# Causality and the initial value problem in Modified Gravity

This dissertation is submitted for the degree of
*Doctor of Philosophy*

**Giuseppe Papallo**

University of Cambridge
Trinity Hall

29th March 2018

*E o fim durou, e o fim durava*

*O que acabou não se acabava*

*A lição foi mal passada*

*Quem aprendeu não sabe nada.*

*No entanto, afinal*

*Chega a fase racional*

*Leia o livro, vai saber…*

— Tim Maia, *Universo em Desencanto*

**Causality and the initial value problem in Modified Gravity**

---

# Abstract

Lovelock and Horndeski theories are natural generalisations of Einstein's theory of General Relativity. They find applications in Astrophysics, Cosmology and String Theory. This dissertation discusses some issues regarding the mathematical consistency of these theories.

In the first part of the thesis we study the Shapiro time delay for gravitons in spherically symmetric spacetimes in Einstein–Gauss–Bonnet gravity (a Lovelock theory). In Lovelock theories, gravitons can propagate faster or slower than light. We show that, thanks to this property, it is possible for them to experience a negative time delay. It was recently argued that this feature could be employed to construct closed causal curves, implying that the theory should be discarded as causally pathological. We show that this construction is unphysical, for it cannot be realised as the evolution of sensible initial data.

The second part investigates the local well-posedness of the initial value problem for Lovelock and Horndeski theories. For the initial value problem to be well-posed it is necessary that the equations of motion be strongly hyperbolic. It is known that when the background fields are large, even weak hyperbolicity may fail. Hence, we consider the weak field regime, in which these equations can be considered as small perturbations of the Einstein equations. We prove that both Lovelock and Horndeski theories are weakly hyperbolic in a generic weak field background in harmonic and generalised harmonic gauge, respectively. We show that Lovelock theories fail to be strongly hyperbolic in this setting. We also prove that the most general Horndeski theory which is strongly hyperbolic is simply a "k-essence" theory coupled to Einstein gravity and that any more general theory would necessarily fail to be so.

Our results imply that the standard methods used to prove the well-posedness of the initial value problem for the Einstein equations cannot be extended to Lovelock or Horndeski theories. This raises the possibility that these theories may not admit a well-posed initial value problem even for weak fields and hence might not constitute a valid alternative to General Relativity.

# Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

The work presented is based on research conducted at the Department of Applied Mathematics and Theoretical Physics, between October 2014 and March 2018. The results discussed were published in:

- G. Papallo and H. S. Reall, *Graviton time delay and a speed limit for small black holes in Einstein-Gauss-Bonnet theory*, *J. High Energy Phys.* **2015** (2015) 109, [1508.05303]
- G. Papallo and H. S. Reall, *On the local well-posedness of Lovelock and Horndeski theories*, *Phys. Rev. D* **96** (2017) 044019, [1705.04370]
- G. Papallo, *On the hyperbolicity of the most general Horndeski theory*, *Phys. Rev. D* **96** (2017) 124036, [1710.10155]

Giuseppe Papallo

# Acknowledgements

First, and foremost, I would like to thank my supervisor, Harvey Reall, for the guidance and support he offered me throughout my Ph.D. Without his precious advice, comments and ideas, this thesis would not exist.

I would like to thank all of my friends at DAMTP, who have made my time here very enjoyable. In particular, I thank Markus Kunesch, Felicity Eperon, Joe Keir and Dejan Gajic for the many scientific (and non-scientific) discussions.

I am grateful to Lavinia, Salvo, Alex, James, Gianni, Giulia, Alessio, Ian, Breno, Erick, Christoph, Beñat, the Trinity Hall MCR Football Club, the Cambridge University Brazilian Society and all my friends with whom I shared many happy moments in the past four years. I am also grateful to Geo, Marco, Giugno, Sergio, Zazza, Stefano, Marta and all my old friends who have always been close to me despite the physical distance separating us.

I thank Bruno for all the adventures we lived together since our undergraduate days and for being the best friend one could ask for.

I thank my whole family, for bringing me up in an always stimulating environment, for encouraging me and for giving me the opportunity to pursue my dreams. Without their love and support I would not be here.

Finally, I thank Aline for being always by my side, for helping me in the difficult moments, for inspiring me and for bringing happiness into my life every single day.

### Financial support

The research presented in this thesis was supported by an STFC studentship and a Cambridge Philosophical Society research grant.

# Contents

# Acronyms

**ΛCDM** Λ Cold Dark Matter;

**ADM** Arnowitt–Deser–Misner;

**AS** Aichelburg–Sexl;

**CEMZ** Camanho–Edelstein–Maldacena–Zhiboedov;

**CMB** Cosmic Microwave Background;

**EGB** Einstein–Gauss–Bonnet gravity;

**EGdB** Einstein–dilaton–Gauss–Bonnet gravity;

**FLRW** Friedmann–Lemaître–Robertson–Walker;

**GR** General Relativity;

**LIGO** Laser Interferometer Gravitational-Wave Observatory.

# Glossary

$\Phi$ scalar field in Horndeski theories;

$X$ norm of the gradient of $\Phi$, $X \equiv -\frac{1}{2}(\nabla\Phi)^2$;

$H_a$ linearised gauge condition;

$M(\xi_i)$ principal part of the first order reduction of the differential operator (cf. Eqs. (2.2.13) and (2.3.13));

$H(\xi_i)$ non-degenerate, hermitian matrix such that $H(\xi_i)M(\xi_i)H(\xi_i) = M(\xi_i)^T$;

$\xi_0^\pm$ eigenvalues of $M(\xi_i)$ in GR;

$\xi_0^\pm$−group eigenvalues of $M(\xi_i)$ in Lovelock/Horndeski that tend to $\xi_0^\pm$, respectively, in the limit of vanishing coupling;

$V^\pm$ total generalised eigenspace for $\xi_0^\pm$-group;

$V$ sum of the total generalised eigenspaces for $\xi_0^\pm$-groups, $V = V^+ \oplus V^-$;

$H^\pm$ restriction of $H(\xi_i)$ to the $V^\pm$ subspaces;

$N^\pm$ $H^\pm$-null subspace of $V^\pm$ defined by “pure gauge” vectors.

# Chapter 1

# Introduction

General Relativity is, without doubt, one of the greatest achievements of modern Physics. It provides a simple and elegant description of gravity compatible with Special Relativity. In General Relativity, the fixed space and absolute time of Newtonian physics are replaced by a dynamical manifold, the *spacetime*, whose geometry describes the gravitational field. The geometry of the spacetime affects how matter "moves" and is in turn influenced by the matter present in it. The relation between the curvature of the spacetime and its matter content is determined by the Einstein field equations:[1]

$$G_{ab} = 8\pi T_{ab}.$$

Since its inception in 1915 [1, 2], General Relativity has performed remarkably well in all of the experimental tests, from Eddington's measurement of the deflection of light in 1919 [3] to the recent direct observation of *gravitational waves* by the LIGO collaboration [4].

Despite its numerous achievements, Einstein's theory still presents some shortcomings which motivate the study of alternative theories of gravity.

The presence of singularities [5, 6] suggests that General Relativity breaks down at curvatures comparable with the Planck scale. At these scales, both gravitational and quantum mechanical effects become important and hence a quantum description of gravity becomes necessary. It is well-known that, due to its non-renormalisability, General Relativity cannot provide such description. A more fundamental theory reconciling gravity with the other forces is needed.

On large scales, General Relativity fails to provide a satisfying explanation for certain cosmological phenomena. The accelerated expansion of our Universe [7, 8] suggests that its energy content must be dominated by *Dark Energy*, a non-clustering

[1] We choose units in which $c = G = 1$.

form of energy which takes the form of a perfect fluid with sufficiently negative pressure ($\omega < -1/3$). A simple model for Dark Energy is provided by adding a positive Cosmological Constant term to the Einstein equations[2]

$$G_{ab} + \Lambda g_{ab} = 8\pi T_{ab}.$$

Observations set the Cosmological Constant to have a small, positive value in this model, known as ΛCDM model.

While giving good phenomenological predictions, the ΛCDM model can be considered unsatisfactory from a theoretical point of view. The main issue, known as the *Cosmological Constant problem* [10], is related to the discrepancy between the natural scale and the measured value of $\Lambda$. In fact, the vacuum energy contribution to $\Lambda$ expected from Particle Physics is more than one-hundred orders of magnitude larger than the observed one. To account for the observed smallness of $\Lambda$, its "bare" value should be such as to *almost* cancel the vacuum energy contribution. The reasons for such "miraculous" cancellation are not known.

Many believe that a modification of General Relativity is needed in order to address these issues. Lovelock theorem [11] — which we will discuss in more detail in Chapter 3 — on one hand provides a nice proof of "uniqueness" for General Relativity (under certain assumptions), on the other it "embeds" the theory in a more general framework. In a way, it provides us with a consistent "guide" on how to generalise General Relativity. Thinking of General Relativity as a "special case" of a more general theory can be interesting from the mathematical point of view. Moreover, the study of more general theories can be beneficial to our understanding of General Relativity itself, by providing insight or by developing new useful tools and techniques. How are these generalisations of Einstein's theory obtained? In brief, Lovelock's theorem tells us that if we want to modify General Relativity whilst maintaining some of its essential properties — such as having second order equations of motion arising from a diffeomorphism-invariant action — we must do so by adding extra degrees of freedom in specific ways. One way to do so consists in including additional fields (scalars, vectors or tensors) in our theory. A second way of achieving this consists instead in maintaining the metric as the only field mediating the gravitational interaction but considering a spacetime of dimension greater than four. We will examine these possibilities shortly.

[2] The Cosmological Constant was first introduced by Einstein himself in order to obtain static Universe solutions [9]. The idea was later dropped after Hubble's discovery of the expansion of the Universe.

### Higher dimensions and Lovelock theories

Let us consider the case of a higher dimensional spacetime. Lovelock proved that in dimensions greater than four, the most general diffeomorphism-covariant metric theory of gravity with second order equations of motion is not General Relativity, but rather a generalisation of it known as *Lovelock gravity* [12, 11]. In this theory the Einstein tensor in the field equations is replaced by a more general tensor, consisting of "powers" of the Riemann curvature tensor. An interesting property of Lovelock gravity is that it automatically reduces to Einstein's theory in four spacetime dimensions. Since this theory basically satisfies the same assumptions as General Relativity and reduces to it in four dimensions, we have no a priory reason to prefer one to the other. It is therefore important to understand the properties of this theory and to look for evidence that either supports it or invalidates it. But why would we be interested in a theory that differs from General Relativity only in higher dimensions?

Over the past few decades several reasons, other than pure mathematical curiosity, have motivated the study of theories of gravity in higher dimensions, initiated by the pioneering work of Kaluza [13] and Klein [14, 15] shortly after the birth of General Relativity. They proposed a model to unify gravity and electromagnetism based on the addition of a fifth (compact) dimension. The five-dimensional equations would then yield the four-dimensional Einstein and Maxwell equations. This theory, however, also predicts the existence of an additional non-minimally coupled scalar field. We will come back to this point later in this section.

A great effort has been devoted to the quest for a *theory of everything* which would reconcile Gravity with the other fundamental forces. Many consider *String Theory* to be the best candidate, as it contains General Relativity and appears to be renormalisable. Consistency of this theory, however, requires the underlying spacetime to have ten (or more) dimensions [16] rather than four, as postulated in Einstein's theory. An understanding of gravity in more than four dimensions is essential in order to make sense of this theory. Interestingly, effective action computations of certain String Theories yield corrections to the Einstein–Hilbert action that take the form of Lovelock theories [17, 18]. We will come back to this last point later.

In the past twenty years, the interest in higher dimensions was further fuelled by the rising popularity of the *"gauge/gravity duality"* [19, 20]. Roughly speaking, this duality relates a strongly coupled $d$–dimensional (conformal) field theory, to a $(d + 1)$–dimensional, weakly-coupled, theory of gravity. In particular, certain four-dimensional quantum field theories are conjectured to be "equivalent" to five-dimensional theories of gravity. Thanks to this equivalence it is possible to map certain challenging computations in the field theory to more tractable ones in the gravitational theory. For this reason, the study of higher dimensional theories of gravity could bring useful contributions to the understanding of strongly coupled field theories.

### Scalar fields and Horndeski theories

Let us now consider the possibility of adding extra degrees of freedom in four spacetime dimensions. We will only be interested in the simplest such modification, that is we will consider the case of a single additional *scalar* degree of freedom. Theories in which the gravitational interaction is mediated by the metric and a scalar field are known as *scalar-tensor* theories. The most general diffeomorphism-covariant scalar-tensor theory of gravity with second order equations of motion is known as *Horndeski gravity* [21].

The idea of considering an additional scalar degree of freedom is not a new one. As mentioned earlier, the model of Kaluza and Klein gives rise to a scalar field non-minimally coupled to gravity. The idea was further explored by Jordan [22], Brans and Dicke [23]. In their theory, the scalar field was taken to represent a dynamical Newton's constant (an idea first put forward by Dirac [24]). Despite being tightly constrained by observation, Jordan–Brans–Dicke theory has played a fundamental role in the development of scalar-tensor theories.

Scalar fields also play an important role in String Theory, Astrophysics and Cosmology. For example, scalar fields — such as the *dilaton* — naturally appear in the dimensional reduction of higher dimensional theories such as String Theory or Kaluza-Klein theory. In fact, certain subclasses of Horndeski theory can be obtained as dimensional reduction of Lovelock theories.

Certain theories — such as quintessence [25, 26] — attempt to solve the issues related with the fine-tuning of the Cosmological Constant and its interpretation as the vacuum energy by modelling Dark Energy as a dynamical scalar field. Dark energy models in the context of Horndeski theory are discussed in e.g. [27].

Furthermore, scalar fields play an essential role in the theory of Cosmological *inflation*, proposed by Guth [28] in order to explain the observed isotropy and (spatial) flatness of the Universe [29, 30]. Horndeski theories provide a framework to study the most general single-field inflation models [31].

The importance of Horndeski theories in Early Universe Cosmology is not limited to inflation models. In fact, they play an important role in alternative models such as *bouncing cosmologies* [32, 33].

More generally, many scalar-tensor theories of interest — such as Jordan–Brans–Dicke, k-essence, Einstein–dilaton–Gauss–Bonnet, Galileons, etc. — can be recovered as special cases of Horndeski gravity.

#### Effective field theory

As mentioned earlier, when performing effective action calculations in certain String Theories, Lovelock terms appear among more general higher order corrections. More generally, both Lovelock and Horndeski theories are obtained by adding higher derivative terms to the Einstein–Hilbert action. One may wonder whether these theories can be recovered as effective theories for some more fundamental theory.

The higher order corrections present in Lovelock and Horndeski theories form a special class of terms that yield second order equations of motion. In the context of effective field theory, all higher order derivatives terms in the action are suppressed by some UV mass scale $M$. This scale does not distinguish between the "special" and the more general terms. This means that the "special" terms only become important when *all* higher order terms do.

The natural question to ask then is how would Lovelock and Horndeski theories be recovered in this framework? A possible solution to this problem is given by postulating the existence of a second mass scale $M'$ associated with the "special" terms. If this scale is much larger than the scale $M$ associated with the more general higher order terms, i.e. $M' \gg M$, then there exist a regime in which the general higher order terms are negligible compared to the "special" ones. Lovelock and Horndeski theories could then be recovered as effective theories at this energy scale. One could, in principle, investigate whether this situation could arise from some UV theory. However, in order to do so we would need to make assumptions about such unknown UV theory.

While many people do, in fact, treat both Lovelock and Horndeski theories as effective theories, this *ad hoc* approach seems rather unnatural and it is not clear how and why the separation of scales needed for this to work would happen in nature. Others, instead, follow the more conservative approach of considering these theories to be the fundamental theories describing the gravitational interaction. In this case, for these to be considered physical theories, they must be shown to be classically self-consistent — a problem that has not received much attention in the literature. The main goal of this thesis will be to shed some light on this issue.

### Experimental constraints

The recent direct observation of gravitational waves [4] — besides being a remarkable scientific and technological achievement — opened up the possibility of studying gravitational physics in regimes never tested before. In particular, testing the behaviour of gravity in regimes where alternative theories are expected to deviate from General Relativity could provide useful constraints for ruling out unphysical models.

As we will see in more detail later on, in these theories, contrary to what happens in General Relativity, gravitational perturbations do *not* in general propagate on null curves of the spacetime metric. In other words, the "speed of gravity" will, in general, differ from the speed of light. The nature and magnitude of this effect depends on the free parameters of these theories, as well as other factors, such as the direction and polarization of the gravitational waves. Providing experimental bounds for the speed of gravitational waves could therefore impose constraints on the free parameters of these theories.

On 17th August 2017, a few seconds after the LIGO and VIRGO detectors picked up a loud signal, the Fermi gamma-ray satellite detected a strong burst of gamma rays coming from the same location. The event, GW170817, was then confirmed to be a neutron star merger [34, 35]. The delay between the arrival of the gravitational and electromagnetic signals can be used to estimate the speed of gravitational waves, which was found to agree with the speed of light within one part in $10^{15}$ [36].

The implications of these observations for Horndeski theories were discussed in Refs. [37, 38, 39, 40]. They consider the Cosmological models commonly used to explain the accelerated expansion of the Universe and show that the observation GW170817 imposes stringent constraints on most free parameters, forcing them to vanish.[3] While the remaining free parameters can still be tuned as to reproduce Cosmic acceleration in the absence of a Cosmological Constant, the resulting models are incompatible with other observations such as the correlation between galaxy distributions and the CMB. They then conclude that in this context, a large subclass of Horndeski (and beyond-Horndeski) theories — including theories such as Galileons, Einstein–dilaton–Gauss–Bonnet, Fab Four, etc. — should be ruled out. The surviving class of Horndeski theories consists simply of Einstein gravity coupled to a "k-essence" field.

The results presented in this Thesis are compatible with the above discussion. We

[3] In these Cosmological models, both the metric the scalar field mediating the gravitational interaction are taken to be functions of time only, an assumption which considerably simplifies the phenomenology.

will conclude that the aforementioned class of Horndeski theories is, in fact, better behaved than the others from the mathematical point of view (cf. Chapter 6).

## Objectives

While both Lovelock and Horndeski theories have been extensively studied in the Physics literature, the issue of their mathematical consistency has received little attention. In this thesis we will be concerned with two main themes: causality and the well-posedness of the initial value problem.

As mentioned earlier, in these theories, the propagation speed of the gravitational perturbations will in general depend on the background fields and will differ from the speed of light. In particular, gravitons can propagate faster than light in these theories. We will focus on Einstein–Gauss–Bonnet gravity (a Lovelock theory) and we will study the causal structure of the theory in a static, spherically symmetric spacetime. One of our objectives will be to verify a recent claim [41] that gravitons can experience a negative Shapiro time delay in Lovelock gravity. Ref. [41] then argued that this property could be used to construct closed causal curves. They conclude that since this theory suffers from causal pathologies, it should be discarded. We will study the proposed construction in detail in order to assess its viability and establish whether these arguments are sufficient to deem the theory "pathological".

The core of the thesis will focus on the initial value problem for Horndeski and Lovelock theories. The *local well-posedness* of the initial value problem is a fundamental requirement for any classical physical theory to make sense. By well-posedness we mean that once the initial state of the system is specified (and is sufficiently "well behaved"), then a solution to the equations of motion must *exist*, be *unique* and *continuously depend on the initial data.* By "locally" we mean that the time of existence could be infinitesimally small, but must be strictly greater than zero. Roughly speaking, these properties ensure that the theory is deterministic and has predictive power. While a local well-posedness theorem for the Einstein equations has been proved in 1952 [42], a corresponding result for Lovelock or Horndeski theories is still lacking.

We will investigate whether these theories admit a well-posed initial value problem. In particular, we will focus on the linearised theory, since its well-posedness is a necessary condition for the well-posedness of the non-linear problem.

The well-posedness of the linearised theory is tightly related to the *hyperbolicity* of the equations of motion. In contrast to General Relativity, these theories may fail to be hyperbolic around certain backgrounds. Examples of this failure usually involve considering backgrounds with "large fields".[4] One may still hope that restricting to "weak fields" could cure this problem and ensure that the initial value problem be well-posed in this regime, where the equation of motion can be considered as "small perturbations" of the Einstein equations.

