{\nabla}_\nu \phi$ which is mapped to zero by the transformation (4.4). Since the resultant mimetic theory has at most three DOFs (see §4.2.1), it must be a DHOST theory. Indeed, one can show that this theory is represented as a combination of ²N-III and ³M-I theories in the terminology of Ref. [20], and hence is degenerate.

The third comment is on an alternative formulation of mimetic gravity. It was pointed out in Ref. [77] that imposing the constraint $X = -1$ (the *mimetic constraint*), as implied by Eq. (4.18), leads to a theory which is equivalent to the one obtained via the noninvertible transformation (4.4). Note that Eq. (4.4) reads $\tilde{g}_{\mu\nu} = g_{\mu\nu}$ when one sets $X = -1$, which means that imposing the constraint $X = -1$ after performing the transformation (4.4) generates the same theory as the one obtained by imposing $\tilde{X} = -1$ from the beginning in the original frame. In the Lagrangian formalism, the equivalence between the two formulations can directly be verified by comparing the EOMs. In contrast, in the language of the Hamiltonian analysis in §4.2.1, one could regard the mimetic constraint as a gauge condition that completely fixes the conformal gauge DOF. Thus, we could safely impose the constraint $X = -1$ from the beginning, which would significantly simplify the analysis. For related arguments on eliminating ghost DOFs by constraints, see Ref. [86].

Finally, we discuss the issue of the unitary gauge. In the previous section, we did not fix the coordinate system and performed the Hamiltonian analysis maintaining general covariance. However, in many cases, it is convenient to impose the unitary gauge $\phi = t$. Note that this gauge choice is compatible with the mimetic constraint (4.18), which ensures that $\nabla_\mu \phi$ is timelike. Let us see that the number of DOFs of mimetic gravity does not change due to the unitary gauge fixing. Under the unitary gauge with the mimetic constraint, one has $N = 1$, $A_* = 1$, and $V_* = 0$. The action of mimetic gravity (4.23) then reduces to

$$S_{\text{unitary}} = \int dt d^3x \sqrt{\gamma} L(\gamma_{ij}, R_{ij}, t; K_{ij}; D_i). \quad (4.53)$$

The canonical variables are $(N^i, \gamma_{ij}; \pi_i, \pi^{ij})$, which form an 18-dimensional phase space. If the Lagrangian L is nondegenerate, i.e., $\det(\partial^2 L / \partial K_{ij} \partial K_{kl}) \neq 0$, we obtain the primary constraints $\pi_i \approx 0$, and then they lead to the momentum constraints $\mathcal{H}_i \approx 0$, with no further constraints. All these six constraints are first class and they reduce the phase-space dimension to yield

$$\frac{1}{2}(18 - 6 \times 2) = 3 \text{ DOFs}. \quad (4.54)$$

This is consistent with the analysis without unitary gauge fixing. On the contrary, if L is degenerate, one obtains additional constraints on π^{ij} . This is different from the case considered in the Hamiltonian analysis in §4.2.1, as the degeneracy of L in Eq. (4.53) implies that L_M in Eq. (4.23) is also degenerate. Thus, we see that the two analyses (with or without unitary gauge fixing) give consistent results in the present context of mimetic theories.

4.3 Cosmological perturbations

The large generalization of mimetic gravity we have obtained has 3 DOFs and hence is free from obvious instabilities of Ostrogradsky ghosts. This does not mean, however, that general mimetic theories are phenomenologically viable. To see this point, let us analyze perturbations around the flat FLRW background in the mimetic gravity models.

According to the last two remarks in §4.2.2, we may take safely the unitary gauge to write $\phi = t$ and impose the constraint $X = -1$ in the action (4.3), with which the calculation is simplified drastically. As a consequence of $\phi = t$ and $X = -1$, any function of (ϕ, X) can be regarded as a function of t only. We also have $N = 1$ since X is written in terms of N as $X = -1/N^2$ in the unitary gauge. Therefore, the ADM form (4.11) of the first two terms in the action (4.3) reduces to

$$\begin{aligned} & \int d^4x \sqrt{-g} (f_2 \mathcal{R} + f_3 \mathcal{G}^{\mu\nu} \nabla_\mu \nabla_\nu \phi) \\ &= \int dt d^3x \sqrt{\gamma} \left[\left(f_2 - \frac{1}{2} \dot{f}_3 \right) R + \left(f_2 + \frac{1}{2} \dot{f}_3 \right) (K_{ij}^2 - K^2) - 2\dot{f}_2 K \right]. \end{aligned} \quad (4.55)$$

As for the third term in Eq. (4.3), its ADM representation (4.10) is now written in terms of scalar quantities composed of γ_{ij} and K_{ij} , i.e., it can be expressed as a function of $\mathcal{K}_n \equiv K_{i_2}^{i_1} K_{i_3}^{i_2} \cdots K_{i_1}^{i_n}$ ($n = 1, 2, \dots, \ell$):

$$\int d^4x \sqrt{-g} F(g_{\mu\nu}, \phi, \nabla_\mu \phi, \nabla_\mu \nabla_\nu \phi) = \int dt d^3x \sqrt{\gamma} \hat{\mathcal{F}}(t, K, \mathcal{K}_2, \mathcal{K}_3, \dots, \mathcal{K}_\ell), \quad (4.56)$$

where note that $\mathcal{K}_1 = K$. Combining Eqs. (4.55) and (4.56), we obtain the following action for the mimetic counterpart of the theory (4.3):

$$S = \int dt d^3x \sqrt{\gamma} \left[\left(f_2 - \frac{1}{2} \dot{f}_3 \right) R + \mathcal{F}(t, K, \mathcal{K}_2, \mathcal{K}_3, \dots, \mathcal{K}_\ell) \right], \quad (4.57)$$

$$\mathcal{F} \equiv \hat{\mathcal{F}} + \left(f_2 + \frac{1}{2} \dot{f}_3 \right) (\mathcal{K}_2 - K^2) - 2\dot{f}_2 K. \quad (4.58)$$

It is useful to define the first and second derivatives of \mathcal{F} as

$$\mathcal{F}_n \equiv \frac{\partial \mathcal{F}}{\partial \mathcal{K}_n}, \quad \mathcal{F}_{mn} \equiv \frac{\partial^2 \mathcal{F}}{\partial \mathcal{K}_m \partial \mathcal{K}_n}, \quad (4.59)$$

respectively.

Now we substitute the following metric ansatz to the action (4.57),

$$N = 1, \quad N_i = \partial_i \chi, \quad \gamma_{ij} = a^2(t) e^{2\zeta} (e^h)_{ij} = a^2 e^{2\zeta} \left(\delta_{ij} + h_{ij} + \frac{1}{2} h_{ik} h_{jk} + \cdots \right), \quad (4.60)$$

where χ and ζ are scalar perturbations and h_{ij} denotes a transverse-traceless tensor perturbation. The background EOM is given by

$$\dot{\mathcal{P}} + 3H\mathcal{P} - \mathcal{F} = 0, \quad \mathcal{P} \equiv \sum_{n=1}^{\ell} n H^{n-1} \mathcal{F}_n, \quad (4.61)$$

where $H \equiv \dot{a}/a$ is the Hubble parameter and \mathcal{F}_n are evaluated at the background, $\mathcal{K}_n = 3H^n$. This equation will be used to simplify the expressions of the quadratic actions for the tensor and scalar perturbations.

The quadratic action for the tensor perturbation h_{ij} is given by

$$S_{\text{T}}^{(2)} = \int dt d^3x \frac{a^3}{4} \left[\mathcal{B} \dot{h}_{ij}^2 - \mathcal{E} \frac{(\partial_k h_{ij})^2}{a^2} \right], \quad (4.62)$$

where

$$\mathcal{B} \equiv \sum_{n=2}^{\ell} \frac{n(n-1)}{2} H^{n-2} \mathcal{F}_n, \quad \mathcal{E} = f_2 - \frac{1}{2} \dot{f}_3. \quad (4.63)$$

The tensor perturbations are stable provided that $\mathcal{B} > 0$ and $\mathcal{E} > 0$.

The quadratic action for the scalar perturbations ζ and χ is

$$S_{\text{S}}^{(2)} = \int dt d^3x a^3 \left[\frac{3}{2} (3\mathcal{A} + 2\mathcal{B}) \dot{\zeta}^2 + 2\mathcal{E} \frac{(\partial_k \zeta)^2}{a^2} + \frac{1}{2} (\mathcal{A} + 2\mathcal{B}) \left(\frac{\partial^2 \chi}{a^2} \right)^2 - (3\mathcal{A} + 2\mathcal{B}) \dot{\zeta} \frac{\partial^2 \chi}{a^2} \right], \quad (4.64)$$

where we have used the background EOM (4.61) and defined

$$\mathcal{A} \equiv \sum_{m=1}^{\ell} \sum_{n=1}^{\ell} m n H^{m+n-2} \mathcal{F}_{mn}. \quad (4.65)$$

The EOM for χ can be solved to give

$$\frac{\partial^2 \chi}{a^2} = \frac{3\mathcal{A} + 2\mathcal{B}}{\mathcal{A} + 2\mathcal{B}} \dot{\zeta}, \quad (4.66)$$

where we have assumed that $\mathcal{A} + 2\mathcal{B} \neq 0$. Substituting Eq. (4.66) to the action (4.64), we obtain

$$S_{\text{S}}^{(2)} = 2 \int dt d^3x a^3 \left[\frac{\mathcal{B}(3\mathcal{A} + 2\mathcal{B})}{\mathcal{A} + 2\mathcal{B}} \dot{\zeta}^2 + \mathcal{E} \frac{(\partial_k \zeta)^2}{a^2} \right]. \quad (4.67)$$

Written in this form, one notices that the stability condition for the tensor perturbations, $\mathcal{E} > 0$, is not compatible with the stability of the scalar perturbation, $\mathcal{E} < 0$. This

indicates that either of the tensor or scalar perturbations exhibits gradient instabilities, even if one circumvents ghosts by choosing the coefficients in front of the time derivative terms in Eqs. (4.62) and (4.67) to be positive. This result generalizes what was found in Refs. [79, 80, 84], and we have thus established that all the mimetic gravity models with 3 DOFs obtained so far are plagued with ghost/gradient instabilities on a cosmological background (except for the special case of nondynamical scalar perturbations mentioned below). To circumvent this problem, one must design the models so that the timescale of the instability is longer than the age of the Universe. It is worth noting that the terms of the form $\sim RK$ introduced in Refs. [84, 87] to resolve this problem are not fully satisfactory, because now it is clear from the covariant analysis of the present chapter that such terms give rise to unwanted extra DOFs on a general background.

Let us now comment on the Hamiltonian structure of the mimetic theories. As was pointed out in Refs. [67, 76], the Hamiltonian (4.36) evaluated on the constraint surface depends only linearly on p_ϕ , which leads to Ostrogradsky-like instability [88]. The above ghost/gradient instabilities presumably originate from this linear momentum.

It should be noted that the linear momentum can be removed by adding terms like $F_0(\phi, X)$ and $F_1(\phi, X)\square\phi$ to the mimetic Lagrangian. The 3-DOF nature of the extended mimetic theories is not changed by these terms since they do not contribute to the kinetic matrix. It should be noted that such terms generically break the conformal symmetry of the mimetic theories. The only change in the above Hamiltonian analysis is the nature of the primary constraint \mathcal{C} : Its time evolution generates a new secondary constraint (hereafter called \mathcal{D}) and both \mathcal{C} and \mathcal{D} are second class, which is related to the broken conformal symmetry. Since this new constraint \mathcal{D} fixes the linear momentum p_ϕ as a function of other canonical variables, the aforementioned (would-be) Ostrogradsky instability no more appears. Nevertheless, one can verify that the same instability in cosmological perturbations as above persists even if we include $F_0(\phi, X)$ and/or $F_1(\phi, X)\square\phi$ in the mimetic Lagrangian.