Note that there are backgrounds around which the equations can be shown to be well-posed at the linearised level. However, these solutions are *non-generic*, i.e. they exhibit a high degree of symmetry. This feature is not sufficient to establish well-posedness at the non-linear level. In fact, it is necessary that the linearised theory be well-posed around a *generic* background.

We will study in detail the hyperbolicity of Lovelock and Horndeski theories in generic weak fields backgrounds. We will discuss the implications for the well-posedness of the initial value problem and how this affects the viability of these theories as alternatives to General Relativity.

## Conventions

We summarise here some of the conventions we will use throughout the thesis. We use units such that $c = G = 1$. The spacetime metric is taken to have signature $(- + \ldots +)$. Latin indices $a, b, c, \ldots$ are abstract indices denoting tensor equations valid in any basis. Greek indices $\mu, \nu, \ldots$ refer to a particular basis, e.g. a coordinate basis. In a basis, Greek indices are "spacetime" indices, running from 0 to $d - 1$, while Latin indices $i, j, k, \ldots$ are "spatial" indices, running from 1 to $d - 1$.

## Outline

This thesis is organised as follows. In the next Chapter we will introduce the initial value problem and review the theory of hyperbolic partial differential equations which will be used in the other chapters. In Chapter 3 we will review some background material on Lovelock and Horndeski theories. The rest of the thesis contains the original results of our research. Chapter 4 is based on Ref. [43]. We will discuss aspects of the causal structure of Lovelock theories. In particular we will study the Shapiro time delay and the possibility of constructing "time machines" in Einstein–Gauss–Bonnet gravity. Chapters 5 and 6 are based on Refs. [44, 45] and contain our main results. There we discuss the local well-posedness of Lovelock and Horndeski equations of motion around a weak-field background. Finally, in Chapter 7 we summarise our results and discuss their implications.

[4] In Lovelock theories by "large fields" we mean large curvature. In Horndeski theories we also have large values of the scalar fields and its gradients.

# Chapter 2

# Hyperbolic PDEs and the initial value problem

In this Chapter we will review some theoretical background on hyperbolic partial differential equations. The first section briefly introduces the initial value problem in Classical Physics. Sections 2.2 and 2.3 will be devoted to the discussion of different notions of hyperbolicity in first and second order systems. These sections are mainly based on [46, 47]. In Section 2.4, mainly based on [48, 49, 50, 51], we will discuss the causal structure induced by an hyperbolic PDE. Finally, we will show as an example how these concepts apply to General Relativity (Section 2.5).

## 2.1 The initial value problem

Many systems in classical physics are modelled through an initial value problem. In simple terms, this means that they are described by a system of differential equations determining the evolution of the system, together with an initial data set, characterising the initial state of the system. A simple example is given by the motion of a particle of mass $m$, subject to a force $F$ in Newtonian mechanics

$$\begin{cases} \dfrac{d^2x(t)}{dt^2} = \dfrac{1}{m} F(x(t), \dot{x}(t), t) & (2.1.1) \\ x(0) = x_0 \qquad \dot{x}(0) = v_0 & (2.1.2) \end{cases}$$

Once the initial data (2.1.2) — consisting of the particle's initial position and velocity — is specified, the ODE (2.1.1) has a *unique* solution determining the position of the particle at any later time $t > 0$.

The requirement that the solution be unique, given the initial data, is of crucial importance: it ensures that classical physics be deterministic. Moreover, one can show that the solution to (2.1.1) will depend continuously on $x_0$ and $v_0$. This is also a desirable property from the physical point of view: the solution should not change discontinuously if we slightly perturb the initial data. If this were not the case, then it would be necessary to measure the initial position and velocity with *infinite* precision in order to extract meaningful predictions out of this model, which is clearly not possible in reality.

These requirements were formalised by Hadamard [52].

**Definition 1.** The initial value problem is *well-posed* if a solution exists, it is unique and it depends continuously on the initial data.

*We will regard (local) well-posedness as a necessary condition for a physical theory to be meaningful.*

If a problem does not satisfy the above definition, it is said to be *ill-posed.* A classical example of an ill-posed initial value problem is given by the *Laplace equation* (in fact, it was the study of this problem that led Hadamard to formulate the above conditions). Let us look at the initial value problem for the two-dimensional Laplace equation:

$$\begin{cases} \Delta\phi(x,y) = 0 & (2.1.3) \\ \phi(x,0) = f(x) \qquad \partial_y\phi(0,\mathbf{x}) = h(x) & (2.1.4) \end{cases}$$

Consider a sequence of initial data $f(x) = f_n(x)$ and $h(x) = h_n(x)$, where

$$f_n(x) = 0, \qquad h_n(x) = e^{-\sqrt{n}}\sin(nx). \tag{2.1.5}$$

The solution is then given by

$$\phi_n(x,y) = \frac{e^{-\sqrt{n}}}{n}\sin(nx)\sinh(ny). \tag{2.1.6}$$

In the $n \to \infty$ limit, the initial data converges to zero,

$$f_n \to \bar{f} \equiv 0, \qquad h_n \to \bar{h} \equiv 0. \tag{2.1.7}$$

Note that the solution of (2.1.3)-(2.1.4) for trivial data, i.e. $f = \bar{f}$ and $h = \bar{h}$, is $\bar{\phi}(x,y) = 0$. On the other hand, in the $n \to \infty$ limit, the solution $\phi_n(x,y)$ blows up exponentially for any non-zero value of $y$. In other words, we have that $||\phi_n - \bar{\phi}|| \to \infty$ while $||f_n - \bar{f}||, ||h_n - \bar{h}|| \to 0$. From this we can conclude that the solution does not depend continuously on the initial data and hence, since the third condition in Definition 1 is violated, we conclude that this problem is ill-posed.[5]

[5] This example is quite special, since the Laplace equation does describe a perfectly physical system. It does, in fact, admit a well-posed *boundary* value problem. The issue here lies in the fact

For the purpose of illustration let us consider two examples of well-posed problems. The linear wave equation (which, as we will see in the following sections, is the prototypical example of a hyperbolic PDE) has a well-posed initial value problem

$$\begin{cases} \Box\phi(t,\mathbf{x}) = 0 & (2.1.8) \\ \phi(0,\mathbf{x}) = \phi_0(\mathbf{x}) \qquad \partial_t\phi(0,\mathbf{x}) = \psi_0(\mathbf{x}) & (2.1.9) \end{cases}$$

where the initial data consists of the initial profile of the wave $\phi_0$ and its time derivative $\psi_0$. Similarly, Maxwell's equations (in vacuum) can be cast as a system of linear wave equations for the electric and magnetic field

$$\begin{cases} \Box E(t,\mathbf{x}) = 0 & (2.1.10) \\ \Box B(t,\mathbf{x}) = 0 & (2.1.11) \\ E(0,\mathbf{x}) = E_0(\mathbf{x}) \qquad \partial_t E(0,\mathbf{x}) = E_1(\mathbf{x}) & (2.1.12) \\ B(0,\mathbf{x}) = B_0(\mathbf{x}) \qquad \partial_t B(0,\mathbf{x}) = B_1(\mathbf{x}). & (2.1.13) \end{cases}$$

together with a set of constraints

$$\nabla \cdot E(t,\mathbf{x}) = 0 \qquad \nabla \cdot B(t,\mathbf{x}) = 0. \tag{2.1.14}$$

In contrast with what happens for the linear wave equation (2.1.8), the presence of such constraints means that the initial data cannot be completely arbitrary. On the other hand, the structure of the equations ensures that if the constraints are satisfied by the initial data, then they will be satisfied by the solution at any time. As we will see, the structure of the Einstein equations will present some similarity with Maxwell's equations.

We will now discuss how one may establish hyperbolicity and prove well-posedness for first and second order systems.

## 2.2 First order systems

Let $(\mathcal{M}, g)$ be a $d-$dimensional spacetime and let $(t, x^i)$ be a coordinate chart on it. Consider a first order linear partial differential equation for an $N-$dimensional vector $u$:[6]

$$Au_t + P^i\partial_i u + Cu = 0, \tag{2.2.1}$$

that we considered the Laplace equation — an elliptic equation — as if it had the character of a wave equation — an hyperbolic equation. Elliptic equations do not describe the "time evolution" of a system, (e.g. they do not account for the propagation of signals with finite speed) and hence, intuitively, it does not really make sense to pose the problem as an initial value problem. Note that, in general, ill-posedness of the initial value problem does not tell us that the equation will instead admit a well-posed boundary value problem.

[6] $u_t := \partial_t u$.

where $A$, $P^i$ and $C$ are real $N \times N$ matrices. We assume that $A$ is invertible. Let us first consider a PDE with constant coefficients, i.e. the case in which the aforementioned matrices have constant entries. Since we are dealing with constant coefficients, we can obtain a formal solution to this system using Fourier transforms. Taking a spatial Fourier transform

$$\tilde{u}(t, \xi) = \frac{1}{(2\pi)^{(d-1)/2}} \int d^{d-1}x \exp(-i\xi_i x^i) u(t, x) \tag{2.2.2}$$

Eq. (2.2.1) becomes

$$\tilde{u}_t - i\mathcal{M}(\xi_i)\tilde{u} = 0, \tag{2.2.3}$$

where we have defined

$$\mathcal{M}(\xi_i) = A^{-1}(-P^i\xi_i + iC). \tag{2.2.4}$$

We can solve Eq. (2.2.3)

$$\tilde{u}(t, \xi_i) = \exp(i\mathcal{M}(\xi_i)t)\tilde{u}(0, \xi_i) \tag{2.2.5}$$

and hence we find a formal solution to (2.2.1)

$$u(t, x) = \frac{1}{(2\pi)^{(d-1)/2}} \int d^{d-1}\xi \exp(i\xi_i x^i) \exp(i\mathcal{M}(\xi_i)t)\tilde{u}(0, \xi_i). \tag{2.2.6}$$

For general initial data, this integral may not converge as the integrand will not necessarily decay fast enough as $|\xi| \to \infty$. Here we have defined

$$|\xi| = \sqrt{\xi_i \xi^i}. \tag{2.2.7}$$

We will now investigate which conditions should $\mathcal{M}$ satisfy in order to ensure convergence of the integral. Convergence is guaranteed if $\mathcal{M}(\xi_i)$ satisfies, for all $\xi_i$ and $t > 0$,

$$|| \exp(i\mathcal{M}(\xi_i)t)|| \leq f(t) \tag{2.2.8}$$

for some continuous function $f(t)$, independent of $\xi_i$. When this condition holds, the integral converges and the resulting solution satisfies (using Plancherel's theorem)

$$||u||(t) \leq f(t)||u||(0) \tag{2.2.9}$$

where $|| \cdots ||$ denotes the spatial $L^2$-norm.

Using this, one can prove that the initial value problem is locally well-posed. To see this, consider two solutions $u_1$ and $u_2$, arising from initial data $u_1^0$ and $u_2^0$, respectively. The above inequality implies

$$||u_1 - u_2||(t) \leq f(t)||u_1^0 - u_2^0|| \tag{2.2.10}$$

from which we immediately deduce that the solution depends continuously on the initial data. Moreover, considering the case $u_1^0 = u_2^0$ we see that the solution is unique. Note that since we are only interested in *local* well-posedness, we do not need to impose any condition on $f(t)$ other than continuity. In fact, even if it were growing exponentially, it could not grow unbounded in a finite time.

We thus need to determine under which circumstances (2.2.8) is satisfied. Since the convergence of (2.2.6) is a high-frequency problem, we will only be interested in the high-frequency behaviour of $\mathcal{M}$. To study this, we let

$$t = \frac{t'}{|\xi|} \qquad \hat{\xi}_i = \frac{\xi_i}{|\xi|} \tag{2.2.11}$$

and consider the $|\xi| \to \infty$ limit at fixed $t'$. Equation (2.2.8) becomes

$$|| \exp(iM(\hat{\xi}_i)t')|| \leq k, \tag{2.2.12}$$

where $k = f(0)$ and

$$M(\xi_i) = -A^{-1}P^i\xi_i \tag{2.2.13}$$

is the *principal part* of $\mathcal{M}$.

Consider an eigenvector $v$ of $M(\hat{\xi}_i)$ with eigenvalue $\lambda = \lambda_1 + i\lambda_2$. We have

$$\exp(iM(\hat{\xi}_i)t')v = e^{i\lambda_1 t'}e^{-\lambda_2 t'}v. \tag{2.2.14}$$

The condition (2.2.12) will be satisfied only if $\lambda_2 \geq 0$ for all $\hat{\xi}_i$. However, $M(\hat{\xi}_i)$ is a real matrix and hence, if $\lambda$ is an eigenvector then so is its complex conjugate $\bar{\lambda}$. Hence, consistency with (2.2.12) requires $\pm\lambda_2 \geq 0$, i.e.

$$\lambda_2 = 0.$$

We deduce that (2.2.12) implies that all eigenvalues of $M(\hat{\xi}_i)$ must be *real*. This motivates the definition of *weak hyperbolicity*: Eq. (2.2.1) is *weakly hyperbolic* if, and only if, for any real $\xi_i$ (with $\xi_i\xi_i = 1$), all eigenvalues of $M(\hat{\xi}_i)$ are real.

A failure of weak hyperbolicity would be a disaster for the initial value problem because the integrand in (2.2.6) would grow exponentially with $|\xi|$ at large $|\xi|$ so convergence would require highly fine-tuned initial data. However, while necessary, weak hyperbolicity is not sufficient to ensure well-posedness.

The matrix $M(\hat{\xi}_i)$ can be brought to Jordan normal form by a similarity transformation

$$M(\hat{\xi}_i) = S(\hat{\xi}_i)J(\hat{\xi}_i)S(\hat{\xi}_i)^{-1} \tag{2.2.15}$$

so

$$\exp(iM(\hat{\xi}_i)t') = S(\hat{\xi}_i)\exp(iJ(\hat{\xi}_i)t')S(\hat{\xi}_i)^{-1}. \tag{2.2.16}$$

Suppose that $M$ were not diagonalisable, i.e., $J$ contains a $n \times n$, $n \geq 2$, Jordan block associated to some eigenvalue $\lambda$. In this case, the RHS of (2.2.6) would exhibit polynomial growth in $t'$. To see why this is the case, consider the following example: let $J_2$ be a $2 \times 2$ block associated to the eigenvalue $\lambda$

$$J_2 = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \tag{2.2.17}$$

then

$$\exp(iJ_2 t') = e^{i\lambda t'} \begin{pmatrix} 1 & it' \\ 0 & 1 \end{pmatrix}. \tag{2.2.18}$$

If the equation is weakly hyperbolic, then we are guaranteed that $\lambda$ is real and hence that there is no exponential growth in $t'$. However, the presence of the term linear in $t'$ implies that (2.2.12) is not satisfied. More generally, the presence of an $n \times n$ block would lead to terms $t'^p$ for some integer $p \leq n$. Using $t' = |\xi| t$ this gives rise to terms proportional to $|\xi|^p$. The presence of such terms implies that it is not possible to obtain a bound of the form (2.2.9). The best one could hope for is that it is possible to modify the RHS to include sufficiently many spatial derivatives of $u$. Whether or not this is possible depends on the form of the zero derivative term $Cu$ in the equation of motion [53].[7] But even if this is possible, the "loss of derivatives" in (2.2.9) would be worrying if we are considering an equation obtained by linearising some non-linear equation. This is because the loss of derivatives would be a serious obstruction to establishing local well-posedness for the *non-linear* equation.

To avoid this problem, $M(\hat{\xi})$ must be diagonalizable, i.e., there exists a matrix $S(\hat{\xi}_i)$ such that $M = SDS^{-1}$ where $D(\hat{\xi}_i)$ is diagonal. Defining a positive definite hermitian matrix $K(\hat{\xi}_i) = (S^{-1})^\dagger S^{-1}$ we then have

$$K(\hat{\xi}_i) M(\hat{\xi}_i) K(\hat{\xi}_i)^{-1} = M(\hat{\xi}_i)^\dagger .^8 \tag{2.2.19}$$

This motivates the definition of strong hyperbolicity. With constant coefficients, equation (2.2.1) is *strongly hyperbolic* if, and only if, there exists a positive definite hermitian matrix $K(\hat{\xi}_i)$ depending smoothly on $\hat{\xi}_i$ such that (2.2.19) holds.

Note that (2.2.19) states that $M(\hat{\xi}_i)$ is Hermitian w.r.t. $K(\hat{\xi}_i)$. This implies that $M(\hat{\xi}_i)$ is diagonalizable with real eigenvalues. Using $K$ one can define an inner product between solutions, and the corresponding norm can be shown to satisfy an inequality of the form (2.2.9). This is called the *energy estimate*. Using this one can

[7] There are examples of weakly hyperbolic systems for which $||\exp(i\mathcal{M}(\xi_i)t)||$ grows as $\exp(c\sqrt{|\xi|}t)$ for some constant $c > 0$ [53], in which case one could not even obtain a bound of this weaker type.

[8] In our case, since these matrices are real, the hermitian conjugate is simply the transpose.

prove that the initial value problem is locally well-posed independently of the form of the zero derivative term $Cu$ in (2.2.1) [53].

So far the discussion has considered first order linear PDEs with constant coefficients. We can now discuss the case of non-constant coefficients.

Let the matrices $A$, $P^i$ and $C$ in (2.2.1) depend smoothly on time and space, with $A$ still invertible. At some point $(t_0, x_0^i)$ we define the "*frozen coefficients*" equation by fixing $A$, $P^i$ and $C$ to their values at $(t_0, x_0^i)$. It is believed that a necessary condition for local well-posedness of the varying coefficients equation near $(t_0, x_0^i)$ is that the "frozen coefficients" equation should be locally well-posed. The idea behind this is that, since well-posedness is a question about the short wavelength behaviour of the coefficients of the differential operator, we can limit our considerations to small neighbourhoods where the coefficients, being smooth, are approximately constant and hence, intuitively, the non-constant coefficient problem can be reduced to the "frozen coefficients" problem.

For this to be the case, the "frozen coefficients" equation must satisfy the above definitions of weak and strong hyperbolicity. For the varying coefficients equation to be locally well-posed, we need these definitions to be satisfied for all $(t_0, x_0^i)$. This motivates extending the definitions of hyperbolicity to equations with non-constant coefficients in the obvious way: we simply allow $M(t, x, \xi_i)$ and $K(t, x, \xi_i)$ to depend smoothly on $(t, x)$ as well as on $\xi$ [53, 46].

**Definition 2.** The system (2.2.1) is *weakly hyperbolic* if, and only if, all eigenvalues of $M(t, x, \xi_i)$ are *real* for any real $\xi_i$ with $\xi_i \xi_i = 1$.

**Definition 3.** The system (2.2.1) is *strongly hyperbolic* if, and only if, there exists a positive definite Hermitian matrix $K(t, x, \hat{\xi}_i)$ depending smoothly on $t, x, \hat{\xi}_i$, such that

$$K(t, x, \hat{\xi}_i) M(t, x, \hat{\xi}_i) K(t, x, \hat{\xi}_i)^{-1} = M(t, x, \hat{\xi}_i)^\dagger \tag{2.2.20}$$

and a real constant $C > 0$ such that for all $t, x, \hat{\xi}_i$

$$C^{-1} I \leq K(t, x, \hat{\xi}_i) \leq CI.^9 \tag{2.2.21}$$

The latter technical condition is required to obtain an energy estimate — it ensures that $K$ does not behave badly for large $t, x$ e.g. it does not become degenerate or blow up asymptotically.

[9] Given two hermitian matrices, $A$ and $B$, we say that $A \leq B$ iff $u^\dagger A u \leq u^\dagger B u$ for any vector $u$. Condition (2.2.21) is equivalent to saying that the eigenvalues of $K$ must be bounded from above and below.

Finally, for completeness, we include the definition of a symmetric hyperbolic system:

**Definition 4.** The system (2.2.1) is *symmetric hyperbolic* if it is strongly hyperbolic and the matrix $K$ is independent of $\hat{\xi}_i$.

## 2.3 Second order systems

Our treatment of second order systems is based on [46]. Consider again a $d-$dimensional spacetime with coordinates $(t, x^i)$. A second order system of linear PDEs takes the form

$$P^{\mu\nu}\partial_\mu\partial_\nu u + Q^\mu\partial_\mu u + Ru = 0 \tag{2.3.1}$$

where $P^{\mu\nu} = P^{(\mu\nu)}, Q^\mu$ and $R$ are $N \times N$ real matrices. For a 1-form $\xi$, the *principal symbol* of (2.3.1) is the $N \times N$ matrix

$$P(\xi) \equiv P^{\mu\nu}\xi_\mu\xi_\nu. \tag{2.3.2}$$

As in the previous section, we will start by studying the *constant coefficients* case. A spatial Fourier transform of (2.3.1) yields

$$A\tilde{u}_{tt} + \left(iB(\xi_i) + Q^0\right)\tilde{u}_t + \left(-C(\xi_i) + iQ^i\xi_i + R\right)\tilde{u} = 0 \tag{2.3.3}$$

where

$$A = P^{00} \qquad B(\xi_i) = 2\xi_i P^{0i} \qquad C(\xi_i) = P^{ij}\xi_i\xi_j \tag{2.3.4}$$

and we assume that $A$ is invertible, i.e. that surfaces of constant $t$ are non-characteristic (characteristic hypersurfaces will be discussed in detail in the next section). We can re-write this equation in first-order form by introducing the vector

$$\tilde{w}^T = \left(\sqrt{1+|\xi|^2}\tilde{u}, -i\tilde{u}_t\right) \tag{2.3.5}$$

where, as above, $|\xi| = \sqrt{\xi_i\xi^i}$. Eq. (2.3.3) is, in fact, equivalent to

$$\tilde{w}_t = i\mathcal{M}(\xi_i)\tilde{w} \tag{2.3.6}$$

where we have defined the $2N \times 2N$ matrix

$$\mathcal{M}(\xi_i) = \begin{pmatrix} 0 & (1+|\xi|^2)^{1/2}I \\ -(1+|\xi|^2)^{-1/2}A^{-1}\left(C(\xi_i) - iQ^i\xi_i - R\right) & -A^{-1}\left(B(\xi_i) - iQ^0\right) \end{pmatrix}. \tag{2.3.7}$$

Note that the $L^2-$norm of $\tilde{w}$ is a measure of the energy of the field $u$ — it is quadratic in the field and its first derivatives:

$$||\tilde{w}|| \sim ||u|| + ||\partial u||. \tag{2.3.8}$$

In order to prove local well-posedness, it is required that the solution obeys an energy estimate of the form

$$||\tilde{w}||(t) \leq f(t)||\tilde{w}||(0), \tag{2.3.9}$$

for some continuous function $f(t)$ independent of $\xi_i$ and $\tilde{w}$.

The solution to (2.3.6) is given by

$$\tilde{w}(t, \xi_i) = \exp(i\mathcal{M}(\xi_i)t)\tilde{w}(0, \xi_i), \tag{2.3.10}$$

so for the energy estimate to hold for any initial data, we need

$$||\exp(i\mathcal{M}(\xi_i)t)|| \leq f(t). \tag{2.3.11}$$

Similarly to the first order case, it is sufficient to study the high-frequency part of $\mathcal{M}(\xi_i)$. We define $t' = t|\xi|$ and $\hat{\xi}_i|\xi| = \xi_i$ and take the $|\xi| \to \infty$ limit at constant $t'$. The previous equation reduces to

$$||\exp(iM(\hat{\xi}_i)t')|| \leq k \tag{2.3.12}$$

where $k = f(0)$ and

$$M(\xi_i) = \begin{pmatrix} 0 & I \\ -A^{-1}C(\xi_i) & -A^{-1}B(\xi_i) \end{pmatrix}. \tag{2.3.13}$$

We can now repeat the argument we used for a first order system: if $M(\hat{\xi}_i)$ had a complex eigenvalue then we could violate (2.3.11). Hence we define weak hyperbolicity as the condition that all eigenvalues of $M(\hat{\xi}_i)$ be real.

It is possible to formulate this condition in an alternative way. Let $\xi_0$ be an eigenvalue of $M(\xi_i)$ with eigenvector $v = (t, t')^T$, where $t$ is a two-tensor. Writing out the eigenvalue equation $Mv = \xi_0 v$ gives

$$t' = \xi_0 t \qquad (A\xi_0^2 + B(\xi_i)\xi_0 + C)t = 0. \tag{2.3.14}$$

This is a *quadratic eigenvalue problem* with eigenvector $t$. We can rewrite this in terms of the principal symbol

$$P(\xi) \cdot t = 0 \tag{2.3.15}$$

where $\xi_\mu = (\xi_0, \xi_i)$. This equation states that the 1-form $\xi$ is *characteristic* (see Section 2.4 below). Hence we see that the system (2.3.1) is *weakly hyperbolic* if, for any real $\xi_i \neq 0$, a characteristic covector $(\xi_0, \xi_i)$ has real $\xi_0$.

As for first order systems, if the Jordan normal form of $M$ involves non-trivial blocks then equation (2.3.12) cannot hold. So we define strong hyperbolicity just as we did in the previous section: equation (2.3.1) is *strongly hyperbolic* if, and only if, there exists a positive definite Hermitian matrix $K(\hat{\xi}_i)$ depending smoothly on $\hat{\xi}_i$ such that $M(\hat{\xi}_i)$ is Hermitian w.r.t. $K$, i.e., satisfies (2.2.19). This implies that $M(\xi_i) = |\xi|^2 M(\hat{\xi}_i)$ is diagonalizable with real eigenvalues.

Finally, we can extend these results to the non-constant coefficients case just as we did for first order systems. Let the coefficients $P^{\mu\nu}$, $Q^\mu$ and $R$ depend smoothly on $(t, x^i)$. We then define $M(t, x, \xi_i)$ using (2.3.13). As for first order systems, it is believed that local well-posedness implies local well-posedness for the equation with frozen coefficients. Hence we define weak and strong hyperbolicity again as

**Definition 5.** The system (2.3.1) is *weakly hyperbolic* if, and only if, all eigenvalues of $M(t, x, \xi_i)$ are *real* for any real $\xi_i$ with $\xi_i\xi_i = 1$.

Equivalently, the system is weakly hyperbolic if, and only if, any characteristic 1-form $(\xi_0, \xi_i)$ with real $\xi_i \neq 0$ has real $\xi_0$.

**Definition 6.** The system (2.3.1) is *strongly hyperbolic* if, and only if, there exists a positive definite Hermitian matrix $K(t, x, \hat{\xi}_i)$ depending smoothly on $t, x, \hat{\xi}_i$, such that

$$K(t, x, \hat{\xi}_i) M(t, x, \hat{\xi}_i) K(t, x, \hat{\xi}_i)^{-1} = M(t, x, \hat{\xi}_i)^\dagger \tag{2.3.16}$$

and a real constant $C > 0$ such that for all $t, x, \hat{\xi}_i$

$$C^{-1} I \leq K(t, x, \hat{\xi}_i) \leq CI. \tag{2.3.17}$$

Again, if $K$ is independent of $\hat{\xi}_i$, the system is said to be *symmetric hyperbolic.*

In the later chapters we will mainly be interested in showing that certain equations are *not* strongly hyperbolic. We will do this by demonstrating that $M(t, x, \hat{\xi}_i)$ is not diagonalizable. Note that $M$ is determined by $P^{\mu\nu}$, i.e., by the principal symbol. So hyperbolicity depends only on the nature of the second derivative terms in the equation. Furthermore, to demonstrate that $M$ is not diagonalizable it is sufficient to work at a single point in spacetime.

## Non-linear systems

To conclude this section we will briefly discuss how the theory we have developed so far can be applied to non-linear systems.

Non-linear systems will, in general, exhibit a more complicated behaviour. In particular, the non-linearities in the equations may induce the formations of shocks or cause a solution to blow-up in finite time, even if the initial data is smooth and "small" (a famous example of this problem was found by John [54]). For these reasons one would in general expect (at best) only finite time existence of solutions.[10] As mentioned in the introduction, we will only be interested in the local well-posedness of the equations.

Let us consider a non-linear system which admits a well-posed initial value problem. Well-posedness ensures that the solution depends continuously on the initial data and hence that, given some solution $u_0$, a sufficiently small perturbation of the initial data will result in a solution lying in a neighbourhood of $u_0$. Let us consider a 1-parameter family of initial data, parametrised by some $\epsilon$ s.t. it coincides with the initial data of $u_0$ at $\epsilon = 0$. For small enough values of $\epsilon$ we have a unique solution, $u_\epsilon$. Expanding

$$u_\epsilon(t,x) = u_0(t,x) + \epsilon v(t,x) + \mathcal{O}(\epsilon^2) \tag{2.3.18}$$

and plugging it back into the equations of motion we obtain a well-posed system of linear equations for $v$.

In other words, a *necessary* condition for the well-posedenss of the non-linear problem is that the corresponding linearised problems, obtained by linearising around any solution in an open neighbourhood of $u_0$, be well-posed.

Moreover, the converse statement, known as the *linearisation principle* [53] also holds: the non-linear problem is well-posed around some solution, if all the linearised problems obtained by linearising near such a solution are well-posed.

In Chapters 5 and 6 we will investigate whether Lovelock and Horndeski theories, respectively, satisfy this necessary condition for well-posedness.

## 2.4 Causality

Hyperbolic partial differential equations describe phenomena with finite propagation speed. For this reason they define a causal structure on the spacetime: given an event $p$, we can define its *causal future* $J^+(p)$ as the set of events that can be influenced by it, i.e. $q \in J^+(p)$ if a change in the solution at $p$ will result in a change of the solution at $q$. Similarly one can define the *causal past*. In general, these notions depend on the solution itself.

The causal structure of a PDE is closely related to its *characteristic hypersurfaces*. Loosely speaking these are hypersurfaces on which all of the highest order derivatives cannot be determined in terms of all the lower derivatives by the equations of

[10]Famously, Christodoulou and Klainerman showed in 1990 that the "small data" non-linear problem for General Relativity does in fact admit global-in-time solutions [55].

motion. If one has a solution which is discontinuous in the highest derivatives (but continuous in the others) across a hypersurface, then that hypersurface is necessarily characteristic. Hence, discontinuities in the solution propagate along characteristics. More "physically", if one considers high frequency waves propagating on top of a "background" solution then the surfaces of constant phase are characteristic hypersurfaces of the background solution. Furthermore, the hypersurfaces bounding the Cauchy development of some initial data set are characteristic.

More formally, consider a system of $N$ second order non-linear PDEs

$$E_I(p, u, \partial u, \partial^2 u) = 0, \qquad I = 1, 2, \ldots, N. \tag{2.4.1}$$

In the previous section we have defined the principal symbol for a linear system. The principal symbol of a non-linear system will be defined as the principal symbol of the linearised system. Given a 1-form $\xi$, the *principal symbol* of the system (2.4.1) is

$$P(p, \xi)_{IJ} = \frac{\partial E_I}{\partial(\partial_\alpha \partial_\beta u^J)} \xi_\alpha \xi_\beta. \tag{2.4.2}$$

The *characteristic polynomial* is defined as

$$Q(p, \xi) = \det P(p, \xi). \tag{2.4.3}$$

It is a homogeneous polynomial in $\xi$ of degree $2N$.

We will define a hypersurface $\Sigma$ to be *characteristic* if, and only if, its normal 1-form $\xi$ satisfies

$$Q(p, \xi) = 0 \qquad \forall p \in \Sigma. \tag{2.4.4}$$

Consider the case in which $\Sigma$ is defined as a level set of some function $\phi$. It will be characteristic if, and only if, $Q(p, d\phi) = 0$. This defines a first-order non-linear equation for the function $\phi$, known as the *eikonal equation*.

Let us consider the scalar linear wave equation (2.1.8) as a simple example. In this case the principal symbol and the characteristic polynomial take the form

$$P(\xi) = g^{\mu\nu} \xi_\mu \xi_\nu \quad \Rightarrow \quad Q(\xi) = g^{\mu\nu} \xi_\mu \xi_\nu \tag{2.4.5}$$

and hence we deduce that a hypersurface is characteristic if, and only if, it is null w.r.t. the spacetime metric $g$.

Characteristic hypersurfaces are generated by *bicharacteristic curves*. These are defined, in a local chart $\{x^\mu\}$, as the curves $(x(s), \xi(s))$ in the cotangent bundle satisfying

$$\frac{dx^\mu}{ds} = \frac{\partial Q(x, \xi)}{\partial \xi_\mu} \qquad \frac{d\xi_\mu}{ds} = -\frac{\partial Q(x, \xi)}{\partial x^\mu}, \tag{2.4.6}$$

with the initial condition $Q(x(0), \xi(0)) = 0$ which, by construction, is preserved along these curves. For the linear wave equation, or for GR, bicharacteristic curves correspond to null geodesics. The geometric optics approximation tells us that high-frequency perturbations of a solution will move on bicharacteristic curves. For example, the high-frequency modes of a perturbation $\delta\phi + \phi_0$ of some solution $\phi_0$ to the wave equation (2.1.8) will propagate along null geodesics.

At a point $p$, the set of characteristic one-forms defines a convex cone in the cotangent space, known as the *normal cone* (or *characteristic subset* of $T^*_p\mathcal{M}$)

$$\mathcal{C}^*_p = \{\xi \in T^*_p\mathcal{M} \;:\; Q(p,\xi) = 0\}. \tag{2.4.7}$$

The *ray cone* at $p$ is defined as the dual of the normal cone:

$$\mathcal{C}_p = \bigcup_{\xi \in \mathcal{C}^*_p} \mathcal{C}_{p,\xi} \tag{2.4.8}$$

where

$$\mathcal{C}_{p,\xi} = \{X \in T_p\mathcal{M} \;:\; \langle X, \xi \rangle = 0\}. \tag{2.4.9}$$

The projections of the bicharacteristics onto the base manifold are called *rays* (recall that the bicharacteristics are curves in the cotangent bundle). Their tangent vectors

$$Y_p = \frac{\partial Q(p,\xi)}{\partial \xi_\mu}\frac{\partial}{\partial x^\mu} \in \mathcal{C}_p \tag{2.4.10}$$

are called *characteristic tangent vectors*. The ray cone is simply the "light cone" defined by null rays.

Generically, the normal and the ray cones will be composed of several sheets. Physically this can be interpreted as the fact that different degrees of freedom will propagate with different speeds. Theories that exhibit this property are referred to as *multi-refringent*. As we will see more formally below, the fastest propagating mode will determine the causal structure of the theory: two points will be in causal contact if (at least) a "fastest" signal could be sent between them.

In order to discuss causality, we need to introduce a special subset of the cotangent space. Take $\xi \in \mathcal{C}^*_p$ and let $\zeta \in T^*_p\mathcal{M}$ be such that $\lambda \mapsto Q(p, \xi + \lambda\zeta)$ has $m \times n$ real roots, where $m$ is the order of the PDE and $n$ is the rank of the principal symbol.[11] We denote by $C^*_p$ the subset of the cotangent space defined by such one-forms $\zeta$. It can be shown that $C^*_p$ consists of two convex, opposite cones $C^*_p = C^{*,+}_p \cup C^{*,-}_p$. These are non-empty and $\partial C^*_p \subset \mathcal{C}^*_p$. Similarly, denote the dual cones in the tangent space by $C_p = C^+_p \cup C^-_p$. In this case we have $\mathcal{C}_p \subset C_p$. If it is possible to continuously

[11] In the case of GR, it must have $d(d-3)/2$ roots. In other words, it must have as many roots as propagating degrees of freedom. The roots need not be distinct.

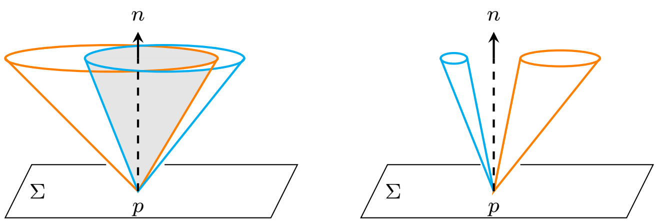


Figure 2.1: Schematic description of the definition of $C^*$. Suppose our theory is such that the normal cone has two sheets, then $C^*$ corresponds to the shaded region in the *left* figure. The hypersurface $\Sigma$ is spacelike, for its normal one-form $n$ lies in $C^*$. In the *right* figure we portray a case in which the theory fails to be hyperbolic, since the sheets of the normal cone have empty intersection.

distinguish between the convex cones $C_p^{\pm}$, we say that the spacetime is *time-orientable*. We can finally define the causal structure:

**Definition 7.** A vector $X \in T_p\mathcal{M}$ is *causal* if $X \in C_p$. A one-form $\xi \in T_p^*\mathcal{M}$ is *causal* if $\xi \in C_p^*$.

It follows, in particular, that a hypersurface $\Sigma$ is *spacelike* if its normal one-form is in the interior of $C_p^*$, $\forall p \in \Sigma$. Thus in discussing the initial value problem we shall prescribe initial data on a hypersurface which is spacelike according to this definition. It is therefore a fundamental requirement for well-posedness that the normal cones have a non-empty intersection, for otherwise a spacelike hypersurface would not exist (Figure 2.1).

Note that for a non-linear system, the principal symbol — and hence the causal cones — will in general depend on the background fields and their derivatives. For this reason, the choice of the initial data may influence the character of the surface on which the data itself is specified. Let us illustrate this problem with a simple example. Consider the following toy model in two-dimensional Minkowski space:

$$(1 + \partial_x^2\varphi)\Box\varphi = 0. \tag{2.4.11}$$

The principal symbol is given by

$$P(\xi) = -\left[1 + \partial_x^2\varphi\right]\xi_t^2 + \left[1 - \partial_t^2\varphi + 2\partial_x^2\varphi\right]\xi_x^2. \tag{2.4.12}$$

Hence the character of a hypersurface of "constant time" $\Sigma_T = \{t = T\}$ (which has normal one-form $\xi \propto dt$) will be determined by the sign of $[1 + \partial_x^2\varphi]$. In particular, consider initial data $(\varphi, \partial_t\varphi)(0, x) = (\varphi_0, \psi_0)(x)$, prescribed on $\Sigma_0$. The equation will be hyperbolic for

$$\frac{1 - \partial_t^2\varphi + 2\partial_x^2\varphi}{1 + \partial_x^2\varphi} > 0 \tag{2.4.13}$$

and the initial surface $\Sigma_0$ will be spacelike only if the initial data is such that

$$(1 + \partial_x^2 \varphi_0) > 0. \tag{2.4.14}$$

From this example we learn that one must be careful in the choice of the initial data, for the character of the initial data surface will in general depend on the initial values of the fields and their derivatives. In other words, in order to have a well-posed initial value problem we must choose initial data for the fields such that the initial surface be spacelike with respect to the causal structure induced by the equations of motion.

## 2.5 The initial value problem in General Relativity

While not evident at first, due to their diffeomorphism-covariant nature, the Einstein equations admit an initial value formulation. In fact, for any diffeomorphism-covariant theory, it is necessary to impose an appropriate gauge fixing condition in order to establish the character of the equations. In her seminal paper [42], Choquet-Bruhat showed that in harmonic gauge[12]

$$\Box_g x^\mu = 0 \tag{2.5.1}$$

the Cauchy problem for the Einstein equations is locally well-posed.

In this gauge, the vacuum Einstein equations reduce to a system of quasilinear wave equations

$$g^{\alpha\beta}\partial_\alpha\partial_\beta g_{\mu\nu} = \mathcal{N}(g, \partial g)_{\mu\nu} \tag{2.5.2}$$

together with a set of elliptic constraint for the initial data. The initial data consists of a triple $(\Sigma, \bar{g}_{ab}, K_{ab})$, where $\Sigma$ is a *spacelike* hypersurface, $\bar{g}_{ab}$ is a Riemannian metric on $\Sigma$ and $K_{ab}$ is the second fundamental form (or extrinsic curvature) of $\Sigma$.[13] A fundamental property of this system is that the gauge condition (2.5.1) is propagated by the evolution equations. This happens because the gauge condition solves a linear wave equation with trivial initial data. In particular, the gauge condition is taken to be satisfied initially, while the constraint equations ensure that its time derivative also vanishes. Once this has been proved, the well-posedness result follows from known results on systems of quasilinear wave equations.