One would notice that if $\mathcal{B}(\mathcal{A} + 2\mathcal{B})(3\mathcal{A} + 2\mathcal{B}) = 0$ then the scalar perturbations appear to be nondynamical. This happens in the mimetic Horndeski theories [37] where $\mathcal{A} + 2\mathcal{B} = 0$. The situation is the same even if we start from GLPV theories. Since such a choice of \mathcal{A} and \mathcal{B} does not change the number of DOFs, the seemingly nondynamical scalar mode should originate from strong coupling of perturbations.

A caveat should be added here. In the case of $\mathcal{A} + 2\mathcal{B} = 0$, it is important to take into account the presence of matter fields other than ϕ to discuss the viability of mimetic cosmology. Let us add to the seed Lagrangian another scalar field ψ whose Lagrangian is of the form

$$L_\psi = P(\psi, Z), \quad Z \equiv g^{\mu\nu} \partial_\mu \psi \partial_\nu \psi. \quad (4.68)$$

This field can also be regarded as a perfect fluid. We split ψ into the background part $\psi(t)$ and the perturbation $\delta\psi$, and then expand the mimetic action to second order in

perturbations. In the case of $\mathcal{A} + 2\mathcal{B} = 0$, we obtain

$$S_S^{(2)} = \int dt d^3x a^3 \left[3\mathcal{A}\dot{\zeta}^2 + 2\mathcal{E} \frac{(\partial_k \zeta)^2}{a^2} - 2\mathcal{A}\dot{\zeta} \frac{\partial^2 \chi}{a^2} - (P_Z + 2ZP_{ZZ})\delta\dot{\psi}^2 + P_Z(\partial_k \delta\psi)^2 - 2\dot{\psi}P_Z\delta\psi \frac{\partial^2 \chi}{a^2} + \dots \right], \quad (4.69)$$

where the ellipses represent the terms that are not relevant to the present argument. Now the EOM for χ is given by

$$-2\mathcal{A}\dot{\zeta} - 2\dot{\psi}P_Z\delta\psi = 0. \quad (4.70)$$

Substituting this back into Eq. (4.69), one can remove $\delta\psi$ as well as χ from the action. The reduced action for ζ contains the term

$$a^3 \mathcal{A}^2 \left(\frac{P_Z + 2ZP_{ZZ}}{ZP_Z^2} \right) \ddot{\zeta}^2, \quad (4.71)$$

showing that the system has two scalar DOFs (one from ϕ and one from the additional matter field ψ). This is the reason why the scalar perturbations revive in mimetic Horndeski gravity in the presence of matter [89]. It is more important to note that one of the two scalar DOFs is a ghost, as is clear from Eq. (4.71).

Chapter 5

Conclusions

Given that one can go beyond the Horndeski theory, i.e., the most general scalar-tensor theory with second-order EL equations, it is intriguing to explore general healthy scalar-tensor theories with 3 DOFs by relaxing the assumptions to allow for higher-order EL equations with a degenerate kinetic matrix. As such “beyond Horndeski” theories, the broadest class known so far is the quadratic/cubic DHOST, which cannot be further extended by disformal transformation (1.4). Within the quadratic/cubic DHOST class, only those that can be mapped to the Horndeski class via disformal transformation can accommodate stable cosmology [42].

In the present thesis, we have first clarified the nature of general invertible transformations. We have shown that there is a one-to-one correspondence between solutions in the old frame and those in the new frame if the field transformation is invertible. Our Theorem does not assume any specific form of Lagrangian or transformation law, nor dimensionality of spacetime. On the other hand, if the transformation is noninvertible, it can change the number physical DOFs of the original seed theory. This suggests that one may obtain degenerate scalar-tensor theory by noninvertible disformal transformation of a nondegenerate theory.

Based on this idea, we have demonstrated that the seed action (4.3), which is nondegenerate and thus has an extra DOF in general, can be transformed to give a degenerate theory through the noninvertible conformal transformation (4.4). The unwanted extra DOF is eliminated by the local conformal symmetry associated with the noninvertibility of the transformation, leaving only 3 DOFs, as implied in Ref. [78]. We have shown this explicitly by means of Hamiltonian analysis. The resultant degenerate scalar-tensor theories obtained thus are novel in the sense that they are related to none of the known healthy theories via disformal transformation (see Figure 5.1) and can be thought of as an extension of mimetic gravity since the original mimetic theory is generated from the Einstein-Hilbert action through the same noninvertible conformal transformation. It should be emphasized that not all nondegenerate scalar-tensor theories can be transformed to mimetic gravity with 3 DOFs. Rather, we have specified the possible form of the seed

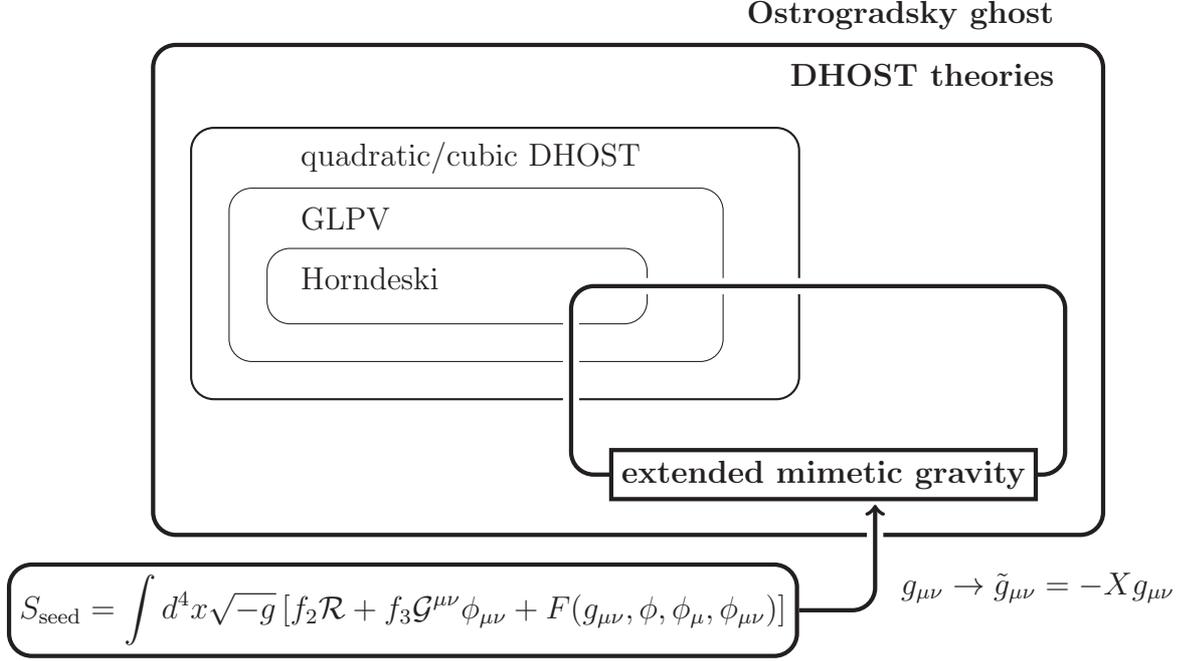


Figure 5.1: The place of the extended mimetic gravity in degenerate higher-order scalar-tensor theories. The extended mimetic gravity models are obtained by performing the noninvertible conformal transformation (4.4) on the seed action (4.3) and they generically lie outside the quadratic/cubic DHOST class.

action as Eq. (4.3). As far as we have investigated, any deviation from this seed leads to unwanted extra DOFs even after the noninvertible transformation.

We have also studied cosmological perturbations in our extended mimetic gravity and found that either of tensor/scalar perturbations is plagued with gradient instabilities in general, except for the special cases where the scalar perturbations would be strongly coupled, or otherwise ghost instabilities appear. In the strongly-coupled case, the inclusion of matter fields other than the scalar field renders the scalar perturbations dynamical and unstable. Our result can be regarded as an extension of the one in Ref. [42], i.e., any quadratic/cubic DHOST theory that cannot be mapped to the Horndeski class via disformal transformation is plagued by gradient instabilities in cosmological perturbations. This is because our extended mimetic gravity models, which were shown to suffer from ghost/gradient instabilities in cosmological perturbations, cannot be obtained from the Horndeski class via disformal transformation in general. Combining these results, one can conjecture that any scalar-tensor theory that is not connected to the Horndeski class through disformal transformation universally has the problem of ghost/gradient instabilities.

In spite of this flaw, the idea of constructing degenerate field theories from nondegenerate higher derivative theories by performing a field transformation is interesting itself and worth pursuing. As we saw in §3.3, the space of possible transformations on scalar-tensor theories is so vast that the disformal transformation (1.4) may capture only a part of its whole aspects. Thus, there would be other types of transformation that can generate a novel class of healthy scalar-tensor theories. Meanwhile, it is not necessary to restrict ourselves to scalar-tensor theories: We expect that, e.g., a noninvertible vector disformal transformation would produce degenerate vector-tensor theories beyond the existing framework. Moreover, once we find a new degenerate theory, we can proceed to a stability analysis of cosmological perturbations and then further investigation into its phenomenology. These issues are left for future works.

Acknowledgments

First, I gratefully acknowledge my supervisor Jun'ichi Yokoyama for offering me an opportunity to spend my Ph.D. years in Research Center for the Early Universe (RESCEU). His insightful comments have led me to a deeper understanding of physics and helped me a lot in starting my research career.

I am grateful to my collaborators, Teruaki Suyama (RESCEU), Tsutomu Kobayashi (Rikkyo University), and Hayato Motohashi (Yukawa Institute for Theoretical Physics), for fruitful discussions. I have acquired not only technical but practical skills in doing research on modified theories of gravity through our collaborations. I am also thankful to all the members of RESCEU.

This work was supported by the Japan Society for the Promotion of Science (JSPS) research fellowships for Young Scientists.

Finally, I would like to express my special thanks to my family. Without their warm and continuous support, I could not have completed this thesis.