These results were later extended to the global-in-time case by Choquet-Bruhat and Geroch [58], and Sbierski (without recurring to Zorn's lemma) [59] who estab-

[12] $\Box_g x^\mu := \frac{1}{\sqrt{-\det g}}\partial_\alpha(\sqrt{-\det g}\, g^{\alpha\mu})$

[13] Rendall [56] and Luk [57] showed that GR also admits a well-posed *characteristic* initial value problem, where the spacelike hypersurface $\Sigma$ is replaced by two intersecting null hypersurfaces.

lished the existence of a unique *maximal* globally hyperbolic development of given initial data for the Einstein equations.

It is worth stressing the importance of the gauge choice. In harmonic gauge, the Einstein equations admit a well-posed initial value problem. In fact, we will show below that, in this gauge, they are strongly hyperbolic. This property, however, does not necessarily hold in other gauges. Other approaches, such as the ADM formulation [60] of the Einstein equations lead to equations that are only weakly hyperbolic. More precisely, the ADM equations are weakly hyperbolic for any fixed choice of shift and *densitized* lapse [47].[14] This implies that these gauges cannot be used to establish local well-posedness.[15] A further consequence is that these equations are unsuitable for numerical simulations. For numerical simulations, strong hyperbolicity is regarded as essential. The first successful binary black hole simulations [62, 63, 64] employed numerical codes based either on (generalised) harmonic gauge [62] or the BSSN formalism [65, 66, 67]. The latter is a modification of the ADM formalism that can be shown to be strongly hyperbolic [68, 47]. The *generalised harmonic gauge* is obtained by replacing the RHS of Eq. (2.5.1) by a source term [69].

## Hyperbolicity of the Einstein equations

We will now see how these concepts introduced in this Chapter can be used to show that the harmonic gauge Einstein equations are strongly hyperbolic.

The vacuum Einstein equations for a metric $g$ read:[16]

$$G_{ab}[g] = 0. \tag{2.5.3}$$

To compute the principal symbol, we consider a perturbation $g \to g + h$ of some background solution $g$ and linearise in the perturbation

$$G_{ab}[g + h] = G_{ab}[g] + G^{(1)}_{ab}[h] + \dots \tag{2.5.4}$$

The linearised equations read

$$G^{(1)}_{ab}[h] = 0 \tag{2.5.5}$$

[14]The densitized lapse is obtained by rescaling the lapse by $(\sqrt{\det \gamma})^b$, where $b$ is a constant and $\gamma$ is the metric induced on the leaves of the time foliation.

[15]However, there exist strongly hyperbolic modifications of these equations which can be used to establish local well-posedness [61].

[16]Equivalently, we could consider the vacuum Einstein equations written in the form $R_{ab} = 0$. However, in view of our discussion of Lovelock and Horndeski theories later on, it will be convenient to consider the Einstein vacuum equations in the form $G_{ab} = 0$.

where

$$G^{(1)}_{ab}[h] = -\frac{1}{2}\partial_c\partial^c h_{ab} + \partial^c\partial_{(a}h_{b)c} - \frac{1}{2}\partial_a\partial_b h^c{}_c - \frac{1}{2}g_{ab}\left(\partial^c\partial^d h_{cd} - \partial^c\partial_c h^d{}_d\right) + \dots \quad (2.5.6)$$

and the ellipsis includes the lower order derivatives terms. We will regard the principal symbol as a linear map between symmetric rank-2 tensors, i.e., as a $N \times N$ matrix, where $N = d(d+1)/2$. We can deduce its form from the structure of the second order derivatives in the linearised equations of motion

$$(P_{\text{Einstein}}(\xi)\cdot t)_{ab} = -\frac{1}{2}\xi^2 t_{ab} + \xi^c\xi_{(a}t_{b)c} - \frac{1}{2}\xi_a\xi_b t^c{}_c - \frac{1}{2}g_{ab}\xi^c\xi^d t_{cd} + \frac{1}{2}g_{ab}\xi^2 t^c{}_c \quad (2.5.7)$$

Recall that the Einstein equations are diffeomorphism-covariant. This means that two solutions $g_1$ and $g_2$ related by a diffeomorphism $\varphi$

$$g_1 = \varphi^*(g_2)$$

are physically equivalent. At the linearised level, this means that we are free to perform the gauge transformation

$$h_{ab} \to h_{ab} + \nabla_{(a}X_{b)}, \quad (2.5.8)$$

where $X$ is a vector field generating the diffeomorphism $\varphi$. The corresponding gauge symmetry at the level of the principal symbol takes the form

$$t_{ab} \to t_{ab} + \xi_{(a}X_{b)}. \quad (2.5.9)$$

It is easy to see that the principal symbol is invariant under such transformation. For this reason the principal symbol is always degenerate. We will deal with this degeneracy by considering equivalence classes of symmetric tensors, where two tensors $t_{ab}$ and $t'_{ab}$ will be considered equivalent if $t'_{ab} = t_{ab} + \xi_{(a}X_{b)}$. In particular

**Definition 8.** A tensor $t_{ab}$ is *pure gauge* if, and only if, $t_{ab} \sim 0$, i.e., if, and only if,

$$t_{ab} = \xi_{(a}X_{b)}, \quad (2.5.10)$$

for some $X^a$.

We will now proceed to fixing the gauge.[17] At the linearised level, the harmonic gauge condition (2.5.1) reduces to[18]

$$H_a \equiv \nabla^c h_{ac} - \frac{1}{2}\nabla_a h^c{}_c = 0. \quad (2.5.11)$$

[17]Showing that this gauge condition can always be imposed via a gauge transformation (2.5.8) is a standard argument. We refer the reader to the discussion below Eq. (6.2.8) where we present a proof in the more general context of Horndeski gravity.

[18]To be precise, this should be referred to as Lorenz gauge condition. See discussion below Eq. (5.2.5).

The harmonic gauge linearised Einstein equations are given by

$$\tilde{G}^{(1)}_{ab}[h] = 0 \tag{2.5.12}$$

where we have defined

$$\tilde{G}^{(1)}_{ab}[h] \equiv G^{(1)}_{ab}[h] - \nabla_{(a}H_{b)} + \frac{1}{2}g_{ab}\nabla_c H^c \tag{2.5.13}$$

$$= -\frac{1}{2}\left(\partial^c\partial_c h_{ab} - \frac{1}{2}g_{ab}\partial^c\partial_c h^d{}_d\right) + \dots . \tag{2.5.14}$$

To see that the gauge condition (2.5.11) is propagated by the harmonic equations of motion (2.5.12), we take the divergence of $\tilde{G}^{(1)}_{ab}$ when the equations of motion are satisfied, obtaining a linear wave equation for $H_a$

$$\nabla^c\nabla_c H_a + R_{ac}H^c = 0. \tag{2.5.15}$$

Requiring the initial data to satisfy the constraint equations ensures that the time derivative of $H_a$ is initially vanishing [42]. Therefore, if we take initial data such that $H_a = 0$ initially, we are guaranteed that the solution will have $H_a \equiv 0$. Hence, the equations of motion (2.5.12) propagate the harmonic gauge condition and hence any solution will also satisfy the original equations of motion (2.5.5).

The gauge fixed principal symbol can be deduced from Eq. (2.5.14) and takes the form

$$(P_{\text{Einstein}}(\xi)\cdot t)_{ab} = -\frac{1}{2}\xi^2 G_{ab}{}^{cd}t_{cd}, \tag{2.5.16}$$

where

$$G^{abcd} = \frac{1}{2}\left(g^{ac}g^{bd} + g^{ad}g^{bc} - g^{ab}g^{cd}\right). \tag{2.5.17}$$

It is now straightforward to find the characteristics for the Einstein equations. Recall that a hypersurface with normal one-form $\xi$ is characteristic if, and only if, $\xi$ is a root of the characteristic polynomial, i.e., $Q(x,\xi) = 0$. This implies that there exists a non-zero (i.e. not pure gauge) symmetric 2-tensor which belongs to the kernel of the principal symbol

$$P_{\text{Einstein}}(\xi)\cdot t = 0, \qquad t \neq 0. \tag{2.5.18}$$

Thanks to the non-degeneracy of $G^{abcd}$ [70] we can conclude that in General Relativity a hypersurface is characteristic if, and only if, it is null ($\xi^2 = 0$). We see that all the sheets of the normal cone coincide with the null cone of the metric. For this reason causality in General Relativity is determined by the lightcone of the spacetime metric.

We can finally discuss the hyperbolicity of the Einstein equations.

In a chart $x^\mu = (t, x^i)$, consider a *characteristic* covector $\xi_\mu = (\xi_0, \xi_i)$ with $\xi_i$ *real*. Denote by $\xi_0^\pm$ the two solutions of $\xi^2 = g^{\mu\nu}\xi_\mu\xi_\nu = 0$. It is easy to see that these must be real. We can conclude that any characteristic covector with real spatial components has a real time component and hence the Einstein equations are *weakly hyperbolic*.

For a characteristic covector $\xi$, we have that $P_{\text{Einstein}}(\xi) \cdot t = 0$ for any symmetric tensor $t_{ab}$. In the language of Section 2.3 this means that, for the Einstein equations, the $2N \times 2N$ matrix $M$ has two real eigenvalues, $\xi_0^\pm$, to which are associated the eigenvectors $(t, \xi_0^\pm t)$. Since there are $N$ linearly independent eigenvectors associated to each eigenvalue, the matrix $M$ has $2N$ linearly independent eigenvectors and hence is diagonalisable. We conclude that the Einstein equations are *strongly hyperbolic* in harmonic gauge, independently of the background solution $g$.

# Chapter 3

# Beyond Einstein's theory

In this chapter we will introduce the main subjects of the dissertation: Lovelock and Horndeski theories.

## 3.1 Lovelock's theorem

In General Relativity the gravitational field is described by the geometry of spacetime. The Einstein field equations relate the curvature of the spacetime to its matter content

$$G_{ab} + \Lambda g_{ab} = 8\pi T_{ab}. \tag{3.1.1}$$

A natural question to ask is whether the choice of the LHS of the field equations is unique. In order to answer this question, let us examine some of the properties of these equations.

The LHS encodes the information relative to the gravitational field. The Einstein tensor is symmetric and depends only on the metric and its first and second derivatives. In other words, the Einstein equations are of *second order* and the gravitational field is mediated solely by the metric tensor. Moreover, these equations can be obtained by an action principle.

A second important property is that, thanks to the diffeomorphism-covariant nature of the theory (cf. general covariance principle), both sides of the Einstein equations must be divergence-free when the the equations of motion for the matter fields are satisfied. That this condition does not impose any further constraints on the spacetime is a consequence of the differential Bianchi identities which ensure that the divergence of the Einstein tensor is always identically vanishing. More generally, this property is a consequence of requiring the field equation to arise as the metric variation of a diffeomorphism-invariant action.

Going back to our initial question, Lovelock proved in 1971 [12] that, essentially, the only theory which satisfies such properties is, in fact, General Relativity.

**Theorem 1** (Lovelock [11]). *In four spacetime dimensions, the unique symmetric rank-2, divergence-free tensors depending only on the metric and its first and second derivatives are the Einstein tensor and the metric tensor itself.*

This means that the only terms we can write on the LHS of (3.1.1) are the Einstein tensor and a cosmological constant term.

In order to obtain a different theory of gravity we will need to relax some of the assumptions of Theorem 1. If we require the theories to have second-order field equations (establishing the well-posedness of higher derivative theories is, in general, a much harder problem; moreover, higher derivative theories are often plagued by Ostrogradsky instabilities [71, 72]), to be diffeomorphism-covariant and to arise from an action principle, we are left with two possible ways of modifying GR:

(i) we can consider spacetimes with more than four dimensions;

(ii) we can consider additional degrees of freedom mediating the gravitational field (in four spacetime dimensions).

In the first case, assuming that we have no other degree of freedom than the metric tensor, we will obtain *Lovelock theories*. The second case is richer, since one could add scalar, vector or tensor degrees of freedom. We will however only consider the simplest case, i.e., the addition of a single, non-minimally coupled scalar field. This will lead to *Horndeski theories*.

In the following sections, we will discuss in more details what these theories are and illustrate some of their properties.

## 3.2 Lovelock theories of gravity

As discussed in the Introduction, modern developments in Theoretical Physics have resulted in an increased interest in higher dimensions. If spacetime has more than four dimensions, Lovelock's theorem does not hold and hence GR will not be the unique theory satisfying the other assumptions. However, Lovelock proved that a modified version of Theorem 1 holds.

**Theorem 2** (Lovelock [12]). *The most general symmetric rank-2 tensor, which is divergence-free and only depends on the metric and its first and second derivatives is the* Lovelock tensor

$$A^a{}_b \equiv \sum_{p\geq 0} k_p \delta^{ac_1\dots c_{2p}}_{bd_1\dots d_{2p}} R_{c_1c_2}{}^{d_1d_2}\cdots R_{c_{2p-1}c_{2p}}{}^{d_{2p-1}d_{2p}}, \tag{3.2.1}$$

*where $k_p$ are dimensionful coupling constants:* $[k_p] = [\text{Length}]^{2(p-1)}$.

The *generalised Kronecker delta* is defined as

$$\delta^{ac_1\ldots c_p}_{bd_1\ldots d_p} = p!\delta^{[a}_b\delta^{c_1}_{c_2}\cdots\delta^{c_p]}_{d_p}. \tag{3.2.2}$$

*Lovelock theories* are the most general diffeomorphism-covariant theories of gravity involving only the metric tensor and having second order equations of motion:

$$A_{ab} = 8\pi T_{ab}. \tag{3.2.3}$$

We will assume that the coupling constants are normalised as

$$k_0 = \Lambda \qquad k_1 = -\frac{1}{4} \tag{3.2.4}$$

so that the first two terms of (3.2.1) coincide with those appearing in the Einstein equations

$$A^a{}_b = \Lambda\delta^a{}_b + G^a{}_b + \sum_{p\geq 2} k_p\delta^{ac_1\ldots c_{2p}}_{bd_1\ldots d_{2p}}R_{c_1c_2}{}^{d_1d_2}\cdots R_{c_{2p-1}c_{2p}}{}^{d_{2p-1}d_{2p}}. \tag{3.2.5}$$

The antisymmetry of the generalised Kronecker delta ensures that the sum is finite, since all terms with $2p > (d-1)$ vanish identically. This implies that in $d = 4$, Lovelock theories automatically reduce to GR.

*Remark.* Note that if one were to require this tensor to be linear in the second derivatives of the metric, then the only possibility would be the Einstein tensor, in *any* number of spacetime dimensions [73, 74], i.e. GR is the unique diffeomorphism-covariant metric theory of gravity with quasilinear second order equations of motion.

The Lagrangian density for Lovelock theories is given by

$$\mathcal{L} = -\sum_{p\geq 0} 2k_p\mathcal{L}_p = \mathcal{L}_{\text{EH}} - \sum_{p\geq 2} 2k_p\mathcal{L}_p \tag{3.2.6}$$

where

$$\mathcal{L}_{\text{EH}} = R - 2\Lambda \tag{3.2.7}$$

is the usual *Einstein-Hilbert* term and

$$\mathcal{L}_p = \delta^{c_1\ldots c_{2p}}_{d_1\ldots d_{2p}}R_{c_1c_2}{}^{d_1d_2}\cdots R_{c_{2p-1}c_{2p}}{}^{d_{2p-1}d_{2p}}. \tag{3.2.8}$$

Each Lagrangian density $\mathcal{L}_p$ corresponds to the *generalised Euler density* of a $2p$–dimensional manifold, and hence the Chern–Gauss–Bonnet theorem [75] implies that the $p^{\text{th}}$ Lovelock term of the action becomes topological — that is to say, the corresponding term in the action is a topological invariant which does not yield any contribution to the equations of motion — in $d = 2p$.[19] For this reason, in $d = 4$ the Lovelock action reduces to the Einstein–Hilbert action.

[19] Analogously, the Einstein-Hilbert term becomes topological in $d = 2$.

## Principal symbol

As discussed in Chapter 2, the hyperbolicity of the theory will be related to some algebraic properties of the principal symbol. Furthermore, the principal symbol also encodes the information regarding the causal structure of the theory. Hence we will now compute its form (without fixing the gauge). Recall that, in a chart $\{x^\mu\}$, the principal symbol is given by

$$P(x,\xi)_{\mu\nu}{}^{\rho\sigma} = \frac{\delta E_{\mu\nu}}{\delta(\partial_\alpha \partial_\beta g_{\rho\sigma})}\xi_\alpha \xi_\beta. \tag{3.2.9}$$

We will regard it as a linear map between symmetric rank-2 tensors. To compute it, we will consider a perturbation $g + \delta g$ of some background solution $g$ and look at the linearised equations for the perturbation. In the above chart we have

$$-\sum_{p\geq 1} 2k_p \delta^{\mu\rho_1\ldots\rho_{2p}}_{\nu\sigma_1\ldots\sigma_{2p}} (\partial_{\rho_1}\partial^{\sigma_1}\delta g_{\rho_2}{}^{\sigma_2}) R_{\rho_3\rho_4}{}^{\sigma_3\sigma_4}\cdots R_{\rho_{2p-1}\rho_{2p}}{}^{\sigma_{2p-1}\sigma_{2p}} + \ldots = 0, \tag{3.2.10}$$

where the ellipsis denotes terms not involving second derivative of the perturbation. We can then read off the principal symbol [70]

$$(P(\xi)\cdot t)^a{}_b = \sum_{p\geq 1} \delta^{ac_1\ldots c_{2p}}_{bd_1\ldots d_{2p}} \xi_{c_1}\xi^{d_1} t_{c_2}{}^{d_2} R_{c_3c_4}{}^{d_3d_4}\cdots R_{c_{2p-1}c_{2p}}{}^{d_{2p-1}d_{2p}}. \tag{3.2.11}$$

It will be convenient to separate the terms coming from the Einstein equations to those coming from the Lovelock contribution

$$P(\xi)^{abcd} = P_{\text{Einstein}}(\xi)^{abcd} + \delta P(\xi)^{abcd}, \tag{3.2.12}$$

where

$$P_{\text{Einstein}}(\xi)_{ab}{}^{cd} t_{cd} = -\frac{1}{2}\left(\xi^2 t_{ab} - 2\xi^c \xi_{(a}t_{b)c} + \xi_a\xi_b t^c{}_c + g_{ab}\xi^c\xi^d t_{cd} - g_{ab}\xi^2 t^c{}_c\right) \tag{3.2.13}$$

$$\delta P(\xi)^a{}_b{}^{cd} t_{cd} = \sum_{p\geq 2} \delta^{ac_1\ldots c_{2p}}_{bd_1\ldots d_{2p}} \xi_{c_1}\xi^{d_1} t_{c_2}{}^{d_2} R_{c_3c_4}{}^{d_3d_4}\cdots R_{c_{2p-1}c_{2p}}{}^{d_{2p-1}d_{2p}}. \tag{3.2.14}$$

*Remark.* This is the general form of the principal symbol *before* fixing the gauge. In order to discuss the initial value problem and the hyperbolicity of the equations, we will need to fix the gauge. In Section 5.2 we will derive the *harmonic gauge* linearised Lovelock equations and their principal symbol.

**Properties of the principal symbol**

The principal symbol enjoys the following symmetries [70]:

$$\delta P^{abcdef} = \delta P^{cdabef} \tag{3.2.15}$$

and

$$\delta P^{(a|bcd|ef)} = \delta P^{a(bc|de|f)} = 0. \tag{3.2.16}$$

It follows that

$$\xi_a \delta P^{abcd}(\xi) = \xi_b \xi_c \xi_f \delta P^{abcdef} = 0. \tag{3.2.17}$$

The validity of these identities can be shown by direct computation from (3.2.14). However, they are a consequence of the gauge symmetry of the theory and the fact that $\delta P$ is not affected by gauge fixing. We will discuss this in detail in Section 6.3.

Note also that the antisymmetrisation in (3.2.14) ensures that the principal symbol be gauge invariant.