Appendices

Appendix A

Theorem of Ostrogradsky

In this appendix, we provide a proof of the theorem of Ostrogradsky [26] in the case of analytical mechanics. Let q^i ($i = 1, \dots, N$) be the coordinate of a point particle moving in n -dimensional space and consider a Lagrangian $L(q^i, \dot{q}^i, \ddot{q}^i)$. The statement of the theorem is as follows:

Theorem of Ostrogradsky. *If the matrix $k_{ij} \equiv \partial^2 L / \partial \ddot{q}^i \partial \ddot{q}^j$ is nondegenerate, then the Hamiltonian of the system contains linear momenta and thus is unbound.*

Proof. We first introduce auxiliary variables Q^i with Lagrange multipliers λ_i to remove the higher derivative from the Lagrangian as follows:^{*1}

$$L'(q^i, \dot{q}^i, Q^i, \dot{Q}^i, \lambda_i) \equiv L(q^i, \dot{q}^i, \ddot{q}^i) + \lambda_i(Q^i - \dot{q}^i). \quad (\text{A.1})$$

This new Lagrangian L' describes the same dynamics of q^i as L since the EOMs for λ_i fix $Q^i = \dot{q}^i$, from which L is reproduced. In the following, L is regarded as a function of $(q^i, \dot{q}^i, \ddot{q}^i)$. Once written in the form (A.1), the nondegeneracy condition reads

$$\det k'_{ij} \neq 0, \quad k'_{ij} \equiv \frac{\partial^2 L(q^k, \dot{q}^k, \ddot{q}^k)}{\partial \ddot{q}^i \partial \ddot{q}^j}. \quad (\text{A.2})$$

Now we move to the Hamiltonian formalism. From the Lagrangian (A.1), the canonical

^{*1}The proof given in Ref. [26] takes a slightly different approach, in which $X^i \equiv q^i, Y^i \equiv \dot{q}^i$ are regarded as canonical variables. Compared with the approach employed in the main text, it has an advantage that we do not need to introduce Lagrange multipliers, though the definition of canonical momenta is rather nontrivial: They are defined as

$$P_i^X \equiv \frac{\partial L}{\partial \dot{q}^i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \ddot{q}^i} \right), \quad P_i^Y \equiv \frac{\partial L}{\partial \ddot{q}^i}.$$

momenta conjugate to q^i , Q^i , and λ_i are computed as

$$p_i \equiv \frac{\partial L'}{\partial \dot{q}^i} = -\lambda_i, \quad (\text{A.3})$$

$$P_i \equiv \frac{\partial L'}{\partial \dot{Q}^i} = \frac{\partial L(q^j, Q^j, \dot{Q}^j)}{\partial \dot{Q}^i}, \quad (\text{A.4})$$

$$\pi^i \equiv \frac{\partial L'}{\partial \dot{\lambda}_i} = 0, \quad (\text{A.5})$$

respectively. These canonical pairs form a $(6N)$ -dimensional phase space. Due to the nondegeneracy condition (A.2), one can apply the implicit function theorem to solve Eq. (A.4) for the velocities \dot{Q}^i :

$$\dot{Q}^i = \dot{Q}^i(q^j, Q^j, P_j). \quad (\text{A.6})$$

On the other hand, since Eqs. (A.3) and (A.5) do not contain \dot{q}^i or $\dot{\lambda}_i$, these velocities cannot be expressed in terms of other canonical variables. Thus, they provide primary constraints as

$$\phi_i \equiv p_i + \lambda_i \approx 0, \quad \pi^i \approx 0. \quad (\text{A.7})$$

Then, the total Hamiltonian becomes

$$H_T = H + u^i \phi_i + v_i \pi^i, \quad H \equiv P_i \dot{Q}^i - L(q^i, Q^i, \dot{Q}^i) - \lambda_i Q^i, \quad (\text{A.8})$$

where u^i and v_i are Lagrange multipliers. Note that \dot{Q}^i in the expression for H should be regarded as functions of (q^j, Q^j, P_j) through Eq. (A.6). Requiring that the primary constraints be maintained under the time evolution, we have

$$0 \approx \dot{\phi}_i = \{\phi_i, H_T\}_P = \frac{\partial L}{\partial q^i} + v^i, \quad (\text{A.9})$$

$$0 \approx \dot{\pi}^i = \{\pi^i, H_T\}_P = Q^i - u^i, \quad (\text{A.10})$$

which fix u^i and v^i and no further constraint arises. Therefore, we have $2N$ second-class constraints and the number of physical DOFs is given by

$$\frac{1}{2}(6N - 2N) = 2N, \quad (\text{A.11})$$

from which we see that the theory has N more DOFs than the system without higher derivatives. The Hamiltonian evaluated on the constraint surface is

$$H \approx P_i \dot{Q}^i - L(q^i, Q^i, \dot{Q}^i) + p_i Q^i, \quad (\text{A.12})$$

where the momenta p_i appear only linearly. \square

Thus, to avoid the linear momenta, the matrix k_{ij} must be degenerate. There is a corresponding argument in the language of Lagrangian formalism. The EL equations from the Lagrangian $L(q^i, \dot{q}^i, \ddot{q}^i)$ are

$$\mathcal{E}_i \equiv \frac{\partial L}{\partial q^i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right) + \frac{d^2}{dt^2} \left(\frac{\partial L}{\partial \ddot{q}^i} \right) = 0, \quad (\text{A.13})$$

which contain fourth time derivatives:

$$\mathcal{E}_i \supset \frac{\partial^2 L}{\partial \ddot{q}^i \partial \ddot{q}^j} q^{(4)j} = k_{ij} q^{(4)j}. \quad (\text{A.14})$$

Hence, if k_{ij} is nondegenerate, the EL equations necessarily acquire such higher derivative terms. Conversely, if k_{ij} is degenerate, (at least some of) the EL equations can be recomposed to yield lower-order differential equations. As such, there is a close relation between Hamiltonian and Lagrangian analyses of degenerate theories.

In the above proof, we restricted ourselves to the case where the Lagrangian depends on at most second derivative of q^i . This is because all the models mentioned in the main text are of this type. For more generic higher derivative Lagrangian $L(q^i, \dot{q}^i, \dots, q^{(n)i})$, one can prove the following theorem in a similar manner as above:

Theorem. *If the matrix $k_{ij} \equiv \partial^2 L / \partial q^{(n)i} \partial q^{(n)j}$ is nondegenerate, then the Hamiltonian of the system contains linear momenta and thus is unbound.*

Appendix B

Disformal transformation

B.1 Disformal transformation of scalar-tensor theories

Here, we study how scalar-tensor theories of the form

$$S = \int d^4x \sqrt{-\tilde{g}} \left[\tilde{f}_2(\phi, \tilde{X}) \tilde{\mathcal{R}} + \tilde{f}_3(\phi, \tilde{X}) \tilde{\mathcal{G}}^{\mu\nu} \tilde{\nabla}_\mu \tilde{\nabla}_\nu \phi + \tilde{F}(\tilde{g}_{\mu\nu}, \phi, \tilde{\nabla}_\mu \phi, \tilde{\nabla}_\mu \tilde{\nabla}_\nu \phi) \right] \quad (\text{B.1})$$

are transformed under the disformal transformation

$$\tilde{g}_{\mu\nu} = A(\phi, X) g_{\mu\nu} + B(\phi, X) \phi_\mu \phi_\nu. \quad (\text{B.2})$$

Note that the quadratic/cubic DHOST action (2.20) amounts to a particular case where the function \tilde{F} is at most cubic order in the second derivative of ϕ . First, if $A(A+XB) \neq 0$, the inverse matrix $\tilde{g}^{\mu\nu}$ is written in terms of $(g_{\mu\nu}, \phi)$ as^{*1}

$$\tilde{g}^{\mu\nu} = \frac{1}{A} \left(g^{\mu\nu} - \frac{B}{A+XB} \phi^\mu \phi^\nu \right), \quad (\text{B.3})$$

which relates \tilde{X} to X as

$$\tilde{X} = \frac{X}{A+XB}. \quad (\text{B.4})$$

As for the determinant of $\tilde{g}_{\mu\nu}$, the following identity between 5×5 matrices is useful:

$$\begin{bmatrix} \delta_\lambda^\mu & 0 \\ \frac{B}{A} \phi^\mu & 1 \end{bmatrix} \begin{bmatrix} \tilde{g}_{\mu\nu} & \phi_\mu \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \delta_\sigma^\nu & 0 \\ -B \phi_\sigma & 1 \end{bmatrix} = \begin{bmatrix} A g_{\lambda\sigma} & \phi_\lambda \\ 0 & \frac{A+XB}{A} \end{bmatrix}. \quad (\text{B.5})$$

Taking the determinant of both sides, we have

$$\tilde{g} = A^3 (A+XB) g, \quad (\text{B.6})$$

^{*1}In the following, the indices of ϕ_μ and $\phi_{\mu\nu}$ are raised by $g_{\mu\nu}$.

which can be more conveniently expressed as

$$\frac{\sqrt{-\tilde{g}}}{\sqrt{-g}} = A^{3/2} \sqrt{A + XB}. \quad (\text{B.7})$$

Here, it should be noted that $A > 0$ and $A + XB > 0$ are required so that the metric signature does not change [32], and thus the square roots of A and $A + XB$ are defined within the real numbers.

The second derivative of ϕ is written in the form

$$\tilde{\nabla}_\mu \tilde{\nabla}_\nu \phi = \phi_{\mu\nu} - C^\lambda{}_{\mu\nu} \phi_\lambda, \quad (\text{B.8})$$

where $C^\lambda{}_{\mu\nu}$ is a tensor defined by the difference between the Christoffel symbols [67]:

$$\begin{aligned} C^\lambda{}_{\mu\nu} &\equiv \tilde{\Gamma}^\lambda{}_{\mu\nu} - \Gamma^\lambda{}_{\mu\nu} = \tilde{g}^{\lambda\sigma} \left(\nabla_{(\mu} \tilde{g}_{\nu)\sigma} - \frac{1}{2} \nabla_\sigma \tilde{g}_{\mu\nu} \right) \\ &= \frac{A_X}{A} \left[\delta^\lambda_{(\mu} X_{\nu)} - \frac{1}{2} X^\lambda g_{\mu\nu} + \frac{B}{A + XB} \phi^\lambda \left(-\phi_{(\mu} X_{\nu)} + \frac{1}{2} \phi^\sigma X_\sigma g_{\mu\nu} \right) \right] \\ &\quad + \frac{B_X}{A} \left[-\frac{1}{2} \phi_\mu \phi_\nu X^\lambda + \frac{A}{A + XB} \phi^\lambda \left(\phi_{(\mu} X_{\nu)} + \frac{B}{2A} \phi_\mu \phi_\nu \phi^\sigma X_\sigma \right) \right] \\ &\quad + \frac{A_\phi}{2A} \left[2\delta^\lambda_{(\mu} \phi_{\nu)} - \frac{1}{A + XB} \phi^\lambda (A g_{\mu\nu} + 2B \phi_\mu \phi_\nu) \right] \\ &\quad + \frac{B}{A + XB} \phi^\lambda \left(\phi_{\mu\nu} + \frac{B_\phi}{2B} \phi_\mu \phi_\nu \right). \end{aligned} \quad (\text{B.9})$$

Here $X_\mu \equiv \nabla_\mu X = 2\phi'_\mu \phi_\nu$ and thus $C^\lambda{}_{\mu\nu}$ depends at most linearly on $\phi_{\mu\nu}$. Note also the following identities:

$$\nabla_\lambda \tilde{g}_{\mu\nu} = 2\tilde{g}_{\alpha(\mu} C^\alpha{}_{\nu)\lambda} \quad \tilde{\nabla}_\lambda g_{\mu\nu} = -2g_{\alpha(\mu} C^\alpha{}_{\nu)\lambda}. \quad (\text{B.10})$$

From these expressions one can deduce two things about the term \tilde{F} in Eq. (B.1): (i) It is stable under disformal transformations since no third covariant derivative of ϕ appears. (ii) The order of second derivative of ϕ does not change under disformal transformations.