## Causality

As discussed in Chapter 2 the causal structure of the theory is determined by the principal symbol. We see that the principal symbol of Lovelock gravity, Eq. 3.2.14, depends on the curvature of the spacetime. This implies that the speed of propagation of gravitational perturbations is determined not only by the metric — as it happens for photons — but also by its first and second derivatives. In particular, this means that gravitons do not necessarily propagate at the speed of light. In fact, it was proven in Ref. [70] that gravitons generically propagate on non-null curves. Note, however, that in Minkowksi space the vanishing of the Riemann curvature tensor implies that gravitons will propagate on null curves (w.r.t. the spacetime metric).

We can illustrate these facts with an example. In Lovelock gravity, there exist *non-generic* spacetimes for which the characteristic polynomial factors into a product of quadratic factors:

$$Q(p,\xi) = \prod_I (G_I^{ab}\xi_a\xi_b)^{p_I}. \tag{3.2.18}$$

It was shown in Ref. [70] that both Ricci flat Type N and static, spherically symmetric spacetimes belong to this class.

Clearly, the characteristic polynomial has a root whenever

$$G_I^{ab}\xi_a\xi_b = 0, \tag{3.2.19}$$

that is to say, a hypersurface is characteristic if, and only if, it is *null* with respect to any of these effective metrics. Similarly, a hypersurface is spacelike if, and only if, it is spacelike with respect to *all* the effective metrics. Each $G_I$ can be interpreted as an (inverse) "effective metric", governing the propagation of a certain "mode" of the graviton. Since different polarisations of the graviton propagate at different speed, we see that Lovelock theories are multi-refringent.

Each effective metric defines through (3.2.19) a "causal cone". These causal cones, together with the causal cone of the spacetime metric form a nested set. Modes whose cone lies inside (outside) the lightcone of the metric will propagate subluminally (superluminally). We will discuss these topics in more detail in Chapter 4.

## 3.3 Horndeski theories of gravity

If we consider a four-dimensional spacetime, Lovelock's Theorem tells us that to obtain a theory of gravity different from GR, whilst still having second order equations of motion, we will need to add extra degrees of freedom. The simplest way to do so is to add a single extra *scalar* degree of freedom. Theories in which gravity is mediated by the metric and a scalar field are known as *scalar-tensor* theories. In the Introduction we discussed several applications of scalar-tensor theories in Cosmology and Astrophysics. Furthermore, scalar-tensor theories can also arise in the dimensional reduction of higher dimensional theories, such as Lovelock theories.

*Horndeski theories* are the most general diffeomorphism-covariant scalar-tensor theories of gravity with second order equations. [21]. The action for these theories has the following form

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} (\mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4 + \mathcal{L}_5) \tag{3.3.1}$$

where

$$\mathcal{L}_1 = R + X - V(\Phi) \tag{3.3.2}$$

$$\mathcal{L}_2 = \mathcal{G}_2(\Phi, X) \tag{3.3.3}$$

$$\mathcal{L}_3 = \mathcal{G}_3(\Phi, X) \Box \Phi \tag{3.3.4}$$

$$\mathcal{L}_4 = \mathcal{G}_4(\Phi, X) R + \partial_X \mathcal{G}_4(\Phi, X) \delta^{ac}_{bd} \nabla_a \nabla^b \Phi \, \nabla_c \nabla^d \Phi \tag{3.3.5}$$

$$\mathcal{L}_5 = \mathcal{G}_5(\Phi, X) G_{ab} \nabla^a \nabla^b \Phi - \frac{1}{6} \partial_X \mathcal{G}_5(\Phi, X) \delta^{ace}_{bdf} \nabla_a \nabla^b \Phi \, \nabla_c \nabla^d \Phi \, \nabla_e \nabla^f \Phi \tag{3.3.6}$$

and we have defined $X = -\frac{1}{2}(\nabla\Phi)^2$. The functions $\mathcal{G}_i(\Phi, X)$ are arbitrary. We will assume that the combinations of these functions that appear in the equations of motion below be smooth.

The first Lagrangian density, $\mathcal{L}_1$ corresponds to the usual Einstein–scalar field theory, i.e. gravity minimally coupled to a scalar field $\Phi$ with potential $V(\Phi)$.

The Horndeski equations of motion can then be obtained by varying the action (3.3.1) with respect to the metric and the scalar field

$$E^{ab}[g,\Phi] \equiv -\frac{1}{\sqrt{-\det g}}\frac{\delta S}{\delta g_{ab}} = 0 \tag{3.3.7}$$

$$E_{\Phi}[g,\Phi] \equiv -\frac{1}{\sqrt{-\det g}}\frac{\delta S}{\delta \Phi} = 0. \tag{3.3.8}$$

In the most general case, these read

$$\begin{aligned}
E^a{}_b \equiv &-\frac{1}{4}\left[1+\mathcal{G}_4-2X\partial_X\mathcal{G}_4+X\partial_\Phi\mathcal{G}_5\right]\delta^{ac_1c_2}_{bd_1d_2}R_{c_1c_2}{}^{d_1d_2}\\
&+\frac{1}{4}\left[\partial_X\mathcal{G}_4-\partial_\Phi\mathcal{G}_5\right]\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\nabla_{c_1}\Phi\nabla^{d_1}\Phi R_{c_2c_3}{}^{d_2d_3}\\
&-\frac{1}{4}\left[X\partial_X\mathcal{G}_5\right]\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\nabla_{c_1}\nabla^{d_1}\Phi R_{c_2c_3}{}^{d_2d_3}\\
&-\frac{1}{2}\left(2X+\mathcal{G}_2+2X\partial_\Phi\mathcal{G}_3+4X\partial^2_\Phi\mathcal{G}_4\right)\delta^a_b\\
&-\frac{1}{2}\left(2+\partial_X\mathcal{G}_2+2\partial_\Phi\mathcal{G}_3+2\partial^2_\Phi\mathcal{G}_4\right)\nabla^a\Phi\nabla_b\Phi\\
&+\left[X\partial_X\mathcal{G}_3+\partial_\Phi\mathcal{G}_4+2X\partial^2_{X\Phi}\mathcal{G}_4\right]\delta^{ac}_{bd}\nabla_c\nabla^d\Phi\\
&+\frac{1}{2}\left[4\partial^2_{X\Phi}\mathcal{G}_4+\partial_X\mathcal{G}_3-\partial^2_\Phi\mathcal{G}_5\right]\delta^{ac_1c_2}_{bd_1d_2}\nabla_{c_1}\Phi\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi\\
&+\frac{1}{2}\left[\partial_X\mathcal{G}_4+2X\partial^2_X\mathcal{G}_4-\partial_\Phi\mathcal{G}_5-X\partial^2_{X\Phi}\mathcal{G}_5\right]\delta^{ac_1c_2}_{bd_1d_2}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi\\
&+\frac{1}{2}\left[\partial^2_X\mathcal{G}_4-\partial^2_{X\Phi}\mathcal{G}_5\right]\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\Phi\nabla^{d_3}\Phi\\
&-\frac{1}{6}[\partial_X\mathcal{G}_5+X\partial^2_X\mathcal{G}_5]\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\nabla^{d_3}\Phi=0
\end{aligned} \tag{3.3.9}$$

and

$$
\begin{aligned}
E_\Phi \equiv &- [1 + \partial_X \mathcal{G}_2 + 2X\partial_X^2 \mathcal{G}_2 + 2\partial_\Phi \mathcal{G}_3 + 2X\partial_{X\Phi}^2 \mathcal{G}_3]\Box\Phi \\
&- [\partial_X^2 \mathcal{G}_2 + 2\partial_{X\Phi}^2 \mathcal{G}_3 + 2\partial_{X\Phi\Phi}^3 \mathcal{G}_4]\delta^{c_1c_2}_{d_1d_2}\nabla_{c_1}\Phi\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi \\
&- [\partial_X \mathcal{G}_3 + X\partial_X^2 \mathcal{G}_3 + 2X\partial_{XX\Phi}^3 \mathcal{G}_4 + 3\partial_{X\Phi}^2 \mathcal{G}_4]\delta^{c_1c_2}_{d_1d_2}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi \\
&- \frac{1}{4}[\partial_X \mathcal{G}_3 + 4\partial_{X\Phi}^2 \mathcal{G}_4 - \partial_\Phi^2 \mathcal{G}_5]\delta^{c_1c_2c_3}_{d_1d_2d_3}\nabla_{c_1}\Phi\nabla^{d_1}\Phi R_{c_2c_3}{}^{d_2d_3} \\
&- [X\partial_X \mathcal{G}_3 + \partial_\Phi \mathcal{G}_4 + 2X\partial_{X\Phi}^2 \mathcal{G}_4]R \\
&- \frac{1}{2}[\partial_X^2 \mathcal{G}_3 + 4\partial_{XX\Phi}^3 \mathcal{G}_4 - \partial_{X\Phi\Phi}^3 \mathcal{G}_5]\delta^{c_1c_2c_3}_{d_1d_2d_3}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\Phi\nabla^{d_3}\Phi \\
&- \frac{1}{2}[\partial_X \mathcal{G}_4 + 2X\partial_X^2 \mathcal{G}_4 - \partial_\Phi \mathcal{G}_5 - X\partial_{X\Phi}^2 \mathcal{G}_5]\delta^{c_1c_2c_3}_{d_1d_2d_3}\nabla_{c_1}\nabla^{d_1}\Phi R_{c_2c_3}{}^{d_2d_3} \\
&- \frac{1}{2}[\partial_X^2 \mathcal{G}_4 - \partial_{X\Phi}^2 \mathcal{G}_5]\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\Phi\nabla^{d_2}\Phi R_{c_3c_4}{}^{d_3d_4} \\
&- \frac{1}{3}\left[3\partial_X^2 \mathcal{G}_4 + 2X\partial_X^3 \mathcal{G}_4 - 2\partial_{X\Phi}^2 \mathcal{G}_5 - X\partial_{XX\Phi}^3 \mathcal{G}_5\right]\delta^{c_1c_2c_3}_{d_1d_2d_3}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\nabla^{d_3}\Phi \\
&- \frac{1}{3}[\partial_X^3 \mathcal{G}_4 - \partial_{XX\Phi}^3 \mathcal{G}_5]\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\nabla^{d_3}\Phi\nabla_{c_4}\Phi\nabla^{d_4}\Phi \\
&+ \frac{1}{12}[2\partial_X^2 \mathcal{G}_5 + X\partial_X^3 \mathcal{G}_5]\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\nabla^{d_3}\Phi\nabla_{c_4}\nabla^{d_4}\Phi \\
&+ \frac{1}{4}[\partial_X \mathcal{G}_5 + X\partial_X^2 \mathcal{G}_5]\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi R_{c_3c_4}{}^{d_3d_4} \\
&+ \frac{1}{16}X\partial_X \mathcal{G}_5 \delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}R_{c_1c_2}{}^{d_1d_2}R_{c_3c_4}{}^{d_3d_4} \\
&+ 2X(\partial_\Phi^2 \mathcal{G}_3 + \partial_{X\Phi}^2 \mathcal{G}_2) - \partial_\Phi \mathcal{G}_2 = 0.
\end{aligned} \tag{3.3.10}
$$

## Principal symbol

As in the previous section on Lovelock theories, we will compute the principal symbol for Horndeski theories. The linearised equations of motion take the form

$$
P_{gg}^{abcdef}\nabla_e\nabla_f h_{cd} + P_{g\Phi}^{abef}\nabla_e\nabla_f\psi + \ldots = 0 \tag{3.3.11}
$$

$$
P_{\Phi g}^{cdef}\nabla_e\nabla_f h_{cd} + P_{\Phi\Phi}^{ef}\nabla_e\nabla_f\psi + \ldots = 0, \tag{3.3.12}
$$

where the ellipses denote terms with fewer than two derivatives of the perturbation. We can then define the principal symbol for this system

$$
P(\xi) = \begin{pmatrix} P_{gg}^{abcdef}\xi_e\xi_f & P_{g\Phi}^{abef}\xi_e\xi_f \\ P_{\Phi g}^{cdef}\xi_e\xi_f & P_{\Phi\Phi}^{ef}\xi_e\xi_f \end{pmatrix}. \tag{3.3.13}
$$

We think of it as acting on vectors of the form $(t_{cd}, \alpha)^T$, where $t_{cd}$ is a symmetric 2-tensor and $\alpha$ is a number.

Similarly to the Lovelock case, the principal symbol can be separated into an Einstein–scalar field part and a Horndeski part

$$
P(\xi) = P_{\mathrm{Esf}}(\xi) + \delta\tilde{P}(\xi). \tag{3.3.14}
$$

The principal symbol for the Einstein–scalar field equations reads

$$P_{\text{Esf}}(\xi) = \begin{pmatrix} P_{\text{Einstein}}(\xi) & 0 \\ 0 & -\xi^2 \end{pmatrix}, \tag{3.3.15}$$

where $P_{\text{Einstein}}(\xi)$ is defined in Eq. (3.2.13). The Horndeski terms read

$$\begin{aligned}
(\delta\tilde{P}_{gg}(\xi)\cdot t)^a{}_b = &-\frac{1}{2}[\mathcal{G}_4 - 2X\partial_X\mathcal{G}_4 + X\partial_\Phi\mathcal{G}_5]\delta^{ac_1c_2}_{bd_1d_2}\xi_{c_1}\xi^{d_1}t_{c_2}{}^{d_2} \\
&-\frac{1}{2}[\partial_X\mathcal{G}_4 - \partial_\Phi\mathcal{G}_5]\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\xi_{c_1}\xi^{d_1}t_{c_2}{}^{d_2}\nabla_{c_3}\Phi\nabla^{d_3}\Phi \\
&+\frac{1}{2}X\partial_X\mathcal{G}_5\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\xi_{c_1}\xi^{d_1}t_{c_2}{}^{d_2}\nabla_{c_3}\nabla^{d_3}\Phi
\end{aligned} \tag{3.3.16a}$$

$$\begin{aligned}
\delta\tilde{P}_{g\Phi}(\xi)^a{}_b = &-\frac{1}{4}\left[X\partial_X\mathcal{G}_5\right]\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\xi_{c_1}\xi^{d_1}R_{c_2c_3}{}^{d_2d_3} \\
&+\left[X\partial_X\mathcal{G}_3 + \partial_\Phi\mathcal{G}_4 + 2X\partial^2_{X\Phi}\mathcal{G}_4\right]\delta^{ac}_{bd}\xi_c\xi^d \\
&+\frac{1}{2}\left[4\partial^2_{X\Phi}\mathcal{G}_4 + \partial_X\mathcal{G}_3 - \partial^2_\Phi\mathcal{G}_5\right]\delta^{ac_1c_2}_{bd_1d_2}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\Phi\nabla^{d_2}\Phi \\
&+\left[\partial_X\mathcal{G}_4 + 2X\partial^2_X\mathcal{G}_4 - \partial_\Phi\mathcal{G}_5 - X\partial^2_{X\Phi}\mathcal{G}_5\right]\delta^{ac_1c_2}_{bd_1d_2}\xi_{c_1}\xi^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi \\
&+\left[\partial^2_X\mathcal{G}_4 - \partial^2_{X\Phi}\mathcal{G}_5\right]\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\Phi\nabla^{d_3}\Phi \\
&-\frac{1}{2}[\partial_X\mathcal{G}_5 + X\partial^2_X\mathcal{G}_5]\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\nabla^{d_3}\Phi
\end{aligned} \tag{3.3.16b}$$

$$\delta\tilde{P}_{\Phi g}(\xi)^a{}_b = \delta\tilde{P}_{g\Phi}(\xi)^a{}_b \tag{3.3.16c}$$

$$\begin{aligned}
\delta\tilde{P}_{\Phi\Phi}(\xi) = &-[1 + \partial_X\mathcal{G}_2 + 2X\partial^2_X\mathcal{G}_2 + 2\partial_\Phi\mathcal{G}_3 + 2X\partial^2_{X\Phi}\mathcal{G}_3]\xi^2 \\
&-[\partial^2_X\mathcal{G}_2 + 2\partial^2_{X\Phi}\mathcal{G}_3 + 2\partial^3_{X\Phi\Phi}\mathcal{G}_4]\delta^{c_1c_2}_{d_1d_2}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\Phi\nabla^{d_2}\Phi \\
&-2[\partial_X\mathcal{G}_3 + X\partial^2_X\mathcal{G}_3 + 2X\partial^3_{XX\Phi}\mathcal{G}_4 + 3\partial^2_{X\Phi}\mathcal{G}_4]\delta^{c_1c_2}_{d_1d_2}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\nabla^{d_2}\Phi \\
&-[\partial^2_X\mathcal{G}_3 + 4\partial^3_{XX\Phi}\mathcal{G}_4 - \partial^3_{X\Phi\Phi}\mathcal{G}_5]\delta^{c_1c_2c_3}_{d_1d_2d_3}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\Phi\nabla^{d_3}\Phi \\
&-\frac{1}{2}[\partial_X\mathcal{G}_4 + 2X\partial^2_X\mathcal{G}_4 - \partial_\Phi\mathcal{G}_5 - X\partial^2_{X\Phi}\mathcal{G}_5]\delta^{c_1c_2c_3}_{d_1d_2d_3}\xi_{c_1}\xi^{d_1}R_{c_2c_3}{}^{d_2d_3} \\
&-\frac{1}{2}[\partial^2_X\mathcal{G}_4 - \partial^2_{X\Phi}\mathcal{G}_5]\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\Phi\nabla^{d_2}\Phi R_{c_3c_4}{}^{d_3d_4} \\
&-\left[3\partial^2_X\mathcal{G}_4 + 2X\partial^3_X\mathcal{G}_4 - 2\partial^2_{X\Phi}\mathcal{G}_5 - X\partial^3_{XX\Phi}\mathcal{G}_5\right]\delta^{c_1c_2c_3}_{d_1d_2d_3}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\nabla^{d_3}\Phi \\
&-[\partial^3_X\mathcal{G}_4 - \partial^3_{XX\Phi}\mathcal{G}_5]\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\nabla^{d_3}\Phi\nabla_{c_4}\Phi\nabla^{d_4}\Phi \\
&+\frac{1}{3}[2\partial^2_X\mathcal{G}_5 + X\partial^3_X\mathcal{G}_5]\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\xi_{c_1}\xi^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\nabla_{c_3}\nabla^{d_3}\Phi\nabla_{c_4}\nabla^{d_4}\Phi \\
&+\frac{1}{2}[\partial_X\mathcal{G}_5 + X\partial^2_X\mathcal{G}_5]\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\nabla^{d_2}\Phi R_{c_3c_4}{}^{d_3d_4}.
\end{aligned} \tag{3.3.16d}$$

*Remark.* As in the previous section, this is the general form of the principal symbol *before* fixing the gauge. In Section 6.2 we will derive the *generalised harmonic gauge* linearised Horndeski equations and their principal symbol.

The principal symbol of Horndeski theory satisfies similar symmetries as that of Lovelock theory, i.e. Equations (3.2.15)–(3.2.17). We will prove these properties in Section 6.3.

# Chapter 4

# Graviton time delay in Einstein–Gauss–Bonnet gravity

The contents of this chapter are the result of original research conducted in collaboration with my supervisor, Harvey Reall, and have been published in [43].

## Introduction

In Lovelock theories, gravitons do not generally propagate at the speed of light [76, 77, 70]. At each point, the speed of propagation depends on the curvature of the spacetime, as well as the metric and the direction of propagation. Moreover the speed will, in general, also depend on the polarisation of the graviton, i.e. these theories are *multi-refringent*.

Consider the special case of gravitons propagating on flat Minkowski space. In this case the Riemann curvature tensor vanishes and the gravitons propagate at the speed of light even in Lovelock gravity. On a flat background, the causal structure of Lovelock gravity coincides with that of General Relativity.

Since gravitons can propagate faster than light in curved space, this property suggests that it could in principle be possible to send a signal between two points in curved space "faster" than if the two points were in flat space.[20] In other words, it seems that one could observe a *negative* Shapiro time delay (i.e. a *time advance*) [78].