Let us next consider the terms with curvature tensors in Eq. (B.1). With the tensor $C^\lambda{}_{\mu\nu}$, the Riemann tensor is calculated as

$$\begin{aligned} \tilde{\mathcal{R}}^\mu{}_{\nu\lambda\sigma} &= \mathcal{R}^\mu{}_{\nu\lambda\sigma} + 2\nabla_{[\lambda} C^\mu{}_{\sigma]\nu} + 2C^\mu{}_{\alpha[\lambda} C^\alpha{}_{\sigma]\nu} \\ &= \mathcal{R}^\mu{}_{\nu\lambda\sigma} + 2\tilde{\nabla}_{[\lambda} C^\mu{}_{\sigma]\nu} - 2C^\mu{}_{\alpha[\lambda} C^\alpha{}_{\sigma]\nu}. \end{aligned} \quad (\text{B.11})$$

Therefore, the term $\tilde{f}_2 \tilde{\mathcal{R}}$ transforms as [67]

$$\begin{aligned} \int d^4x \sqrt{-\tilde{g}} \tilde{f}_2 \tilde{\mathcal{R}} &= \int d^4x \sqrt{-\tilde{g}} \tilde{g}^{\mu\nu} \left[\tilde{f}_2 (\mathcal{R}_{\mu\nu} - 2C^\lambda{}_{\alpha[\lambda} C^\alpha{}_{\nu]\mu}) - 2 \left(\tilde{\nabla}_{[\lambda} \tilde{f}_2 \right) C^\lambda{}_{\nu]\mu} \right] \\ &= \int d^4x \sqrt{-g} \sqrt{A(A+XB)} \left(g^{\mu\nu} - \frac{B}{A+XB} \phi^\mu \phi^\nu \right) \\ &\quad \times \left[\tilde{f}_2 (\mathcal{R}_{\mu\nu} - 2C^\lambda{}_{\alpha[\lambda} C^\alpha{}_{\nu]\mu}) - 2 \left(\tilde{f}_{2\phi} \phi_{[\lambda} + \tilde{f}_{2X} X_{[\lambda} \right) C^\lambda{}_{\nu]\mu} \right]. \end{aligned} \quad (\text{B.12})$$

Note that in the last line \tilde{f}_2 is regarded as a function of (ϕ, X) through Eq. (B.4). In Eq. (B.12), the only term that does not appear in Eq. (B.1) is $\mathcal{R}_{\mu\nu} \phi^\mu \phi^\nu$. However, from the identity

$$\mathcal{R}_{\mu\nu} \phi^\mu \phi^\nu = (\square\phi)^2 - \phi_\mu^\nu \phi_\nu^\mu + \nabla_\mu (\phi^\nu \phi_\nu^\mu - \phi^\mu \square\phi), \quad (\text{B.13})$$

we find that this term can be absorbed in Eq. (B.1) after integration by parts. Note also that the contribution from the term $\tilde{f}_2 \tilde{\mathcal{R}}$ is at most quadratic in $\phi_{\mu\nu}$. The calculation for the term $\tilde{f}_3 \tilde{\mathcal{G}}^{\mu\nu} \tilde{\nabla}_\mu \tilde{\nabla}_\nu \phi$ is tedious and we do not reproduce it here, but one can check that this term is also closed under disformal transformation and the resultant expression is at most cubic order in $\phi_{\mu\nu}$. Combining these results, one can conclude that the quadratic/cubic DHOST action is closed under the disformal transformation (B.2), and so does the action (B.1).

Note that we did not use the invertibility of the disformal transformation in deriving the above result and thus it still holds for the case of noninvertible disformal transformations.

B.2 Noninvertible disformal transformation

In this section, we show that any noninvertible disformal transformation can essentially be reduced to the simplest form of Eq. (4.4). Let us consider a disformal transformation (B.2) with $B \neq 0$. This transformation is invertible [namely, Eq. (B.2) is uniquely solvable for $g_{\mu\nu}$] if the functions A and B satisfy $A(A - XA_X - X^2 B_X) \neq 0$ [31, 43, 53]. Below we study *noninvertible* transformations with

$$A - XA_X - X^2 B_X = 0, \quad A(A + XB) \neq 0, \quad (\text{B.14})$$

where the latter condition guarantees the existence of the inverse matrix $\tilde{g}^{\mu\nu}$ [31]. In the language of the derivative-operator-valued matrix \hat{P} in §3.3.1, the noninvertibility of the transformation (B.2) amounts to the existence of a “zero-eigenvalue” of \hat{P} . To see this, we note that the matrix \hat{P} in Eq. (3.79) satisfy

$$\hat{P}v = (A - XA_X - X^2 B_X)v, \quad v \equiv \begin{bmatrix} A_X g_{\alpha\beta} + B_X \phi_\alpha \phi_\beta \\ 0 \end{bmatrix}, \quad (\text{B.15})$$

where v plays the role of an ‘‘eigenvector’’ of \hat{P} . Now it is clear that the eigenvalue $(A - XA_X - X^2B_X)$ is zero under the condition (B.14).

In the following, we demonstrate any noninvertible disformal transformation with the condition (B.14) can be reduced to a noninvertible *conformal* transformation. Equation (B.14) is equivalent to

$$B = -\frac{A}{X} - f(\phi), \quad (\text{B.16})$$

with $f(\phi)$ being some nonzero function of ϕ . As in Eq. (4.18), the scalar kinetic term in the new frame, \tilde{X} , is constrained as $\tilde{X} = -1/f$. Note that A is proportional to X if there is no disformal part B , i.e., in the case of noninvertible conformal transformation.

Now let us consider another disformal transformation

$$g_{\mu\nu} = \bar{A}(\phi, \bar{X})\bar{g}_{\mu\nu} + \bar{B}(\phi, \bar{X})\phi_\mu\phi_\nu. \quad (\text{B.17})$$

Suppose that a theory $\tilde{S}[\tilde{g}_{\mu\nu}, \phi]$ is mapped to another theory of the form $S[g_{\mu\nu}, \phi]$ by the noninvertible disformal transformation (B.2), and then to $\bar{S}[\bar{g}_{\mu\nu}, \phi]$ by the transformation (B.17). We choose the functions \bar{A} and \bar{B} in Eq. (B.17) so that the composition of the transformations (B.2) and (B.17) reduces to a noninvertible conformal transformation. In doing so we require that the transformation (B.17) is invertible. The following choice of \bar{A} and \bar{B} satisfies these requirements:

$$\bar{A}(\phi, \bar{X}) \equiv \bar{Q} - \bar{X}\bar{B}(\phi, \bar{X}), \quad \bar{B}(\phi, \bar{X}) \equiv -\frac{B(\phi, \bar{X}/\bar{Q})}{A(\phi, \bar{X}/\bar{Q})}. \quad (\text{B.18})$$

To ensure the invertibility of the transformation (B.17), \bar{Q} must not be of the form $\bar{Q} = q(\phi)\bar{X}$ [with arbitrary $q(\phi)$], but otherwise it is an arbitrary function of (ϕ, \bar{X}) . For the choice (B.18), the relation between $\tilde{g}_{\mu\nu}$ and $\bar{g}_{\mu\nu}$ is found to be

$$\tilde{g}_{\mu\nu} = -f(\phi)\bar{X}\bar{g}_{\mu\nu}. \quad (\text{B.19})$$

We see that the disformal part has been eliminated. It should be noted that Eq. (B.19) is independent of the function \bar{Q} . Once written in this form, the function $f(\phi)$ can be absorbed into the redefinition of the scalar field in the following way: Introducing a new scalar field $\hat{\phi}$ so that $d\hat{\phi}/d\phi = f(\phi)^{1/2}$, Eq. (B.19) is rewritten as

$$\tilde{g}_{\mu\nu} = -\left(\bar{g}^{\alpha\beta}\partial_\alpha\hat{\phi}\partial_\beta\hat{\phi}\right)\bar{g}_{\mu\nu}, \quad (\text{B.20})$$

which has the same form as Eq. (4.4). Thus, instead of $S[g_{\mu\nu}, \phi]$ which is obtained from $\tilde{S}[\tilde{g}_{\mu\nu}, \phi]$ by the noninvertible disformal transformation (B.2), we may consider $\bar{S}[\bar{g}_{\mu\nu}, \phi(\hat{\phi})]$ obtained by the noninvertible conformal transformation (B.20).

B.3 Multi-disformal transformation

A multi-disformal transformation for \mathcal{N} -scalar theories is defined by [71]

$$\tilde{g}_{\mu\nu} = A(\phi^I, X^{IJ})g_{\mu\nu} + B_{KL}(\phi^I, X^{IJ})\phi_\mu^K\phi_\nu^L, \quad (\text{B.21})$$

where

$$X^{IJ} \equiv g^{\mu\nu}\phi_\mu^I\phi_\nu^J. \quad (\text{B.22})$$

Note that the coefficients B_{KL} are symmetric under interchange of the indices: $B_{KL} = B_{LK}$. If $A \neq 0$ and the matrix determinant of $D_J^I \equiv A\delta_J^I + X^{IK}B_{KJ}$ is nonvanishing, then the inverse matrix $\tilde{g}^{\mu\nu}$ is written in terms of $(g_{\mu\nu}, \phi)$ as

$$\tilde{g}^{\mu\nu} = \frac{1}{A} \left[g^{\mu\nu} - (D^{-1})_K^M B_{ML}\phi^{K\mu}\phi^{L\nu} \right], \quad (\text{B.23})$$

which relates \tilde{X}^{IJ} to X^{IJ} as

$$\tilde{X}^{IJ} = (D^{-1})_K^I X^{KJ}. \quad (\text{B.24})$$

As for the determinant of $\tilde{g}_{\mu\nu}$, the following identity between $(4 + \mathcal{N}) \times (4 + \mathcal{N})$ matrices is useful:

$$\begin{bmatrix} \delta_\lambda^\mu & 0 \\ \frac{1}{A}B_{JM}\phi^{M\mu} & \delta_J^I \end{bmatrix} \begin{bmatrix} \tilde{g}_{\mu\nu} & \phi_\mu^K \\ 0 & \delta_I^K \end{bmatrix} \begin{bmatrix} \delta_\sigma^\nu & 0 \\ -B_{KN}\phi_\sigma^N & \delta_K^L \end{bmatrix} = \begin{bmatrix} Ag_{\lambda\sigma} & \phi_\lambda^L \\ 0 & \frac{1}{A}D_J^L \end{bmatrix}. \quad (\text{B.25})$$

Taking the determinant of both sides and then their square root, we obtain

$$\frac{\sqrt{-\tilde{g}}}{\sqrt{-g}} = A^{2-\frac{\mathcal{N}}{2}} \sqrt{\det D}. \quad (\text{B.26})$$

The above equations are helpful in studying how multi-field scalar-tensor theories are transformed under the multi-disformal transformation (B.21).

Next, let us consider the inverse transformation of the multi-disformal transformation (B.21). We define

$$\begin{aligned} \mathcal{J}^{IJ}{}_{KL} &\equiv D_N^I \frac{\partial}{\partial X^{KL}} \left[(D^{-1})_M^N X^{MJ} \right] \\ &= \delta^{IJ}{}_{KL} - \frac{\partial D_N^I}{\partial X^{KL}} (D^{-1})_M^N X^{MJ}. \end{aligned} \quad (\text{B.27})$$

This quantity can be regarded as a matrix with indices IJ and KL which move from 1 to $\frac{\mathcal{N}(\mathcal{N}+1)}{2}$. If the determinant of the matrix $\mathcal{J}^{IJ}{}_{KL}$ is nonvanishing, then one can solve Eq. (B.24) for X^{IJ} in terms of \tilde{X}^{IJ} and the inverse multi-disformal transformation is given by

$$g_{\mu\nu} = \tilde{A}(\phi^I, \tilde{X}^{IJ})\tilde{g}_{\mu\nu} + \tilde{B}_{KL}(\phi^I, \tilde{X}^{IJ})\phi_\mu^K\phi_\nu^L, \quad (\text{B.28})$$

where

$$\tilde{A}(\phi^I, \tilde{X}^{IJ}) = \frac{1}{A(\phi^I, X^{IJ})}, \quad \tilde{B}_{KL}(\phi^I, \tilde{X}^{IJ}) = -\frac{B_{KL}(\phi^I, X^{IJ})}{A(\phi^I, X^{IJ})}. \quad (\text{B.29})$$

B.4 Vector disformal transformation

In the context of vector-tensor theories, we can consider a similar transformation as the disformal transformation in scalar-tensor theories. Here, we study the nature of the vector disformal transformation introduced in Ref. [23]. Let us consider a transformation

$$\tilde{g}_{\mu\nu} = \Omega(Y)g_{\mu\nu} + \Gamma(Y)A_\mu A_\nu, \quad \tilde{A}_\mu = \Upsilon(Y)A_\mu, \quad (\text{B.30})$$

where $Y \equiv g^{\mu\nu}A_\mu A_\nu$. Note that the vector field is also transformed in contrast to the case of scalar disformal transformation. For the inverse metric to exist, it is necessary that

$$\Omega\Xi \neq 0, \quad \Xi \equiv \Gamma + \frac{\Omega}{Y}. \quad (\text{B.31})$$

If this condition is met, the inverse metric $\tilde{g}^{\mu\nu}$ and the contravariant component of the vector field \tilde{A}^μ are given by [23]

$$\tilde{g}^{\mu\nu} = \frac{1}{\Omega} \left(g^{\mu\nu} - \frac{\Gamma}{Y\Xi} A^\mu A^\nu \right), \quad \tilde{A}^\mu = \frac{\Upsilon}{Y\Xi} A^\mu. \quad (\text{B.32})$$

Moreover, one can verify that the transformation (B.30) is invertible if the functions Ω , Γ , and Υ satisfy

$$\Omega\Upsilon\Xi \neq 0 \quad \text{and} \quad \frac{\Upsilon^2}{\Xi} \neq \text{constant}. \quad (\text{B.33})$$

Indeed, the inverse transformation of Eq. (B.30) is given by

$$g_{\mu\nu} = \frac{1}{\Omega(Y)} \left[\tilde{g}_{\mu\nu} - \frac{\Gamma(Y)}{\Upsilon^2(Y)} \tilde{A}_\mu \tilde{A}_\nu \right], \quad A_\mu = \frac{1}{\Upsilon(Y)} \tilde{A}_\mu, \quad \tilde{Y} = \frac{\Upsilon^2(Y)}{\Xi(Y)}, \quad (\text{B.34})$$

where Y in the first two equations is written in terms of \tilde{Y} through the third equation. Note that the second condition in Eq. (B.33) ensures the solvability of Y for \tilde{Y} . On the other hand, we call the transformation (B.30) noninvertible if

$$\Omega\Upsilon\Xi \neq 0 \quad \text{and} \quad \frac{\Upsilon^2}{\Xi} = \text{constant}. \quad (\text{B.35})$$

As was the case with disformal transformation for scalar-tensor theories, any noninvertible disformal transformation for vector-tensor theories with the conditions (B.35) can be reduced to a simpler form. Now let us start from the transformation (B.30) with $\Gamma \neq 0$ and $\Upsilon \neq 1$ and consider another disformal transformation

$$g_{\mu\nu} = \bar{\Omega}(\bar{Y})\bar{g}_{\mu\nu} + \bar{\Gamma}(\bar{Y})\bar{A}_\mu\bar{A}_\nu, \quad A_\mu = \bar{\Upsilon}(\bar{Y})\bar{A}_\mu. \quad (\text{B.36})$$

As in the previous section, we choose the functions $\bar{\Omega}$, $\bar{\Gamma}$, and $\bar{\Upsilon}$ in Eq. (B.36) so that the composition of the transformations (B.30) and (B.36) reduces to a noninvertible conformal transformation for the metric while the vector field remains unchanged, namely,

$$\tilde{g}_{\mu\nu} = -\bar{Y}\bar{g}_{\mu\nu}, \quad \tilde{A}_\mu = \bar{A}_\mu. \quad (\text{B.37})$$

In doing so, we require that the transformation (B.36) is invertible. The following choice satisfies these requirements:

$$\bar{\Omega}(\bar{Y}) \equiv \bar{Q}\bar{\Upsilon}^2 - \bar{Y}\bar{\Gamma}, \quad \bar{\Gamma} \equiv -\frac{\Gamma(\bar{Y}/\bar{Q})}{\Omega(\bar{Y}/\bar{Q})\Upsilon^2(\bar{Y}/\bar{Q})}, \quad \bar{\Upsilon} \equiv \frac{1}{\Upsilon(\bar{Y}/\bar{Q})}, \quad (\text{B.38})$$

where \bar{Q} is an arbitrary function of \bar{Y} . To ensure the invertibility of the transformation (B.36), \bar{Q} must not be of the form $\bar{Q} \propto \bar{Y}$, but otherwise it is an arbitrary function of \bar{Y} . For the choice (B.38), the relation between $(\tilde{g}_{\mu\nu}, \tilde{A}_\mu)$ and $(\bar{g}_{\mu\nu}, \bar{A}_\mu)$ is found to be

$$\tilde{g}_{\mu\nu} = -\left(\frac{-\Xi}{\Upsilon^2}\right)\bar{Y}\bar{g}_{\mu\nu}, \quad \tilde{A}_\mu = \bar{A}_\mu. \quad (\text{B.39})$$

We see that the disformal part has been eliminated. Note that the constant factor $-\Xi/\Upsilon^2$ can be absorbed into the redefinition of the vector field. Thus, we can always recast a noninvertible vector disformal transformation to the simplest form (B.37).

Appendix C

Unique solvability of DAEs

In Chap. 3, we saw that the invertibility of a transformation between D -dimensional fields ϕ^i and ψ^j , $\phi^i = f^i(\psi^j, \partial_\mu \psi^j, \partial_\mu \partial_\nu \psi^j, \dots, \partial_{(m)} \psi^j)$, is characterized by the *unique solvability* of the associated DAE system

$$\hat{P}_j^i \delta \psi^j = \delta \phi^i, \quad \hat{P}_j^i = \sum_{s=0}^m \frac{\partial f^i}{\partial (\partial_{(s)} \psi^j)} \partial_{(s)} \equiv \sum_{s=0}^m u_j^{i(s)} \partial_{(s)}. \quad (\text{C.1})$$

Here, a DAE system (C.1) is said to be uniquely solvable if and only if the following conditions are satisfied:

- [A] The system is well posed, i.e., it has a solution for any inhomogeneity g^I .
- [B] No ordinary differential equation appears.

In this appendix, following Ref. [63], we discuss the criterion for a system of DAEs to be uniquely solvable. Note that, unless otherwise stated, we consider ordinary DAEs (ODAEs) where the fields are functions only of t . This amounts to the case of analytical mechanics, or $D = 1$. The reason why we focus on ODAEs is that, as shown below, there exists an algorithm to judge whether a given ODAE system is uniquely solvable or not. Although there are some necessary/sufficient conditions for unique solvability in the case of $D \geq 2$, i.e., partial DAEs (PDAEs), it remains an open issue whether a similar algorithm as $D = 1$ exists for $D \geq 2$ to the best of our knowledge.

C.1 Conditions for the unique solvability

For simplicity, we consider a first-order ODAE system of the form

$$\hat{\Pi}_J^I \Psi^J = \Phi^I, \quad \hat{\Pi}_J^I \equiv M_J^I \frac{d}{dt} + N_J^I, \quad (\text{C.2})$$

since any higher-order ODAE system can be recast into a first-order system by introducing enough number of auxiliary variables. We will hereafter denote the number of extended fields Ψ^I and Φ^I by \mathcal{N} .

It is obvious that the requirement [B] cannot be met if $\det M_J^I \neq 0$. We thus obtain a necessary (but not sufficient) condition for the unique solvability of the system:

$$\det M_J^I = 0. \quad (\text{C.3})$$

Moreover, the following condition is also necessary:

$$\det N_J^I \neq 0. \quad (\text{C.4})$$

The reason for this is not so much obvious. Suppose $\det N_J^I = 0$. Then there exists a null eigenvector (hereafter called v_I) of the matrix N_J^I . Multiplying both sides of Eq. (C.2) by v_I , we obtain

$$v_I M_J^I \frac{d\Psi^J}{dt} = v_I \Phi^I. \quad (\text{C.5})$$

If v^I is a null eigenvector of M_J^I , then the left-hand side of Eq. (C.5) vanishes, which contradicts the requirement [A]. On the other hand, if $v_I M_J^I$ is nonvanishing, then an ordinary differential equation appears, which conflicts with [B]. Therefore, the condition (C.4) is necessary for the unique solvability of the DAE system (C.2).

Below we discuss sufficient conditions for the unique solvability. In the case of $\mathcal{N} = 1$, the criterion is trivial, i.e., $M = 0$ and $N \neq 0$. Then, even in the case of $\mathcal{N} > 1$, one may naively think that Ψ^I are uniquely determined only when

$$M_J^I = 0 \quad \text{and} \quad \det N_J^I \neq 0. \quad (\text{C.6})$$

However, as we will see below, this condition is too restrictive. While (C.6) is a sufficient condition, it is not a necessary condition. Another uniquely solvable example is such that

$$M_J^I = K_J^I \quad \text{and} \quad N_J^I = \delta_J^I, \quad (\text{C.7})$$

where K_J^I is a strictly lower (upper) triangular matrix, i.e., all the components of the matrix on and above (below) the diagonal are vanishing. Note that K_J^I can depend on time. If K_J^I is strictly lower triangular, one can first determine Ψ^1 uniquely. One then determines Ψ^2 , as one can treat nonvanishing derivative $\dot{\Psi}^1$ as a source term. Likewise, one can continue to determine all the components of Ψ^I uniquely. On the other hand, if K_J^I is strictly upper triangular, one can start from $\Psi^{\mathcal{N}}$, and proceed to determine $\Psi^{\mathcal{N}-1}, \Psi^{\mathcal{N}-2}, \dots, \Psi^1$ in order. In terms of the operator matrix $\hat{\Pi}$ in (C.2), the case of (C.7) amounts to

$$\hat{\Pi}_J^I = \delta_J^I + K_J^I \frac{d}{dt}. \quad (\text{C.8})$$

Clearly, $\hat{\Pi}$ has an inverse matrix

$$\left(\hat{\Pi}^{-1}\right)_J^I = \delta_J^I + \sum_{s=1}^{N-1} \left[\left(-K \frac{d}{dt}\right)^s \right]_J^I, \quad (\text{C.9})$$

as expected. Indeed, as explained in §3.2, the existence of $\hat{\Pi}^{-1}$ without integral is equivalent to the unique solvability of the corresponding system of equations. Besides the case of (C.7), there are still other forms of (M, N) for which (C.2) is uniquely solvable. For instance, δ_J^I can be relaxed to some diagonal matrix whose diagonal components are all nonvanishing, and then N can be added by any strictly lower (upper) triangular matrix. Therefore, the sufficient condition (C.7) can be generalized as

$$M_J^I = K_J^I \quad \text{and} \quad N_J^I = D_J^I + L_J^I, \quad (\text{C.10})$$

where K_J^I and L_J^I are time-dependent strictly lower (upper) triangular matrices, and D_J^I is a time-dependent regular diagonal matrix.

In summary, we have obtained the necessary conditions (C.3) and (C.4), together with the sufficient conditions (C.6) and (C.10). These conditions can be straightforwardly generalized to the case of $D \geq 2$, i.e., to the PDAE system (C.1). The necessary conditions (C.3) and (C.4) are respectively generalized as

$$\forall s \in \{1, \dots, m\} \quad \det u_j^{i(s)} = 0 \quad (\text{C.11})$$

and

$$\det u_j^{i(0)} \neq 0, \quad (\text{C.12})$$

while the sufficient conditions (C.6) and (C.10) as

$$\det u_j^{i(0)} \neq 0 \quad \text{and} \quad \forall s \in \{1, \dots, m\} \quad u_j^{i(s)} = 0 \quad (\text{C.13})$$

and

$$u_j^{i(0)} = D_j^i + L_j^i \quad \text{and} \quad \forall s \in \{1, \dots, m\} \quad u_j^{i(s)} = K_j^{i(s)}. \quad (\text{C.14})$$

Here, $K_j^{i(s)}$ and L_j^i are spacetime-dependent strictly lower (upper) triangular matrices, and D_j^i is a spacetime-dependent regular diagonal matrix.