In 2014, Camanho, Edelstein, Maldacena and Zhiboedov (CEMZ) studied the Shapiro time delay in Einstein–Gauss–Bonnet gravity (EGB, the simplest Lovelock

[20]Note that it is not always possible to perform a unique identification between points on a curved manifold and points in Minkowski space. This point will be discussed in more detail later on.

theory) and showed that it is indeed possible to observe a time advance. They considered gravitons propagating on the Aichelburg–Sexl (AS) [79, 80] shock-wave spacetime, an exact solution to EGB theory which can be interpreted as the geometry describing the gravitational field of a highly-boosted particle. They showed that for an appropriate choice of polarisation and impact parameter, a graviton scattering off the shock-wave can experience a time advance. Furthermore, they showed that the same result is also obtained from a scattering amplitude calculation.

These results should however be treated with care. The background spacetime considered — the AS solution — is singular: curvature presents a delta function supported on the worldline of the particle generating the shock. For this reason, one must ask whether such solution can be considered "physical". Usually it is regarded as such since it can be obtained as the limit of smooth black hole solutions. Consider a boosted black hole, taking the boost to infinity, while scaling the mass of the black hole to zero and keeping the energy constant yields the shock-wave spacetime. One can then "regulate" the shock-wave solution by replacing it with a small, highly boosted black hole.

This is a valid motivation in GR, but it suffers problems in EGB theory. First of all, "small" black holes are unphysical in $d = 5, 6$:[21] it was shown in [70] that in this case small black holes present a region outside the event horizon in which the equations are not hyperbolic. Moreover in $d = 5$ one cannot consider black holes of arbitrarily small mass, for there exists a mass gap. In $d \geq 7$, instead, we will show that one cannot boost a black hole arbitrarily fast: considering initial data describing a small boosted black hole, we will see that if the boost is too large then the surface on which the initial data is specified will not be everywhere spacelike. This implies that we cannot have a well-posed initial value problem. In particular, if we consider a generic smooth perturbation of the initial data, then the initial value problem with this initial data will have no solution. This means that we cannot consider the solution as physical since it cannot arise as the Cauchy development of a "good" initial data set without fine tuning. We thus deduce that in EGB theory it is not possible to boost a small black hole arbitrarily close to the speed of light: there is a speed limit. These reasons suggest that the AS geometry may not be physical and that an independent confirmation of the possibility of time advance would be desirable.

In this chapter we will study the Shapiro time delay for gravitons in static, spherically symmetric, black hole solutions of EGB theory. A further motivation for considering these spacetime comes from the fact that, in general, it is not possible to

[21] Here "small" is defined in comparison to the length scale set by the EGB coupling constant.

define the Shapiro time delay in a gauge-invariant manner [81]. The problem lies in the fact that there is no gauge-invariant way of identifying points of a curved spacetime with points in Minkowski space. It is however possible to overcome such problem if the spacetime enjoys sufficient symmetry. In particular, for static, spherically symmetric spacetimes it *is* possible to define the Shapiro time delay *unambiguously* [82].[22] The idea is to consider a spherical cavity surrounding the black hole and calculate the proper time it takes for a graviton to cross the cavity. This is then compared to the time it takes for the graviton to travel between the corresponding points of a spherical cavity in flat spacetime.

The main result of this Chapter will be to confirm that gravitons can indeed experience a time advance when they propagate around small EGB black holes. This occurs for gravitons of a certain polarisation incident on the black hole with an impact parameter comparable to the length scale set by the Gauss–Bonnet coupling, analogously to what was found by CEMZ. We further show that such gravitons can undergo a deflection through an angle smaller than $\pi$. These features indicate that certain polarisations of the graviton experience a *repulsive* gravitational interaction at distances comparable to the scale set by the coupling constant. This is confirmed by the analysis of the "effective potential" which determines the graviton trajectories. Close to the event horizon, where the impact parameter is small, the gravitons experience an attractive gravitational interaction and hence the deflection angle is larger than $\pi$.

If the net deflection angle is non-vanishing then we can only have a time advance if the size of the cavity is not too large. If the graviton suffers a net deflection, then its path on the curved manifold will necessarily be longer than the corresponding straight line path in flat spacetime.[23] This results in a *positive* contribution to the time delay which grows with the radius of the cavity. On the other hand the negative contribution to the time delay only comes from a bounded region (whose size is set by the Gauss–Bonnet coupling). This means that for large enough radius of the cavity, the positive contribution will overcome the negative one, hence resulting in a positive time delay. However, since for different values of the impact parameter the graviton can experience deflections larger or smaller than $\pi$, by continuity there must be choices of the impact parameter for which the graviton suffers no net deflection. There are in fact two possible choices of impact parameter for which this is the case and the time delay remains finite in the infinite cavity radius limit. For the larger

[22] The large symmetry of the AS solution implies that the time delay should be unambiguous in this case too.

[23] CEMZ evaded this effect by considering a graviton propagating between two AS shocks, resulting in zero deflection.

of these values, the time delay is positive. However, for the smaller one it can be negative. Hence we find an example of a graviton incident from, and returning to, infinity in less time than it would in flat space.

According to CEMZ, one of the main implications of the negative time delay is that it would make it possible to violate causality, rendering EGB unviable as a classical theory. They employ an argument asserting that superluminal propagation in a Lorentz covariant theory can be exploited to construct "time machines", i.e., spacetimes containing *closed causal curves* (where "causal" is defined with respect to the causal structure induced by the equations of motion, cf. Chapter 2.4). Arguments of this sort have been applied to various non-gravitational field theories in Ref. [83]. They have however been criticised by Geroch [84], who notes that the existence of a causally pathological solution is not enough to reject the theory (see also Ref. [85]). After all, even GR admits causally pathological solutions. A simple example of this is given by Minkowski spacetime with a periodic time direction. However, nobody would discard GR because of this. The reason why such pathological solutions should not be considered problematic for the theory is that they cannot be formed dynamically, i.e. they do not arise as the Cauchy evolution of some "good" initial data. In other words, they are not "physical" solutions.

We will argue that, in fact, the "time machine" constructions of Adams et al. [83] and CEMZ [41] do not arise as the Cauchy development of a legitimate initial data set. The reason behind this is that the initial data surface will fail to be everywhere spacelike. Hence it is not a well-posed problem to specify a solution in terms of such data: either no solution exists, or it is infinitely fine-tuned.

This Chapter is organised as follows. In Section 4.1 we review the static, spherically symmetric, black hole solutions of EGB and derive a "speed limit" for small black holes. In Section 4.2 we will investigate the Shapiro time delay and deflection of gravitons propagating in a static, spherically symmetric, black holes solution of EGB. In Section 4.3 we discuss the "time machine" constructions of Ref. [83] and CEMZ. The details of our perturbative calculation of the time delay and deflection angle are treated in Appendix 4.A.

## 4.1 Spherically symmetric EGB black holes

The equation of motion of Einstein–Gauss–Bonnet (EGB) theory is obtained by varying the action

$$S = \frac{1}{16\pi}\int d^d x\sqrt{-g}\left[R + \lambda_{\rm GB}(R^2 - 4R_{ab}R^{ab} + R_{abcd}R^{abcd})\right] \tag{4.1.1}$$

where $\lambda_{GB}$ is the Gauss–Bonnet (GB) coupling constant and we have set $G = 1$.[24] This theory admits static, spherically symmetric solutions with metric [86]:

$$g = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 g_{\mathbb{S}^{d-2}} \tag{4.1.2}$$

where $g_{\mathbb{S}^{d-2}}$ is the standard round metric on the unit $(d-2)$-sphere, $\mathbb{S}^{d-2}$, and

$$f(r) = 1 + \frac{r^2}{\alpha}(1 - q(r)), \tag{4.1.3}$$

$$q(r) = \sqrt{1 + \frac{2\alpha\mu}{r^{d-1}}}, \tag{4.1.4}$$

$$\alpha = 2(d-3)(d-4)\lambda_{\rm GB}. \tag{4.1.5}$$

We assume $\alpha > 0$ since for $\alpha < 0$ the metric has a naked singularity [86].[25] The parameter $\mu$ is proportional to the ADM mass $M$:

$$M = \frac{(d-2)\text{Vol}(\mathbb{S}^{d-2})}{16\pi}\mu. \tag{4.1.6}$$

The event horizon is a hypersurface $r = r_H$,[26] where $r_H$ is the largest root of

$$\mu = r_H^{d-5}\left(r_H^2 + \frac{\alpha}{2}\right). \tag{4.1.7}$$

The coupling $\alpha$ has dimensions of length squared so EGB theory has a length scale $\sqrt{\alpha}$. We will say that a black hole is "small" if $r_H \ll \sqrt{\alpha}$ and "large" if $r_H \gg \sqrt{\alpha}$. Equivalently, for $d > 5$, a black hole is small if

$$\mu \ll \alpha^{(d-3)/2} \tag{4.1.8}$$

and large if $\mu \gg \alpha^{(d-3)/2}$. Note that black holes with arbitrarily small mass do not exist for $d = 5$ because there is a mass gap (however $r_H$ can be arbitrarily small): $\mu > \alpha/2$. The function $q$ varies over a length scale

$$L \equiv (\mu\alpha)^{\frac{1}{d-1}}. \tag{4.1.9}$$

For $r \gg L$ we have

$$f \approx 1 - \frac{\mu}{r^{d-3}} \tag{4.1.10}$$

i.e., the solution reduces to the higher-dimensional Schwarzschild solution. For a large black hole $r_H \gg L$ so this approximation is valid everywhere outside the horizon.

[24] In the notation of Section 3.2, this theory corresponds to $\lambda_{\rm GB} = -8k_2$, $k_p = 0$ for $p > 2$.

[25] A priori, the sign in front of $q$ in the expression for $f$ is arbitrary. We choose the negative branch, corresponding to asymptotically flat solutions.

[26] Since some modes of gravitational perturbations can travel faster than light, it could be possible, a priori, that such perturbations could escape the black hole region (as defined by the causal structure given by the physical metric). However it was shown in [87, 70] that a Killing horizon is always a characteristic hypersurface for all graviton polarizations, excluding this possibility when the event horizon is a Killing horizon, as is the case here.

### 4.1.1 Effective metrics and bicharacteristic curves

Characteristic hypersurfaces of the above solution were determined in Ref. [70]. The symmetries of the solution imply that the characteristic polynomial factorizes into a product of quadratic factors, each associated to an "effective metric" as discussed above. A hypersurface is characteristic iff it is null w.r.t. one of these effective metrics. The explicit form of the effective metrics was determined by considering linear perturbations of such solutions. Such perturbations can be classified into scalar ($S$), vector ($V$) and tensor ($T$) types w.r.t. the spherical symmetry. For each type, one can obtain a "master equation" [88, 89] and from these one can read off the effective metric for that type. Hence the characteristic polynomial factorizes as

$$Q(p,\xi) = (G_S^{ab}\xi_a\xi_b)^{p_S}(G_V^{cd}\xi_c\xi_d)^{p_V}(G_T^{ef}\xi_e\xi_f)^{p_T} \tag{4.1.11}$$

where $p_S, p_V, p_T$ denote the number of degrees of freedom of each type of modes. Viewing $G_S^{ab}$, etc., as inverse metrics, the corresponding metrics are given by

$$G_A = -f(r)dt^2 + f(r)^{-1}dr^2 + \frac{r^2}{c_A(r)}g_{\mathbb{S}^{d-2}} \tag{4.1.12}$$

for certain smooth functions $c_A(r)$ given by [70]

$$c_S(r) = 3\left(1-\frac{1}{d-2}\right)\mathcal{A}(r) + \left(1-\frac{3}{d-2}\right)\frac{1}{\mathcal{A}(r)} - 3\left(1-\frac{2}{d-2}\right), \tag{4.1.13}$$

$$c_V(r) = \mathcal{A}(r), \tag{4.1.14}$$

$$c_T(r) = -\left(1+\frac{1}{d-4}\right)\mathcal{A}(r) - \left(1-\frac{1}{d-4}\right)\frac{1}{\mathcal{A}(r)} + 3. \tag{4.1.15}$$

where

$$\mathcal{A}(r) = q(r)^{-2}\left(\frac{1}{2}+\frac{1}{d-3}\right) + \left(\frac{1}{2}-\frac{1}{d-3}\right), \tag{4.1.16}$$

It is convenient to take the index $A \in \{0, S, V, T\}$ where 0 refers to the physical metric, i.e.,

$$c_0(r) \equiv 1. \tag{4.1.17}$$

For a large black hole, the functions $c_A(r)$ are positive everywhere outside the horizon. This ensures that the effective metrics have Lorentzian signature, and their null cones form a nested set, with the outermost cone (in the tangent space) corresponding to the effective metric with the largest value of $c_A(r)$. (The physical metric $G_0$ is included in this nested set.) This ensures the hyperbolicity of the theory, and causality is determined by this outermost null cone. If $c_A > 1$ then the associated modes propagate faster than light [70].

For $d = 5, 6$, for small enough $r_H$, it turns out that one of the $c_A(r)$ vanishes at some value $r = r_* > r_H$ and becomes negative for $r < r_*$ [70]. The corresponding inverse effective metric $G_A^{ab}$ is smooth at $r = r_*$ but becomes degenerate there. For $r < r_*$ it has Lorentzian signature, but with the opposite overall sign. This implies that the theory is non-hyperbolic in such black hole spacetimes. Therefore small black holes are unphysical for $d = 5, 6$. This does not occur for EGB theory with $d \geq 7$.[27]

Note that the function $c_A$ are determined entirely by the length scale $L$ defined by (4.1.9). If $r \gg L$ then

$$c_A(r) = 1 + 2\beta_A \left(\frac{L}{r}\right)^{d-1} + \mathcal{O}\left(\left(\frac{L}{r}\right)^{2(d-1)}\right) \tag{4.1.18}$$

where the constants $\beta_A$ are given by

$$\beta_S = -\frac{(d-1)}{(d-3)}, \quad \beta_V = -\frac{1}{2}\frac{(d-1)}{(d-3)}, \quad \beta_T = \frac{(d-1)}{(d-3)(d-4)}, \quad \beta_0 = 0. \tag{4.1.19}$$

We see that $c_S, c_V < 1$ at large $r$ so the scalar and vector polarizations propagate slower than light in this region. However $c_T > 1$ at large $r$ so tensor polarizations propagate faster than light at large $r$. Hence causality at large $r$ is determined by the effective metric for the tensor modes.

We will now prove that $c_S < 1$ and $c_V < 1$ *everywhere*. Since $q(r)^{-2}$ is monotonically increasing we see that also $\mathcal{A}(r)$ is monotonically increasing. We also have $\mathcal{A}(\infty) = 1$. It follows that $\mathcal{A}(r) < 1$ hence $c_V(r) < 1$. Now we look at $c_S(r)$. We have:

$$c_S'(r) = \left[3\left(1 - \frac{1}{d-2}\right) - \left(1 - \frac{3}{d-2}\right)\mathcal{A}(r)^{-2}\right]\mathcal{A}'(r). \tag{4.1.20}$$

Since $\mathcal{A}(r)$ is monotonically increasing, the sign is determined by the terms in parentheses. For $d = 5$ this is constant and positive, hence $c_S' > 0$ so $c_S(r) < c_S(\infty) = 1$. For $d \neq 5$, the expression in parentheses is negative at small $r$ and positive at large $r$. Hence, starting from $r = 0$, $c_S(r)$ decreases to a minimum and then increases monotonically with $r$. Hence $c_S(r) < \max\{c_S(\infty), c_S(0)\} = c_S(\infty) = 1$.

The same argument allows us to determine an upper bound for $c_T$. We have:

$$c_T'(r) = -\left[\left(1 + \frac{1}{d-4}\right) - \left(1 - \frac{1}{d-4}\right)\mathcal{A}(r)^{-2}\right]\mathcal{A}'(r). \tag{4.1.21}$$

If $d = 5$ then the expression in square brackets is constant and positive so we see that $c_T$ is monotonically decreasing hence

$$c_T(r) < c_T(0) = 3 \qquad (d = 5). \tag{4.1.22}$$

[27]However, it does occur for more general Lovelock theories with $d \geq 7$. Generically, it occurs when the equation of motion includes the highest order Lovelock term [70].

If $d > 5$ then $c_T$ has a maximum at $r_0$ where

$$\mathcal{A}(r_0) = \sqrt{\frac{1 - 1/(d-4)}{1 + 1/(d-4)}} \tag{4.1.23}$$

and hence

$$c_T(r) < c_T(r_0) = 3 - 2\sqrt{1 - \frac{1}{(d-4)^2}} \tag{4.1.24}$$

Note that the RHS is greater than 1.

### 4.1.2 Speed limit for small black holes

We will now consider the effect of boosting one of these black holes. To construct initial data describing a boosted black hole, we can consider the data induced on a boosted hypersurface in the black hole spacetime. Such a hypersurface is spacelike w.r.t. the metric for any boost velocity $v$ such that $|v| < 1$. However, since the null cone for the tensor modes can lie outside the null cone of the physical metric, it is possible that, for $|v|$ close to 1, the hypersurface may fail to be everywhere spacelike w.r.t. the tensor effective metric. This implies that it will fail to be spacelike in the sense defined in the Introduction and hence it would not be a valid initial data surface. We will now show that this is indeed what happens.

First we introduce an "isotropic" radial coordinate $\tilde{r}$ defined by

$$\frac{d \log \tilde{r}}{dr} = \frac{1}{r\sqrt{f}} \tag{4.1.25}$$

so that the physical metric is

$$g = -f dt^2 + H\left(d\tilde{r}^2 + \tilde{r}^2 g_{\mathbb{S}^{d-2}}\right) \tag{4.1.26}$$

where

$$H = \frac{r^2}{\tilde{r}^2}. \tag{4.1.27}$$

For $r \gg L$ we can use the approximation (4.1.10) to obtain

$$\tilde{r} \approx r\left(1 - \frac{\mu}{2(d-3)r^{d-3}}\right) \tag{4.1.28}$$

and hence

$$f \approx 1 - \frac{\mu}{\tilde{r}^{d-3}} \qquad H \approx 1 + \frac{\mu}{(d-3)\tilde{r}^{d-3}}. \tag{4.1.29}$$

To construct initial data describing a boosted black hole we convert to Cartesian coordinates $x^i$ so that $x^1 = \tilde{r}\cos\theta_1$, etc., (where $\theta_1, \theta_2, \ldots$ are the angles on $S^{d-2}$) and then perform the Lorentz transformation

$$x^1 = \gamma(x^{1'} - vt'), \quad t = \gamma(t' - vx^{1'}), \qquad \gamma = (1-v^2)^{-1/2}. \tag{4.1.30}$$

We now consider the data induced on a surface of constant $t'$. By inverting the Lorentz transformation, we see that this is the same as the data induced on a surface of constant $t + vx^1$, i.e., a surface of constant $t + v\tilde{r}\cos\theta_1$. Let $\Sigma$ be such a surface. Define a 1-form $\xi$ normal to $\Sigma$:

$$\xi = dt + v\cos\theta_1 d\tilde{r} - v\tilde{r}\sin\theta_1 d\theta_1. \tag{4.1.31}$$

We want to take the data induced on $\Sigma$ as initial data describing a boosted black hole. To do this, we must check that $\Sigma$ is spacelike in the sense defined in the Introduction. This is equivalent to $\Sigma$ being spacelike w.r.t. all of the effective metrics, i.e., $\xi$ must be timelike w.r.t. to all of the effective metrics. To investigate whether this is this case, consider the norm of $\xi$ w.r.t. $G_A$ at $\theta_1 = \pi/2$:

$$G_A^{\mu\nu}\xi_\mu\xi_\nu|_{\theta_1=\pi/2} = -f^{-1} + v^2 c_A \frac{\tilde{r}^2}{r^2}. \tag{4.1.32}$$

Now assume that the black hole is small and consider the region $L \ll r \ll \sqrt{\alpha}$ where $\mu/r^{d-3} \ll L^{d-1}/r^{d-1}$. Using our expansion for $c_A$ we then obtain

$$G_A^{\mu\nu}\xi_\mu\xi_\nu|_{\theta_1=\pi/2} \approx -(1-v^2) + 2\beta_A v^2 \left(\frac{L}{r}\right)^{d-1}. \tag{4.1.33}$$

For the scalar, vector and physical metrics we have $\beta_A \leq 0$ so the RHS is always negative. However, for the tensor effective metric we have $\beta_T > 0$ and hence if $v$ is too close to 1 then the second term above, although small, will overwhelm the first and the RHS will be positive, i.e., $\xi$ will be spacelike w.r.t. $G_T$ and hence $\Sigma$ will not be everywhere spacelike.