C.2 Standard canonical form of ODAEs

A necessary and sufficient condition for the unique solvability of ODAEs is rather non-trivial [90]. The idea is to recast the pair (M_J^I, N_J^I) into the form of Eq. (C.10) by transformation of variables with some regular matrices S_J^I and T_J^I . Let us multiply S_J^I by both sides of Eq. (C.2) and write

$$S_K^I M_L^K T_J^L \dot{\Psi}^J + (S_K^I N_L^K T_J^L + S_K^I M_L^K \dot{T}_J^L) \tilde{\Psi}^J = \tilde{\Phi}^I, \quad (\text{C.15})$$

where $\tilde{\Psi}^I$ and $\tilde{\Phi}^I$ are defined as

$$\tilde{\Psi}^I \equiv (T^{-1})^I_J \Psi^J, \quad \tilde{\Phi}^I \equiv S^I_J \Phi^J. \quad (\text{C.16})$$

Therefore, if we define \tilde{M}^I_J and \tilde{N}^I_J by

$$\tilde{M}^I_J \equiv S^I_K M^K_L T^L_J, \quad \tilde{N}^I_J \equiv S^I_K N^K_L T^L_J + S^I_K M^K_L \dot{T}^L_J, \quad (\text{C.17})$$

Eq. (C.15) becomes

$$\tilde{M}^I_J \dot{\tilde{\Psi}}^J + \tilde{N}^I_J \tilde{\Psi}^J = \tilde{\Phi}^I, \quad (\text{C.18})$$

which has the same form as Eq. (C.2). It was shown in Ref. [90] that, one can always choose S^I_J and T^I_J so that the pair of matrices $(\tilde{M}^I_J, \tilde{N}^I_J)$ takes the following ‘‘standard canonical form’’ (SCF):

$$(\tilde{M}, \tilde{N}) = \left(\begin{bmatrix} I_{n_1} & 0 \\ 0 & K_{n_2}(t) \end{bmatrix}, \begin{bmatrix} J_{n_1}(t) & 0 \\ 0 & I_{n_2} \end{bmatrix} \right), \quad (\text{C.19})$$

if and only if the system is well posed. Here, I_{n_i} denotes an $n_i \times n_i$ identity matrix, $K_{n_2}(t)$ is an $n_2 \times n_2$ matrix which is strictly lower triangular, $J_{n_1}(t)$ is some $n_1 \times n_1$ matrix, and $n_1 + n_2 = \mathcal{N}$. From the block-diagonal structure of Eq. (C.19), it is clear that in Eq. (C.18) the equations for the first n_1 variables and the last n_2 are decoupled. If $n_1 \neq 0$, since the upper-left $n_1 \times n_1$ submatrix of \tilde{M}^I_J is the identity matrix, the first n_1 equations are inevitably ordinary differential equations and thus the unique solvability of the system is spoiled. If $n_1 = 0$, one is left with equations of the form

$$\begin{bmatrix} 0 & \cdots & \cdots & 0 \\ * & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ * & \cdots & * & 0 \end{bmatrix} \begin{bmatrix} \dot{\tilde{\Psi}}^1 \\ \vdots \\ \vdots \\ \dot{\tilde{\Psi}}^{\mathcal{N}} \end{bmatrix} + \begin{bmatrix} \tilde{\Psi}^1 \\ \vdots \\ \vdots \\ \tilde{\Psi}^{\mathcal{N}} \end{bmatrix} = \begin{bmatrix} \tilde{\Phi}^1 \\ \vdots \\ \vdots \\ \tilde{\Phi}^{\mathcal{N}} \end{bmatrix}. \quad (\text{C.20})$$

This precisely satisfies the sufficient condition (C.10). We can uniquely solve this ODAE system for $\tilde{\Psi}^I$ from the first-line equation to the \mathcal{N} th-line equation without any integration constant, and then obtain Ψ^I through $\Psi^I = T^I_J \tilde{\Psi}^J$.

In conclusion, the necessary and sufficient condition for the unique solvability of the DAE system (C.2) with a matrix pair (M, N) is that the corresponding SCF (C.19) has $n_1 = 0$, namely,

$$(\tilde{M}, \tilde{N}) = (K_{\mathcal{N}}, I_{\mathcal{N}}). \quad (\text{C.21})$$

Obviously, for a vanishing source term Φ^I , the unique solution is given by $\Psi^I = 0$.

As an application of the above methodology, let us consider the following ODAE system:

$$\begin{bmatrix} 2 & t+2 & -t-1 \\ -2t & -t(t+2) & t(t+1) \\ 2t & t(t+1) & -t^2 \end{bmatrix} \begin{bmatrix} \dot{\Psi}^1 \\ \dot{\Psi}^2 \\ \dot{\Psi}^3 \end{bmatrix} + \begin{bmatrix} 1 & -t+1 & t \\ -t+2 & t^2+1 & -t(t+1) \\ 0 & t+1 & -t-1 \end{bmatrix} \begin{bmatrix} \Psi^1 \\ \Psi^2 \\ \Psi^3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad (\text{C.22})$$

which corresponds to the case of

$$(M, N) = \left(\begin{bmatrix} 2 & t+2 & -t-1 \\ -2t & -t(t+2) & t(t+1) \\ 2t & t(t+1) & -t^2 \end{bmatrix}, \begin{bmatrix} 1 & -t+1 & t \\ -t+2 & t^2+1 & -t(t+1) \\ 0 & t+1 & -t-1 \end{bmatrix} \right), \quad \Phi^I = 0. \quad (\text{C.23})$$

In this case, we can find the regular transformation matrices S, T as

$$S = \begin{bmatrix} t & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad T = \begin{bmatrix} 1 & -2t-1 & -1 \\ -1 & 3t+2 & 2 \\ -1 & 3t+1 & 2 \end{bmatrix}. \quad (\text{C.24})$$

These matrices actually transform the pair (M, N) into the SCF:

$$(\tilde{M}, \tilde{N}) = \left(\begin{bmatrix} 0 & 0 & 0 \\ t & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right), \quad (\text{C.25})$$

which satisfies the sufficient condition (C.10). This means that the system (C.22) is uniquely solvable and the solution is given by $\Psi^I = 0$.

C.3 Adjoint ODAE

As we saw in §3.2, if a DAE system is uniquely solvable, then its adjoint DAE system is also uniquely solvable. Here, we show the fact for ODAEs in a more direct manner. The adjoint ODAE system to (C.2) has the form of

$$\frac{d}{dt}(\Omega_J M_I^J) - \Omega_J N_I^J = X_I. \quad (\text{C.26})$$

If the ODAE system (C.2) is uniquely solvable, there exists a pair of matrices (S, T) that transforms (M, N) into the form of (C.21). Using the pair (S, T) , we can rewrite (C.26) as

$$\frac{d}{dt}(\tilde{\Omega}_J \tilde{M}_I^J) - \tilde{\Omega}_J \tilde{N}_I^J = \tilde{X}_I, \quad (\text{C.27})$$

where we have defined

$$\tilde{\Omega}_I \equiv \Omega_J (S^{-1})_I^J, \quad \tilde{X}_I \equiv X_J T_I^J. \quad (\text{C.28})$$

More explicitly, (C.27) can be written as

$${}^t K_N \dot{{}^t \tilde{\Omega}} - (I_N - {}^t \dot{K}_N) {}^t \tilde{\Omega} = {}^t \tilde{X}, \quad (\text{C.29})$$

namely,

$$\begin{bmatrix} 0 & * & \cdots & * \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & * \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \begin{bmatrix} \tilde{\Omega}_1 \\ \vdots \\ \vdots \\ \tilde{\Omega}_{\mathcal{N}} \end{bmatrix} - \begin{bmatrix} 1 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{\Omega}_1 \\ \vdots \\ \vdots \\ \tilde{\Omega}_{\mathcal{N}} \end{bmatrix} = \begin{bmatrix} \tilde{X}_1 \\ \vdots \\ \vdots \\ \tilde{X}_{\mathcal{N}} \end{bmatrix}. \quad (\text{C.30})$$

Similarly to (C.20), this system satisfies the sufficient condition (C.10). It can be solved for $\tilde{\Omega}_I$ from the \mathcal{N} th-line equation to the first-line equation without any integration constant. In particular, for a homogeneous system with $X_I = 0$, the unique solution is $\Omega_I = 0$.

Appendix D

Noether identity

In this appendix, we briefly review the concept of Noether identity, which is a consequence of local gauge symmetry. Let us now consider a general field theory defined by the Lagrangian $L = L(\phi^i, \partial_\mu \phi^i, \partial_\mu \partial_\nu \phi^i, \dots; x^\mu)$ with multiple fields $\phi^i = \phi^i(x^\mu)$ in D -dimensional spacetime, which is invariant up to total derivative under a general gauge transformation

$$\phi^i \rightarrow \phi^i + \Delta_\xi \phi^i, \quad (\text{D.1})$$

where $\Delta_\xi \phi^i$ depend on gauge functions $\xi^I(x^\mu)$ and their derivatives. Here, $i = 1, \dots, n$ labels the fields and $I = 1, \dots, m$ labels the gauge symmetries, with $m < n$.

In such a theory with gauge symmetries, there exists an identity between the EOMs, which is known as Noether's second theorem [91]. Let us consider an infinitesimal gauge transformation

$$\phi^i \rightarrow \phi^i + \Delta_\epsilon \phi^i, \quad (\text{D.2})$$

where $\Delta_\epsilon \phi^i$ are linearized as

$$\Delta_\epsilon \phi^i = \sum_{p=0}^k F_I^{i(p)} \partial_{(p)} \epsilon^I. \quad (\text{D.3})$$

Here, we suppress indices for p th-order coefficients and derivative as

$$\begin{aligned} F_I^{i(p)} &\equiv F_I^{i\mu_1 \dots \mu_p}, \\ \partial_{(p)} &\equiv \partial_{\mu_1} \dots \partial_{\mu_p}. \end{aligned} \quad (\text{D.4})$$

Note that $F_I^{i(p)}$ are functions of the fields ϕ^i and their derivatives, and they can also depend explicitly on x^μ . In Eq. (D.3), $p = 0$ term is understood as $F_I^i \epsilon^I$ without derivative, and all the other terms with $p \geq 1$ have p th derivative of the infinitesimal gauge functions ϵ^I . Since the action is invariant under the infinitesimal gauge transformation (D.2), we obtain

$$0 = \Delta_\epsilon S = \int d^D x \mathcal{E}_i \Delta_\epsilon \phi^i. \quad (\text{D.5})$$

Here, $\mathcal{E}_i = 0$ are the EOMs for ϕ^i , i.e., the EL equations derived by the variational principle:

$$\mathcal{E}_i \equiv \frac{\partial L}{\partial \phi^i} - \partial_\mu \left(\frac{\partial L}{\partial (\partial_\mu \phi^i)} \right) + \partial_\mu \partial_\nu \left(\frac{\partial L}{\partial (\partial_\mu \partial_\nu \phi^i)} \right) - \dots \quad (\text{D.6})$$

Plugging Eq. (D.3) into Eq. (D.5) and integration by parts yield

$$0 = \int d^D x \left[\sum_{p=0}^k (-1)^p \partial_{(p)} \left(\mathcal{E}_i F_I^{i(p)} \right) \right] \epsilon^I. \quad (\text{D.7})$$

Since ϵ^I are arbitrary functions, we obtain the following identity (Noether identity) between the EOMs:

$$\sum_{p=0}^k (-1)^p \partial_{(p)} \left(\mathcal{E}_i F_I^{i(p)} \right) = 0, \quad (\text{D.8})$$

for $I = 1, \dots, m$.

Let us remark that one can verify that the gauge transformation of the EOMs \mathcal{E}_i can be written as a linear combination of \mathcal{E}_i and their derivatives. This means that, as it should be, if a configuration of ϕ^i satisfies the EOMs, then its gauge transformation $\phi^i + \Delta_\epsilon \phi^i$ also satisfies the same set of EOMs.

In the context of scalar-tensor theories with general covariance, the argument proceeds as follows. By virtue of the general covariance, the action is invariant under an infinitesimal transformation of coordinates $x^\mu \rightarrow x^\mu + \epsilon^\mu$. The gauge transformation of the metric and the scalar field is then given by

$$\begin{aligned} g_{\mu\nu} &\rightarrow g_{\mu\nu} - \nabla_\mu \epsilon_\nu - \nabla_\nu \epsilon_\mu, \\ \phi &\rightarrow \phi - \epsilon^\mu \nabla_\mu \phi. \end{aligned} \quad (\text{D.9})$$

Indeed, the gauge transformation of the Lagrangian density becomes total derivative:

$$\begin{aligned} \Delta_\epsilon(\sqrt{-g}L) &= (-\sqrt{-g}\nabla_\mu \epsilon^\mu)L + \sqrt{-g}(-\epsilon^\mu \nabla_\mu L) = -\sqrt{-g}\nabla_\mu(\epsilon^\mu L) \\ &= -\partial_\mu(\epsilon^\mu \sqrt{-g}L), \end{aligned} \quad (\text{D.10})$$

where we assumed that L transforms as a scalar. Now we consider the gauge transformation of the action of scalar-tensor theories:

$$\begin{aligned} 0 = \Delta_\epsilon S[g_{\mu\nu}, \phi] &= \int d^D x \sqrt{-g} [E^{\mu\nu}(-2\nabla_\mu \epsilon_\nu) + E_\phi(-\epsilon_\nu \nabla^\nu \phi)] \\ &= \int d^D x \sqrt{-g} [2\nabla_\mu E^{\mu\nu} - E_\phi \nabla^\nu \phi] \epsilon_\nu, \end{aligned} \quad (\text{D.11})$$

where we have defined

$$E^{\mu\nu} \equiv \frac{1}{\sqrt{-g}} \frac{\delta S}{\delta g_{\mu\nu}} = 0, \quad E_\phi \equiv \frac{1}{\sqrt{-g}} \frac{\delta S}{\delta \phi} = 0. \quad (\text{D.12})$$

From Eq. (D.11) the Noether identity can be read off as

$$2\nabla_\mu E^{\mu\nu} - E_\phi \nabla^\nu \phi = 0. \quad (\text{D.13})$$

This means that the scalar EOM $E_\phi = 0$ is redundant if $\nabla_\mu \phi \neq 0$, as was stated in, e.g., Ref. [14].

In Chap. 4, we introduced extended mimetic gravity models, whose action has local conformal symmetry other than general covariance. The action of the extended mimetic gravity is characterized by the part $S_M[N, \gamma_{ij}, \phi, A_*, B_{ij}]$, which is given in Eq. (4.25) (for the detailed notations, see Chap. 4). Below, we summarize the Noether identities associated with the spatial diffeomorphism and the conformal symmetry, which play an important role in the Hamiltonian analysis.

D.1 Spatial diffeomorphism

Under an infinitesimal spatial diffeomorphism $x^i \rightarrow x^i + \epsilon^i(x^j)$, the relevant variables N , γ_{ij} , ϕ , A_* , and B_{ij} transform as

$$\Delta_\epsilon N = -\epsilon^i D_i N, \quad \Delta_\epsilon \gamma_{ij} = -2D_{(i} \epsilon_{j)}, \quad \Delta_\epsilon \phi = -\epsilon^i D_i \phi, \quad (\text{D.14})$$

$$\Delta_\epsilon A_* = -\epsilon^i D_i A_*, \quad \Delta_\epsilon B_{ij} = -\epsilon^k D_k B_{ij} - 2B_{k(i} D_j) \epsilon^k, \quad (\text{D.15})$$

respectively. Note that B_{ij} replaces V_{ij} [see Eq. (4.19)] and therefore transforms in the same way as V_{ij} . Since S_M is invariant under this transformation, we have

$$0 = \Delta_\epsilon S_M = - \int dt d^3x \left[\frac{\delta S_M}{\delta N} \epsilon^i D_i N + 2 \frac{\delta S_M}{\delta \gamma_{ij}} D_i \epsilon_j + \frac{\delta S_M}{\delta \phi} \epsilon^i D_i \phi + \frac{\delta S_M}{\delta A_*} \epsilon^i D_i A_* + \frac{\delta S_M}{\delta B_{ij}} (\epsilon^k D_k B_{ij} + 2B_{ki} D_j \epsilon^k) \right]. \quad (\text{D.16})$$

Integrating by parts, we obtain the following relation among the variations of S_M :

$$\begin{aligned} \sqrt{\gamma} L_M D_i N = & 2\sqrt{\gamma} \gamma_{ij} D_k \left(\frac{1}{\sqrt{\gamma}} \frac{\delta S_M}{\delta \gamma_{jk}} \right) - \frac{\delta S_M}{\delta \phi} D_i \phi - \frac{\delta S_M}{\delta A_*} D_i A_* \\ & - \frac{\delta S_M}{\delta B_{jk}} D_i B_{jk} + 2\sqrt{\gamma} D_j \left(\frac{B_{ik}}{\sqrt{\gamma}} \frac{\delta S_M}{\delta B_{jk}} \right). \end{aligned} \quad (\text{D.17})$$

D.2 Conformal symmetry

Similarly, under an infinitesimal conformal transformation $\Delta_\epsilon g_{\mu\nu} = \epsilon(x^\lambda)g_{\mu\nu}$, we have the following transformation law:

$$\Delta_\epsilon N = \frac{\epsilon}{2}N, \quad \Delta_\epsilon \gamma_{ij} = \epsilon \gamma_{ij}, \quad \Delta_\epsilon \phi = 0, \quad (\text{D.18})$$

$$\Delta_\epsilon A_* = -\frac{\epsilon}{2}A_*, \quad \Delta_\epsilon B_{ij} = \frac{\epsilon}{2}B_{ij} - \frac{D^k \phi}{2A_*} \gamma_{ij} D_k \epsilon. \quad (\text{D.19})$$

These expressions can be obtained by replacing $(-X)$ with $1 + \epsilon$ in Eqs. (4.14) and (4.22). Following the same procedure as in the case of spatial diffeomorphism, we obtain

$$N \sqrt{\gamma} L_M = -\gamma_{ij} \left[2 \frac{\delta S_M}{\delta \gamma_{ij}} + \sqrt{\gamma} D_k \left(\frac{D^k \phi}{\sqrt{\gamma} A_*} \frac{\delta S_M}{\delta B_{ij}} \right) \right] + A_* \frac{\delta S_M}{\delta A_*} - B_{ij} \frac{\delta S_M}{\delta B_{ij}}. \quad (\text{D.20})$$

Bibliography

- [1] C. M. Will, “The Confrontation between General Relativity and Experiment,” *Living Rev. Rel.* **17**, 4 (2014).
- [2] A. G. Riess *et al.* (Supernova Search Team Collaboration), “Observational evidence from supernovae for an accelerating universe and a cosmological constant,” *Astron. J.* **116**, 1009–1038 (1998).
- [3] S. Perlmutter *et al.* (Supernova Cosmology Project Collaboration), “Measurements of Omega and Lambda from 42 high redshift supernovae,” *Astrophys. J.* **517**, 565–586 (1999).
- [4] B. P. Abbott *et al.* (Virgo, LIGO Scientific Collaboration), “Observation of Gravitational Waves from a Binary Black Hole Merger,” *Phys. Rev. Lett.* **116**, 061102 (2016).
- [5] B. P. Abbott *et al.* (Virgo, LIGO Scientific Collaboration), “GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence,” *Phys. Rev. Lett.* **116**, 241103 (2016).
- [6] B. P. Abbott *et al.* (VIRGO, LIGO Scientific Collaboration), “GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2,” *Phys. Rev. Lett.* **118**, 221101 (2017).
- [7] B. P. Abbott *et al.* (Virgo, LIGO Scientific Collaboration), “GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence,” *Phys. Rev. Lett.* **119**, 141101 (2017).
- [8] B. P. Abbott *et al.* (Virgo, LIGO Scientific Collaboration), “GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral,” *Phys. Rev. Lett.* **119**, 161101 (2017).
- [9] B. P. Abbott *et al.* (Virgo, LIGO Scientific Collaboration), “GW170608: Observation of a 19-solar-mass Binary Black Hole Coalescence,” *Astrophys. J.* **851**, L35 (2017).

- [10] K. Koyama, “Cosmological Tests of Modified Gravity,” *Rept. Prog. Phys.* **79**, 046902 (2016).
- [11] P. Jordan, “The present state of Dirac’s cosmological hypothesis,” *Z. Phys.* **157**, 112–121 (1959).
- [12] C. Brans and R. H. Dicke, “Mach’s principle and a relativistic theory of gravitation,” *Phys. Rev.* **124**, 925–935 (1961).
- [13] H. A. Buchdahl, “Non-linear Lagrangians and cosmological theory,” *Mon. Not. Roy. Astron. Soc.* **150**, 1 (1970).
- [14] G. W. Horndeski, “Second-order scalar-tensor field equations in a four-dimensional space,” *Int. J. Theor. Phys.* **10**, 363–384 (1974).
- [15] C. Deffayet, X. Gao, D. A. Steer, and G. Zahariade, “From k -essence to generalised Galileons,” *Phys. Rev. D* **84**, 064039 (2011).
- [16] T. Kobayashi, M. Yamaguchi, and J. Yokoyama, “Generalized G-inflation: Inflation with the most general second-order field equations,” *Prog. Theor. Phys.* **126**, 511–529 (2011).
- [17] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, “Healthy theories beyond Horndeski,” *Phys. Rev. Lett.* **114**, 211101 (2015).
- [18] D. Langlois and K. Noui, “Degenerate higher derivative theories beyond Horndeski: evading the Ostrogradski instability,” *JCAP* **1602**, 034 (2016).
- [19] M. Crisostomi, K. Koyama, and G. Tasinato, “Extended Scalar-Tensor Theories of Gravity,” *JCAP* **1604**, 044 (2016).
- [20] J. Ben Achour, M. Crisostomi, K. Koyama, D. Langlois, K. Noui, and G. Tasinato, “Degenerate higher order scalar-tensor theories beyond Horndeski up to cubic order,” *JHEP* **12**, 100 (2016).
- [21] L. Heisenberg, “Generalization of the Proca Action,” *JCAP* **1405**, 015 (2014).
- [22] L. Heisenberg, R. Kase, and S. Tsujikawa, “Beyond generalized Proca theories,” *Phys. Lett. B* **760**, 617–626 (2016).
- [23] R. Kimura, A. Naruko, and D. Yoshida, “Extended vector-tensor theories,” *JCAP* **1701**, 002 (2017).
- [24] S. F. Hassan and R. A. Rosen, “Bimetric Gravity from Ghost-free Massive Gravity,” *JHEP* **02**, 126 (2012).