As discussed in the Introduction, it is not a well-posed problem to evolve initial data if $\Sigma$ is not spacelike. Of course, we know that this particular data on $\Sigma$ *can* be evolved - the resulting solution is just the black hole solution described above. However, the lack of well-posedness implies that this procedure is infinitely fine-tuned: if we make a generic (smooth) perturbation to the initial data on $\Sigma$ (for $v$ very close to 1) then it will not be possible to evolve the perturbed data either forwards or backwards in time. Hence there is a speed limit for small black holes: they cannot be boosted to velocities arbitrarily close to the speed of light.

One might criticise this argument on the grounds that there is no unique way to boost a black hole. One could consider a different surface which is asymptotic to $\Sigma$ but differs in the region $L \ll r \ll \sqrt{\alpha}$ in which $\Sigma$ can fail to be spacelike. However, note that in our argument, the physical metric is actually flat to the level of approximation used because we neglected terms of order $\mu/r^{d-3}$. The boost used above is a symmetry of this flat metric. Therefore our surface $\Sigma$ conforms to the usual idea of a boosted hypersurface in the region relevant for the above argument.

Of course the effective metrics are not flat to this level of approximation because they include larger terms, of order $\alpha\mu/r^{d-1} = (L/r)^{d-1}$.

It is interesting to determine the critical value of $v$ below which $\Sigma$ is spacelike. For $\Sigma$ to be spacelike w.r.t. the tensor metric we need $\xi$ to be timelike everywhere, which is true iff[28]

$$v^2 < v_{\rm max}^2 \equiv \min \frac{r^2}{\tilde{r}^2 f c_T} \tag{4.1.34}$$

where the minimum is taken over, say, $r > r_H$. For a small black hole, the minimum is achieved for $L \sim r \ll \sqrt{\alpha}$, for which $f \approx 1$ and $\tilde{r} \approx r$ hence

$$v_{\rm max}^2 \approx \frac{1}{\max c_T} = \frac{1}{3 - 2\sqrt{1 - \frac{1}{(d-4)^2}}} < 1. \tag{4.1.35}$$

This is the speed limit for an arbitrarily small black hole. More generally, $v_{\rm max}$ will depend on the mass of the black hole, with $v_{\rm max} \to 1$ for a large black hole.

It is natural to ask what would happen if one attempted to accelerate a small black hole to a speed greater than $v_{\rm max}$. As emphasized in Ref. [84], one would have to specify the details of how one would attempt to achieve this acceleration using only the fields present in the theory. Perhaps one could set up initial data consisting of several black holes in the hope that a "gravitational slingshot" effect could be used to accelerate a small black hole to a speed greater than $v_{\rm max}$. However, as we will see in more detail below, the gravitational interaction associated with small black holes in EGB is very different from GR so there is no reason to believe that a small black hole in such a system would behave in the same was as it would in GR. Whatever the system does, it will not result in a small black hole moving with a speed arbitrarily close to the speed of light at some later time. This is because an "instant of time" corresponds to a spacelike hypersurface and the argument above excludes the possibility of a small black hole moving arbitrarily close to the speed of light on such a surface.

### 4.1.3 Graviton trajectories

As discussed above, characteristic hypersurfaces are generated by bicharacteristic curves and, in the present case, these are simply the null geodesics of the effective metrics. Hence, in the geometric optics approximation, the worldlines of high-frequency gravitons are null geodesics of the effective metrics. We will need to determine these geodesics in order to calculate the time delay.

[28] We used $\theta_1 = \pi/2$ to derive this. It is not hard to show that other values of $\theta_1$ gives less stringent constraints.

Consider a null geodesic of $G_{Aab}$. Introducing polar coordinates $(\theta_1, \theta_2, \ldots, \theta_{d-3}, \phi)$ on $S^{d-2}$, spherical symmetry allows us to assume that the geodesic is confined to the equatorial plane $\theta_1 = \cdots = \theta_{d-3} = \pi/2$. Associated to the Killing fields $\partial/\partial t$ and $\partial/\partial\phi$ are the conserved quantities

$$E = f(r)\dot{t} \qquad J = \frac{r^2}{c_A(r)}\dot{\phi} \tag{4.1.36}$$

where a dot denotes a derivative w.r.t. an affine parameter $\lambda$. $E$ and $J$ are not physical because they depend on the choice of affine parameter. However their ratio is independent of this choice:

$$b = \frac{J}{E} \tag{4.1.37}$$

and this is the impact parameter of the geodesic. The null condition gives

$$\frac{1}{2}\dot{r}^2 + J^2 V_{\text{eff}}^{(A)}(r) = \frac{1}{2}E^2, \tag{4.1.38}$$

where the effective potential is given by

$$V_{\text{eff}}^{(A)}(r) = \frac{f(r)c_A(r)}{2r^2}. \tag{4.1.39}$$

Plots of the effective potential for some different cases are given in Ref. [70] and also in Figure 4.1.

The effective potentials exhibit a local maximum corresponding to an unstable graviton orbit analogous to the photon sphere in GR. Hence in EGB there is a distinct "graviton sphere" for each graviton polarisation. We will refer to these as the "scalar sphere", "vector sphere" and "tensor sphere". In some cases, it turns out that there are two local maxima and a local minimum of the effective potential and hence three graviton spheres. If $V_{\max}^{(A)}$ denotes the maximum of the effective potential then a graviton incident from large distance will be absorbed by the black hole if

$$b^2 < \frac{1}{2V_{\max}^{(A)}} \qquad \Rightarrow \qquad \text{absorption.} \tag{4.1.40}$$

We will consider only gravitons with larger impact parameter, which are scattered by the black hole.

For $r \gg L$, equations (4.1.10) and (4.1.18) imply that the effective potentials have the expansion

$$V_{\text{eff}}^{(A)}(r) = \frac{1}{2r^2} - \frac{\mu}{2r^{d-1}} + \beta_A \frac{\alpha\mu}{r^{d+1}} + \ldots \tag{4.1.41}$$

The first two terms are familiar from GR: the first is a centrifugal barrier and the second is responsible for the deflection of light rays and the time delay of photons. The third term arises from the Gauss–Bonnet interaction. For the effects of this

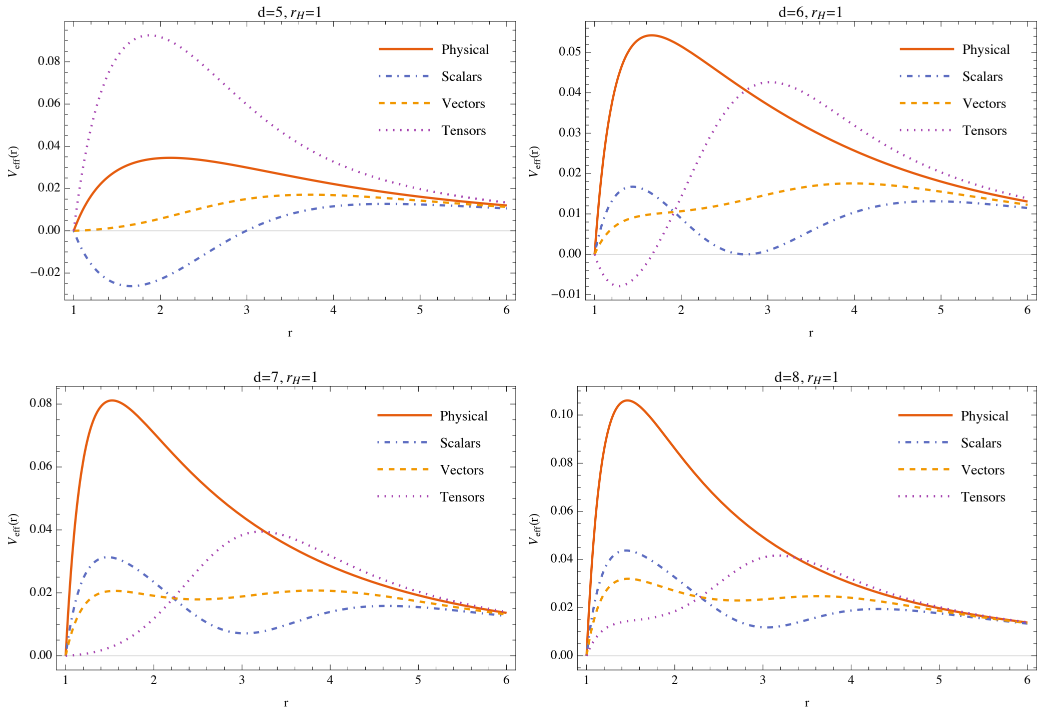


Figure 4.1: Effective potentials for a black hole with $r_H = 1$ in $d = 5, 6, 7, 8$ dimensions. We fix the Gauss–Bonnet coupling $\lambda_{\rm GB} = 2$. The red curve corresponds to the effective potential for photons, i.e., null geodesics of the physical metric ($c = 1$). Superluminal propagation ($c_A > 1$) corresponds to an effective potential which is larger than that for photons. This happens only for tensor polarizations. The violation of hyperbolicity is associated with the region in which one of the effective potentials becomes negative. This happens near the horizon for small black holes in five and six dimensions.

term to be non-negligible compared to the second term we need $r \lesssim \sqrt{\alpha}$. Since we have assumed $r \gg L$ this requires $L \ll \sqrt{\alpha}$, which implies (4.1.8), i.e., the black hole has to be small compared to the GB scale for this term to be important.[29] Notice that this term is negative for vectors and scalars but positive for tensors. Hence, for a small EGB black hole, tensor-polarized gravitons experience a new *repulsive* interaction for $L \ll r \lesssim \sqrt{\alpha}$. It is this repulsive interaction that allows for the possibility of a time advance.

For a small black hole, the effective potential also simplifies in the region $L \sim r \ll \sqrt{\alpha}$. In this region we can approximate $f \approx 1$ and hence

$$V^{(A)}(r) \approx \frac{c_A(r)}{2r^2} \tag{4.1.42}$$

which depends only on the length scale $L$. For tensor modes, the RHS typically has

[29]This is not possible for $d = 5, 6$ because of the failure of hyperbolicity for small black holes with $d = 5, 6$.

a maximum at some value $r \sim L$ so we deduce that the "tensor sphere" has $r \sim L$. Since $c_A \sim 1$, we deduce from (4.1.40) that tensor-polarized gravitons which scatter from the black hole (rather than being absorbed) must have $b \gtrsim L$.[30] Note also that the presence of this maximum implies that the interaction between tensor-polarized gravitons and the black hole must become attractive again for $r \lesssim L$.

## 4.2 Time delay and time advance

### 4.2.1 Photon time delay in GR

As discussed in the Introduction, there is no gauge-invariant definition of the Shapiro time delay applicable to a large class of spacetimes [81]. However, for a static, spherically symmetric spacetime, there is an unambiguous definition of this quantity [82]. The idea is to compare the time it takes a photon to travel between two points of a spherical cavity with the corresponding time in Minkowski spacetime. In more detail, one can introduce coordinates $(t, r, \theta, \phi)$ (in 4-dimensional GR) so that the metric takes the form

$$ds^2 = -A(r)dt^2 + B(r)dr^2 + r^2 d\Omega^2 \tag{4.2.1}$$

with $A, B > 0$. Now one can consider a photon trajectory which starts and ends at $r = R$, with $r \leq R$ along the trajectory. Using the spherical symmetry, we can assume that the motion is confined to the equatorial plane $\theta = \pi/2$. Assume it starts at $t = t_0$, $\phi = \phi_0$ and ends at $t = t_0 + \Delta t$, $\phi = \phi_0 + \Delta\phi$ (Fig. 4.2, left). The coordinate time to traverse the cavity is $\Delta t$. The proper time (according to a cavity observer) is

$$\Delta\tau = \sqrt{A(R)}\Delta t. \tag{4.2.2}$$

One can compare this with corresponding quantity in Minkowski spacetime where the trajectory is simply a straight line traversing the cavity (Fig. 4.2, right), which takes proper times

$$\Delta\tau_{\text{Mink}} = 2R\sin\left(\frac{\Delta\phi}{2}\right). \tag{4.2.3}$$

Hence the time delay can be defined as

$$D \equiv \Delta\tau - \Delta\tau_{\text{Mink}}. \tag{4.2.4}$$

A simple argument (based on [82]) shows that, in GR, it is impossible to have $D < 0$ for a large class of spacetimes of the above form. Consider spacetimes for

[30] Hence the absorption cross-section for (high frequency) tensor-polarized gravitons by a small black hole scales as $L^{d-2}$ rather than $r_H^{d-2}$.

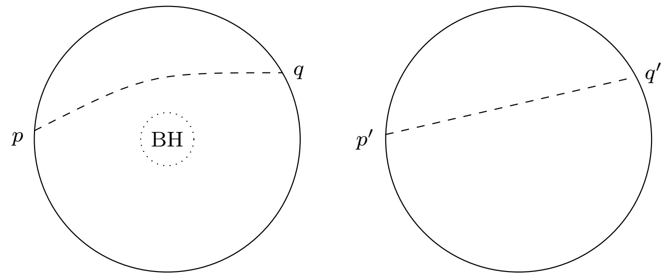


Figure 4.2: Schematic set-up of the problem: the dashed line represents the trajectory of the graviton. The solid circle represents the spherical cavity of area-radius $R$ in the curved spacetime (*left*) and in Minkowski (*right*). The endpoints of the path in the curved spacetime, $p$ and $q$, are identified with $p'$ and $q'$ in Minkowksi, using the spherical symmetry of the problem.

which $A'(r) \geq 0$ and $B(r) \geq 1$. Examples of such spacetimes include the (positive mass) Schwarzschild solution and perfect fluid stars with positive energy density and pressure. Now let $\lambda \in [0,1]$ be an arbitrary parameter along the photon trajectory (which has $r(\lambda) \leq R$). Since the photon trajectory is null we have

$$\begin{aligned}\Delta t = \int_0^1 \sqrt{A^{-1}B\dot{r}^2 + A^{-1}r^2\dot{\phi}^2}\, d\lambda &\geq \frac{1}{\sqrt{A(R)}}\int_0^1 \sqrt{\dot{r}^2 + r^2\dot{\phi}^2}\, d\lambda \\ &\geq \frac{2R}{\sqrt{A(R)}}\sin\left(\frac{\Delta\phi}{2}\right)\end{aligned} \tag{4.2.5}$$

where the first inequality follows from $B(r) \geq 1$ and $A'(r) \geq 0$, so $A(r) \leq A(R)$ for $r \leq R$. The second inequality follows from the fact that the distance in Euclidean space is minimized by a straight line. It follows immediately that $D \geq 0$.[31]

### 4.2.2 Time delay in EGB

We can now calculate the Shapiro time delay for gravitons propagating across a spherical cavity in the geometry (4.1.2). The cavity is taken to be the surface $r = R$.[32] Consider a graviton worldline parametrised by $\lambda \in [0,1]$ that has $r \leq R$ and starts and ends at $r = R$ with $\phi(0) = 0$, $\phi(1) = \Delta\phi$. From the fact that this world line is null w.r.t. the relevant effective metric, the coordinate time $t$ taken for the graviton to traverse the cavity is

$$\Delta t = \int_0^1 \sqrt{\frac{\dot{r}^2}{f^2} + \frac{r^2}{f c_A}\dot{\phi}^2}\, d\lambda. \tag{4.2.6}$$

[31] If we normalize $t$ so that $A(r) \to 1$ as $r \to \infty$ then $A'(r) \geq 0$ implies $A(r) \leq 1$ everywhere so $\Delta\tau \leq \Delta t$. Hence the time delay defined using $\Delta t$ instead of $\Delta\tau$ also will be positive [82].

[32] Note that $R$ is the area-radius of the cavity w.r.t. the physical metric, not w.r.t. an effective metric.

One can show that $f$ is monotonically increasing[33] so $f(r) \leq f(R) \leq f(\infty) = 1$. Hence

$$\Delta t > \frac{1}{\sqrt{f(R)}} \int_0^1 \sqrt{\frac{\dot{r}^2}{f} + \frac{r^2}{c_A}\dot{\phi}^2}\, d\lambda > \frac{1}{\sqrt{f(R)}} \int_0^1 \sqrt{\dot{r}^2 + \frac{r^2}{c_A}\dot{\phi}^2}\, d\lambda. \tag{4.2.7}$$

We showed above that $c_S < 1$ and $c_V < 1$. This implies that for scalar or vector polarizations (or for photons, which have $c_0 = 1$) we have

$$\Delta t > \frac{1}{\sqrt{f(R)}} \int_0^1 \sqrt{\dot{r}^2 + r^2\dot{\phi}^2}\, d\lambda \geq \frac{2R}{\sqrt{f(R)}} \sin\left(\frac{\Delta\phi}{2}\right) \equiv \frac{\Delta\tau_{\text{Mink}}}{\sqrt{f(R)}} \tag{4.2.8}$$

where, as before, $\Delta\tau_{\text{Mink}}$ is the time it takes a photon (or graviton) in Minkowski spacetime to travel across the cavity between the same two points. Converting to proper time we therefore have

$$\Delta\tau = \sqrt{f(R)}\Delta t > \Delta\tau_{\text{Mink}} \tag{4.2.9}$$

so the time delay is always positive for scalar or vector polarized gravitons. More physically: gravitons with these polarizations travel slower than photons so, since photons experience a positive time delay, these gravitons must also experience positive time delay.

The story is different for tensor modes: as we have shown, $c_T(r)$ can be larger than one, so one cannot rule out the possibility of time *advance* (e.g. in $d = 5$ we would have $c_T(r) \in [1, 3]$ and thus $\Delta\tau \geq \Delta\tau_{\text{Mink}}/\sqrt{3}$). In the next subsection we will show that time advance is indeed possible, in agreement with [41].

### 4.2.3 Time advance in EGB: perturbative results

We will now show how one can achieve a negative time delay, i.e., a time advance, for gravitons of tensor polarizations in the space time of a small black hole. We will calculate the time delay explicitly. For completeness, we will also present results for the time delay for vector and scalar graviton polarizations, and also for photons. Consider a graviton trajectory, given by a null geodesic of the relevant effective metric. As before, we assume that this starts at $t = 0$, $r = R$, $\phi = 0$ and ends at $t = \Delta t$, $r = R$, $\phi = \Delta\phi$ with $r \leq R$ along the trajectory. Let $R_0$ be the minimum value of $r$ along the trajectory. As in GR, $R_0$ uniquely labels the trajectory. This is related to the impact parameter $b$ as:

$$b^2 = \frac{R_0^2}{f(R_0)c_A(R_0)}. \tag{4.2.10}$$

[33] Use $q' = -(d-1)(q^2-1)/(2rq)$ to show that $f' > 0$.

We can compute the proper time and the deflection angle from the geodesic equations:

$$\Delta\phi = 2b\int_{R_0}^{R} dr\ c_A(r)\left(r^2\sqrt{1-\frac{f(r)c_A(r)b^2}{r^2}}\right)^{-1}, \tag{4.2.11}$$

$$\Delta t = 2\int_{R_0}^{R} dr\ \left(f(r)\sqrt{1-\frac{f(r)c_A(r)b^2}{r^2}}\right)^{-1}. \tag{4.2.12}$$

As before, this includes results for photons (with $c_0 = 1$). Recall that $R_0$ must be larger than the radius of the *photon/graviton sphere* for the physical/effective metrics. Both the time delay and the deflection angle will diverge as $R_0$ approaches this value since the corresponding trajectories will orbit the black hole many times.