- [25] H. Motohashi, T. Suyama, and M. Yamaguchi, “Ghost-free theory with third-order time derivatives,” (2017).
- [26] R. P. Woodard, “Ostrogradsky’s theorem on Hamiltonian instability,” *Scholarpedia* **10**, 32243 (2015).
- [27] H. Motohashi, K. Noui, T. Suyama, M. Yamaguchi, and D. Langlois, “Healthy degenerate theories with higher derivatives,” *JCAP* **1607**, 033 (2016).
- [28] M. Crisostomi, R. Klein, and D. Roest, “Higher Derivative Field Theories: Degeneracy Conditions and Classes,” *JHEP* **06**, 124 (2017).
- [29] X. Gao, “Unifying framework for scalar-tensor theories of gravity,” *Phys. Rev. D* **90**, 081501 (2014).
- [30] T. Fujita, X. Gao, and J. Yokoyama, “Spatially covariant theories of gravity: disformal transformation, cosmological perturbations and the Einstein frame,” *JCAP* **1602**, 014 (2016).
- [31] J. D. Bekenstein, “The Relation between physical and gravitational geometry,” *Phys. Rev. D* **48**, 3641–3647 (1993).
- [32] J.-P. Bruneton and G. Esposito-Farèse, “Field-theoretical formulations of MOND-like gravity,” *Phys. Rev. D* **76**, 124012 (2007), [Erratum: *Phys. Rev. D* **76**, 129902 (2007)].
- [33] D. Bettoni and S. Liberati, “Disformal invariance of second order scalar-tensor theories: Framing the Horndeski action,” *Phys. Rev. D* **88**, 084020 (2013).
- [34] A. H. Chamseddine and V. Mukhanov, “Mimetic Dark Matter,” *JHEP* **11**, 135 (2013).
- [35] L. Sebastiani, S. Vagnozzi, and R. Myrzakulov, “Mimetic gravity: a review of recent developments and applications to cosmology and astrophysics,” *Adv. High Energy Phys.* **2017**, 3156915 (2017).
- [36] S. Nojiri and S. D. Odintsov, “Mimetic $F(R)$ gravity: inflation, dark energy and bounce,” *Mod. Phys. Lett. A* **29**, 1450211 (2014).
- [37] F. Arroja, N. Bartolo, P. Karmakar, and S. Matarrese, “The two faces of mimetic Horndeski gravity: disformal transformations and Lagrange multiplier,” *JCAP* **1509**, 051 (2015).
- [38] A. H. Chamseddine, V. Mukhanov, and A. Vikman, “Cosmology with Mimetic Matter,” *JCAP* **1406**, 017 (2014).

- [39] L. Mirzaghali and A. Vikman, “Imperfect Dark Matter,” *JCAP* **1506**, 028 (2015).
- [40] F. Capela and S. Ramazanov, “Modified Dust and the Small Scale Crisis in CDM,” *JCAP* **1504**, 051 (2015).
- [41] J. Klusoň, “Canonical Analysis of Inhomogeneous Dark Energy Model and Theory of Limiting Curvature,” *JHEP* **03**, 031 (2017).
- [42] D. Langlois, M. Mancarella, K. Noui, and F. Vernizzi, “Effective Description of Higher-Order Scalar-Tensor Theories,” *JCAP* **1705**, 033 (2017).
- [43] K. Takahashi, H. Motohashi, T. Suyama, and T. Kobayashi, “General invertible transformation and physical degrees of freedom,” *Phys. Rev. D* **95**, 084053 (2017).
- [44] K. Takahashi and T. Kobayashi, “Extended mimetic gravity: Hamiltonian analysis and gradient instabilities,” *JCAP* **1711**, 038 (2017).
- [45] P. G. Bergmann, “Comments on the scalar tensor theory,” *Int. J. Theor. Phys.* **1**, 25–36 (1968).
- [46] W. Hu and I. Sawicki, “Models of $f(R)$ Cosmic Acceleration that Evade Solar-System Tests,” *Phys. Rev. D* **76**, 064004 (2007).
- [47] A. A. Starobinsky, “Disappearing cosmological constant in $f(R)$ gravity,” *JETP Lett.* **86**, 157–163 (2007).
- [48] T. P. Sotiriou and V. Faraoni, “ $f(R)$ Theories Of Gravity,” *Rev. Mod. Phys.* **82**, 451–497 (2010).
- [49] A. De Felice and S. Tsujikawa, “ $f(R)$ theories,” *Living Rev. Rel.* **13**, 3 (2010).
- [50] D. Lovelock, “The Einstein tensor and its generalizations,” *J. Math. Phys.* **12**, 498–501 (1971).
- [51] T. Narikawa, T. Kobayashi, D. Yamauchi, and R. Saito, “Testing general scalar-tensor gravity and massive gravity with cluster lensing,” *Phys. Rev. D* **87**, 124006 (2013).
- [52] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, “Essential Building Blocks of Dark Energy,” *JCAP* **1308**, 025 (2013).
- [53] M. Zumalacárregui and J. García-Bellido, “Transforming gravity: from derivative couplings to matter to second-order scalar-tensor theories beyond the Horndeski Lagrangian,” *Phys. Rev. D* **89**, 064046 (2014).

- [54] D. Blas, O. Pujolàs, and S. Sibiryakov, “On the Extra Mode and Inconsistency of Hořava Gravity,” *JHEP* **10**, 029 (2009).
- [55] D. Blas, O. Pujolàs, and S. Sibiryakov, “Models of non-relativistic quantum gravity: The Good, the bad and the healthy,” *JHEP* **04**, 018 (2011).
- [56] M. Crisostomi, M. Hull, K. Koyama, and G. Tasinato, “Horndeski: beyond, or not beyond?,” *JCAP* **1603**, 038 (2016).
- [57] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, “Exploring gravitational theories beyond Horndeski,” *JCAP* **1502**, 018 (2015).
- [58] J. M. Ezquiaga, J. García-Bellido, and M. Zumalacárregui, “Field redefinitions in theories beyond Einstein gravity using the language of differential forms,” *Phys. Rev. D* **95**, 084039 (2017).
- [59] X. Gao, “Hamiltonian analysis of spatially covariant gravity,” *Phys. Rev. D* **90**, 104033 (2014).
- [60] R. Saitou, “Canonical invariance of spatially covariant scalar-tensor theory,” *Phys. Rev. D* **94**, 104054 (2016).
- [61] C. Deffayet, G. Esposito-Farèse, and D. A. Steer, “Counting the degrees of freedom of generalized Galileons,” *Phys. Rev. D* **92**, 084013 (2015).
- [62] D. Langlois and K. Noui, “Hamiltonian analysis of higher derivative scalar-tensor theories,” *JCAP* **1607**, 016 (2016).
- [63] H. Motohashi, T. Suyama, and K. Takahashi, “Fundamental theorem on gauge fixing at the action level,” *Phys. Rev. D* **94**, 124021 (2016).
- [64] A. Padilla and V. Sivanesan, “Covariant multi-galileons and their generalisation,” *JHEP* **04**, 032 (2013).
- [65] S. Ohashi, N. Tanahashi, T. Kobayashi, and M. Yamaguchi, “The most general second-order field equations of bi-scalar-tensor theory in four dimensions,” *JHEP* **07**, 008 (2015).
- [66] H. Motohashi and T. Suyama, “Third order equations of motion and the Ostrogradsky instability,” *Phys. Rev. D* **91**, 085009 (2015).
- [67] J. Ben Achour, D. Langlois, and K. Noui, “Degenerate higher order scalar-tensor theories beyond Horndeski and disformal transformations,” *Phys. Rev. D* **93**, 124005 (2016).

- [68] N. Deruelle and J. Rua, “Disformal Transformations, Veiled General Relativity and Mimetic Gravity,” *JCAP* **1409**, 002 (2014).
- [69] M. Minamitsuji, “Disformal transformation of cosmological perturbations,” *Phys. Lett. B* **737**, 139–150 (2014).
- [70] S. Tsujikawa, “Disformal invariance of cosmological perturbations in a generalized class of Horndeski theories,” *JCAP* **1504**, 043 (2015).
- [71] Y. Watanabe, A. Naruko, and M. Sasaki, “Multi-disformal invariance of non-linear primordial perturbations,” *Europhys. Lett.* **111**, 39002 (2015).
- [72] H. Motohashi and J. White, “Disformal invariance of curvature perturbation,” *JCAP* **1602**, 065 (2016).
- [73] G. Domènech, A. Naruko, and M. Sasaki, “Cosmological disformal invariance,” *JCAP* **1510**, 067 (2015).
- [74] G. Domènech, S. Mukohyama, R. Namba, A. Naruko, R. Saitou, and Y. Watanabe, “Derivative-dependent metric transformation and physical degrees of freedom,” *Phys. Rev. D* **92**, 084027 (2015).
- [75] T. Damour and G. Esposito-Farèse, “Tensor multiscalar theories of gravitation,” *Class. Quant. Grav.* **9**, 2093–2176 (1992).
- [76] M. Chaichian, J. Klusoň, M. Oksanen, and A. Tureanu, “Mimetic dark matter, ghost instability and a mimetic tensor-vector-scalar gravity,” *JHEP* **12**, 102 (2014).
- [77] A. Golovnev, “On the recently proposed Mimetic Dark Matter,” *Phys. Lett. B* **728**, 39–40 (2014).
- [78] H. Liu, K. Noui, E. Wilson-Ewing, and D. Langlois, “Effective loop quantum cosmology as a higher-derivative scalar-tensor theory,” *Class. Quant. Grav.* **34**, 225004 (2017).
- [79] A. Ijjas, J. Ripley, and P. J. Steinhardt, “NEC violation in mimetic cosmology revisited,” *Phys. Lett. B* **760**, 132–138 (2016).
- [80] H. Firouzjahi, M. A. Gorji, and A. Hosseini Mansoori, “Instabilities in Mimetic Matter Perturbations,” *JCAP* **1707**, 031 (2017).
- [81] P. Hořava, “Quantum Gravity at a Lifshitz Point,” *Phys. Rev. D* **79**, 084008 (2009).
- [82] S. Ramazanov, F. Arroja, M. Celería, S. Matarrese, and L. Pilo, “Living with ghosts in Hořava-Lifshitz gravity,” *JHEP* **06**, 020 (2016).

- [83] T. P. Sotiriou, M. Visser, and S. Weinfurtner, “Quantum gravity without Lorentz invariance,” *JHEP* **10**, 033 (2009).
- [84] S. Hirano, S. Nishi, and T. Kobayashi, “Healthy imperfect dark matter from effective theory of mimetic cosmological perturbations,” *JCAP* **1707**, 009 (2017).
- [85] G. W. Horndeski, “Conformally Invariant Scalar-Tensor Field Theories in a Four-Dimensional Space,” (2017).
- [86] T.-j. Chen, M. Fasiello, E. A. Lim, and A. J. Tolley, “Higher derivative theories with constraints: Exorcising Ostrogradski’s Ghost,” *JCAP* **1302**, 042 (2013).
- [87] Y. Zheng, L. Shen, Y. Mou, and M. Li, “On (in)stabilities of perturbations in mimetic models with higher derivatives,” *JCAP* **1708**, 040 (2017).
- [88] M. Crisostomi, K. Noui, C. Charmousis, and D. Langlois, “Beyond Lovelock: on higher derivative metric theories,” (2017).
- [89] F. Arroja, T. Okumura, N. Bartolo, P. Karmakar, and S. Matarrese, “Large-scale structure in mimetic Horndeski gravity,” (2017).
- [90] T. Berger and A. Ilchmann, “On the standard canonical form of time-varying linear DAEs,” *Quart. Appl. Math.* **71**, 69 (2013).
- [91] E. Noether, “Invariant Variation Problems,” *Gott. Nachr.* **1918**, 235–257 (1918), [*Transp. Theory Statist. Phys.* **1**, 186 (1971)].