We will first calculate the above quantities for a graviton trajectory that has $R_0 \gg L$ and also $R_0 \gg \mu^{1/(d-3)}$, and hence $r \gg L$, $r \gg \mu^{1/(d-3)}$ along the trajectory.[34] Under these assumptions, $b$ and $R_0$ are related by

$$b^2 = R_0^2\left[1+\frac{\mu}{R_0^{d-3}}\left(1-\frac{2\alpha\beta_A}{R_0^2}\right)+\dots\right]. \tag{4.2.13}$$

We will assume also that the cavity radius is large: $R \gg R_0$. The perturbative calculation of the deflection angle is explained in Appendix 4.A. The result is

$$\Delta\phi \approx \pi - 2\frac{R_0}{R} + \frac{\mu}{R_0^{d-3}}\sqrt{\pi}(d-1)\left(1-\frac{2\alpha\beta_A}{R_0^2}\frac{(d-2)}{(d-1)}\right)\frac{\Gamma\left(\frac{d}{2}\right)}{2\Gamma\left(\frac{d+1}{2}\right)}. \tag{4.2.14}$$

This has a well-defined limit when the cavity radius is taken to infinity at fixed $R_0$. Using $b \approx R_0$ to write the result in terms of $b$ gives

$$\Delta\phi_\infty \equiv \lim_{R\to\infty}\Delta\phi \approx \pi + \frac{\mu}{b^{d-3}}\sqrt{\pi}(d-1)\left(1-\frac{2\alpha\beta_A}{b^2}\frac{(d-2)}{(d-1)}\right)\frac{\Gamma\left(\frac{d}{2}\right)}{2\Gamma\left(\frac{d+1}{2}\right)}. \tag{4.2.15}$$

This is analogous to the result for the deflection of light by the Schwarzschild solution in GR. Note that $\Delta\phi_\infty > \pi$ if $\beta_A \le 0$ so scalar and vector polarised gravitons, and photons, are always deflected towards the black hole. However, for tensor polarised gravitons, since $\beta_T > 0$, we see that $\Delta\phi_\infty < \pi$ when

$$b < \sqrt{\frac{2(d-2)\alpha}{(d-3)(d-4)}}. \tag{4.2.16}$$

This is consistent with our previous assumptions if, and only if, (4.1.8) holds, i.e., iff the black hole is small. Hence, for a small black hole, tensor-polarized gravitons with $b$ obeying (4.2.16) (and $b \approx R_0 \gg L$ from our previous assumptions) are deflected

[34]For a large black hole, i.e., $r_H \gg \sqrt{\alpha}$, these conditions reduce to $R_0 \gg r_H$. For a small black hole, i.e., $r_H \ll \sqrt{\alpha}$, they reduce to $R_0 \gg r_H^{\frac{d-5}{d-3}}\alpha^{\frac{1}{d-3}}$.

*away* from the black hole. This is precisely because such gravitons experience the repulsive short-distance interaction in (4.1.41) that we discussed above.

Using the same approximations as above we find that the time it takes a graviton to cross the cavity is

$$\Delta t = 2R + \frac{\mu}{R_0^{d-4}} \left[ \sqrt{\pi} \left( 1 - \frac{2\alpha\beta_A}{R_0^2} \frac{(d-4)}{(d-3)} \right) \frac{(d-1)(d-3)}{(d-4)} \frac{\Gamma\left(\frac{d}{2}\right)}{2\Gamma\left(\frac{d+1}{2}\right)} \right] + \ldots \quad (4.2.17)$$

and $R \gg R_0$ implies that $\Delta\tau = \Delta t$ to this level of approximation. The corresponding time in flat spacetime, with deflection angle $\Delta\phi$ is given by (4.2.3), which can be written as

$$\Delta\tau_{\text{Mink}} = 2R\cos\left(\frac{\pi - \Delta\phi}{2}\right) = 2R\left[1 - \frac{1}{2}\left(\frac{\pi-\Delta\phi}{2}\right)^2 + \ldots\right] \quad (4.2.18)$$

so plugging in our perturbative result (4.2.14) gives

$$\Delta\tau_{\text{Mink}} = 2R + \frac{\mu}{R_0^{d-4}} \left[ \sqrt{\pi} \left( 1 - \frac{2\alpha\beta_A}{R_0^2} \frac{(d-2)}{(d-1)} \right) (d-1) \frac{\Gamma\left(\frac{d}{2}\right)}{2\Gamma\left(\frac{d+1}{2}\right)} \right] \quad (4.2.19)$$

$$+ \mathcal{O}\left( R \left( \frac{\mu}{R_0^{d-3}} + \frac{L^{d-1}}{R_0^{d-1}} \right)^2 \right). \quad (4.2.20)$$

Note that second order corrections to this result grow linearly with $R$. For these to be small compared to the terms that we have retained we need

$$\frac{R}{R_0} \ll \frac{R_0^{d-3}}{\mu}, \qquad \frac{R}{R_0} \ll \left(\frac{R_0}{L}\right)^{d-1} = \frac{R_0^{d-1}}{\alpha\mu} \quad (4.2.21)$$

i.e. the cavity is large ($R/R_0 \gg 1$) but not too large. If these assumptions are not satisfied then most of the trajectory is in a region where spacetime is almost flat and a large positive time delay (proportional to $R$) results simply because, in this flat region, there is a shorter (straight line) path available which remains far from the black hole.

Combining the above results gives the time delay as (using $R_0 \approx b$)

$$D = \Delta\tau - \Delta\tau_{\text{Mink}} \approx \frac{\mu}{b^{d-4}} \left[ \sqrt{\pi} \left( 1 - \frac{2\alpha\beta_A}{b^2} \frac{(d-4)}{(d-1)} \right) \frac{(d-1)}{(d-4)} \frac{\Gamma\left(\frac{d}{2}\right)}{2\Gamma\left(\frac{d+1}{2}\right)} \right]. \quad (4.2.22)$$

We see that scalar or vector polarised gravitons, or photons, suffer a positive time delay as expected. However, for tensor polarised gravitons, a negative time delay, i.e. a time *advance* results when

$$b < \sqrt{\frac{2\alpha}{(d-3)}}. \quad (4.2.23)$$

Since $b \approx R_0$, this is consistent with $R_0 \gg L$, as assumed above, only for a small black hole. For such a black hole, stated in terms of $b$, our assumptions in deriving (4.2.22) are

$$b \gg L, \qquad 1 \ll \frac{R}{b} \ll \frac{b^{d-3}}{\mu}, \left(\frac{b}{L}\right)^{d-1}. \tag{4.2.24}$$

Note that this overlaps the region for which $\Delta\phi_\infty < \pi$, i.e., the gravitons that experience a time advance are also deflected away from the black hole. Both effects arise from the repulsive term in the effective potential discussed above.

As discussed above, we need to impose an upper bound on $R/b$ to see the time advance since for very large $R$ there will be a large time delay, proportional to $R$, which occurs because the trajectory has undergone a deflection. However, for the special case of a trajectory which saturates (4.2.16), we have $\Delta\phi_\infty = \pi$, i.e., there is no net deflection (the effect of the short-distance repulsion is cancelled by the effect of the long-distance attraction). In this case, we no longer need to impose an upper bound on $R$: it is easy to see that the above derivation holds for arbitrarily large $R$. Hence the result (4.2.22) is valid for $R \to \infty$ in this special case. It is easy to see that, for this value of $b$, the expression (4.2.22) is positive, so this special trajectory experiences a time delay. Hence, in this special case, we have a gauge-invariant definition of the time delay for a graviton propagating in from infinity and returning to infinity.

### 4.2.4 Time advance in EGB: numerical results

The above perturbative calculation demonstrates that a time advance is possible for tensor-polarized gravitons propagating in the geometry of a small black hole for $d \geq 7$. However, several questions remain. As discussed above, small black holes are unphysical for $d = 5, 6$. So for $d = 5, 6$ we will have to study black holes which are not small in order to demonstrate that a time advance is possible. Furthermore, we would like to determine (for any $d$) how large the time advance can be. The perturbative result indicates that the time advance increases as the impact parameter decreases so we would like to consider $b$ as small as possible. The lower bound on the impact parameter is $b \sim L$ but the above calculation assumes $b \gg L$. Hence to determine the largest possible time advance we will need to use a different method. Finally, we discussed above the case of special trajectories which experience no net deflection, for which the time delay is finite as $R \to \infty$. We saw that, the resulting time delay is always positive when the perturbative calculation is valid. But what about trajectories with $b \sim L$? Could these exhibit zero net deflection? If so, can they exhibit a time advance?

To address the above questions, we will resort to numerical integration.[35] We will compute numerically both the deflection angle (4.2.11) and the time delay (4.2.12) for the tensor modes as functions of the impact parameter and plot the results for different parameters. In practice, we do this calculation by using $R_0$, the minimum value of $r$, to label the trajectory and determine $b$ from $R_0$ using (4.2.10).

We start by calculating the deflection angle for tensor-polarized gravitons by a small black hole with $d \geq 7$, in the limit of infinite cavity radius. We will compute the deflection angle $\Delta\phi_\infty$ as a function of the impact parameter $b$. For $d \geq 8$, complicated behaviour arises at small $b$ because the tensor effective potential for small black holes has a complicated form with a local minimum and two local maxima - see Figure 4.3. The local minimum corresponds to *stable* circular graviton orbits around the black hole. The maxima correspond to unstable circular graviton orbits. Hence there are three "tensor spheres". Only the local maxima are relevant for scattering of (high frequency) gravitons. As noted above, the deflection angle must diverge at impact parameters corresponding to these maxima and this can be seen in in Fig. 4.3. This figure also shows that $\Delta\phi_\infty > \pi$ when $R_0$ lies between the two maxima.

Figure 4.4 shows plots of $\Delta\phi_\infty$ for $d = 7, 8$ outside the outer tensor sphere. At large $b$, our perturbative results show that $\Delta\phi_\infty > \pi$ although this is not apparent from the plots because $\Delta\phi_\infty - \pi$ is very small. As $b$ decreases, our perturbative result shows that $\Delta\phi_\infty - \pi$ becomes negative, as seen in the plots. The plots show that $\Delta\phi_\infty - \pi$ decreases to a negative minimum and then increases, becoming positive as $b$ is decreased further. This lies outside the validity of the perturbative calculation. Note that there are two values of $b$ for which $\Delta\phi_\infty = \pi$, i.e., for which there is no net deflection. The larger of these, with $b \gg L$ is encompassed by our perturbative approximation, and we showed above that it gives a trajectory with a positive time delay in the infinite cavity limit. However, we will show that the smaller value of $b \sim L$ can give a time advance in the infinite cavity limit.

Figure 4.5 plots the time delay for different values of the cavity radius $R$ with $d = 7, 8$.[36] At large $b$, the time delay is small and positive, but becomes negative as $b$ is decreased, i.e. there is a time advance as predicted by our perturbative calculation. The size of the time advance increases as $b$ is decreased further but at small enough $b$, the time delay becomes positive again. As expected, increasing $R$ tends to increase the time delay. The only trajectories for which this does not happen correspond to the two special values of $b$ for which the trajectory does not undergo a net deflection.

[35] We perform the numerical integration with `Wolfram Mathematica` using a `GlobalAdaptive` method.

[36] For $d = 8$ we only show results for $R_0$ outside the outer tensor sphere. The time delay is positive for $R_0$ between the two maxima of the effective potential.

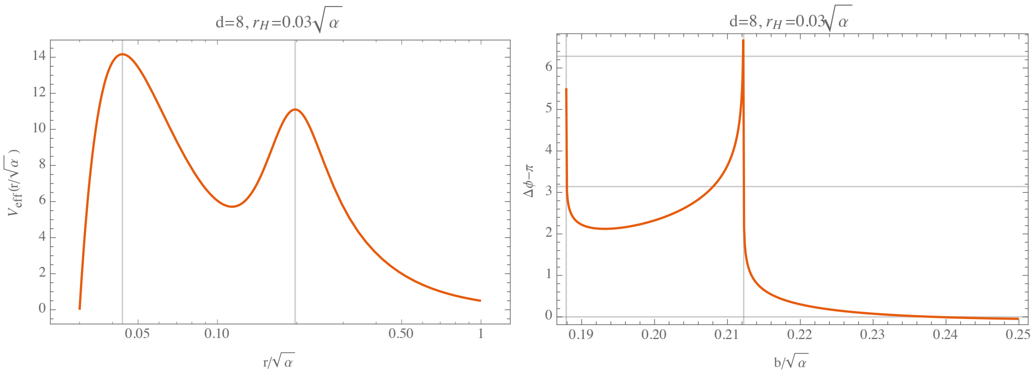


Figure 4.3: Effective potential (*left*) and deflection angle (*right*) for tensor polarized gravitons scattered by a small black hole in $d = 8$. We set $r_H = 0.03\sqrt{\alpha}$, which gives $\mu \approx 1.4 \times 10^{-5}\alpha^{5/2}$ and $L \approx 0.2\sqrt{\alpha}$.

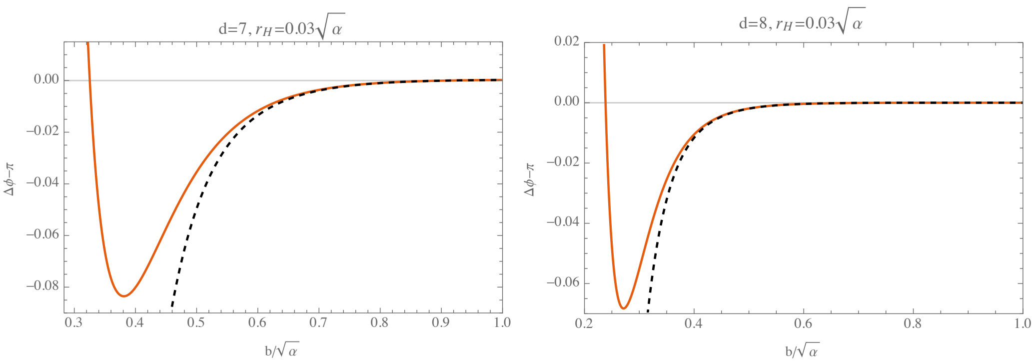


Figure 4.4: Deflection angle for small black holes in $d = 7, 8$. We set $r_H = 0.03\sqrt{\alpha}$ which gives $\mu \approx 4.5 \times 10^{-4}\alpha^2$, $L \approx 0.28\sqrt{\alpha}$ in $d = 7$, and $\mu \approx 1.4 \times 10^{-5}\alpha^{5/2}$, $L \approx 0.2\sqrt{\alpha}$ in $d = 8$. The dashed line represents the perturbative approximation (4.2.15).

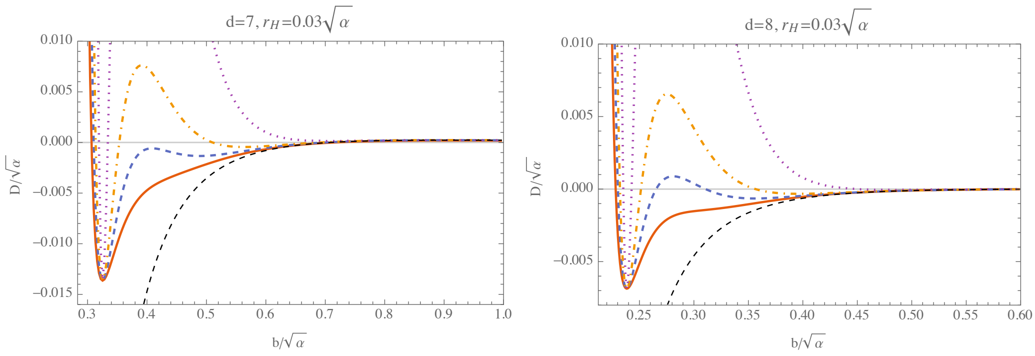


Figure 4.5: Time delay for small black holes in $d = 7, 8$. We set $r_H = 0.03\sqrt{\alpha}$ as before. The solid, dashed, dot–dashed and dotted lines correspond to $R = 2.5\sqrt{\alpha}$, $R = 5\sqrt{\alpha}$, $R = 10\sqrt{\alpha}$ and $R = 50\sqrt{\alpha}$ respectively. The black dashed line corresponds to the perturbative approximation (4.2.22).

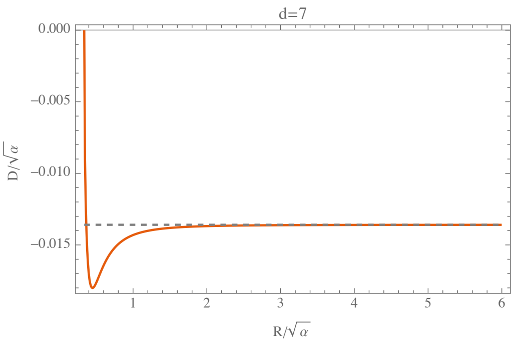


Figure 4.6: Time delay for fixed $b$ (such that $\Delta\phi_\infty = \pi$) expressed as a function of the radius of the cavity for a small black hole in $d = 7$. We set $r_H = 0.03\sqrt{\alpha}$ as before. The dashed line corresponds to the limit $R \to \infty$.

Hence as $R$ is increased, the minimum in these plots, corresponding to the largest time advance, becomes more and more localized around the smaller value of $b$ for which $\Delta\phi_\infty = \pi$.

Figure 4.6 shows how the maximum time advance, corresponding to the minimum in Fig. 4.5, behaves as $R$ is increased. As $R \to \infty$ we see that the (negative) time delay converges to a finite limit $D_\infty$, as expected for a trajectory with zero net deflection. We would like to understand what scale determines the amplitude of $D_\infty$. The obvious guess is the scale $L$ and this turns out to be correct. In Figure 4.7 we plot $|D_\infty|$ against the mass parameter $\mu$ for small black holes in $d = 7$. From the plot we deduce that the relation should be a power law: $|D_\infty| \sim \mu^\kappa$ (in units $\alpha = 1$). By estimating the value of $\kappa$ in different dimensions (Figure 4.7) we obtain numerically $\kappa \approx \frac{1}{d-1}$. Recalling that $L \sim \mu^{1/(d-1)}$ (since $\alpha = 1$), we have found:

$$|D_\infty| \sim L. \tag{4.2.25}$$

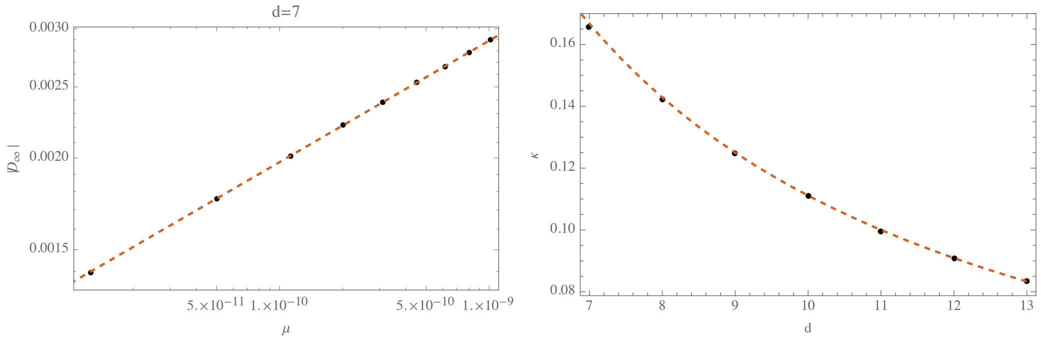


Figure 4.7: (*Left*) Log-log plot of the absolute value of the time advance for an undeflected geodesic in $d = 7$ in the limit $R \to \infty$ against the mass parameter $\mu$ (units $\alpha = 1$). The dashed line is given by $a\mu^\kappa$ with $a \approx 0.09$ and $\kappa \approx 0.165$. (*Right*) Plot of $\kappa$ against $d$. The dashed line is $\frac{1}{d-1}$.

One can explain this result analytically as follows. Denote by $b_*$ the value of $b$ close to the tensor sphere (i.e. $b_* \sim L$) for which $\Delta\phi_\infty = \pi$. From Fig. 4.6 it is clear that most of the time advance arises from the region $r \sim L$. For a small black hole, we have $f \approx 1$ in this region (see end of section 4.1.3). This suggests that we can calculate $D_\infty$ for a small black hole by approximating $f = 1$ in the integral for the time delay. If we do this then the integrand now varies only over the scale $L$ and so all quantities in the problem are of order $L$. Hence, by dimensional analysis, $D_\infty$ must be proportional to $L$. To get an idea of the error made in setting $f = 1$ we note that this approximation eliminates the usual GR time delay effect. We can estimate the error made by our approximation by estimating the size of this delay as $\mu/b_*^{d-4} \sim \mu/L^{d-4} = L^3/\alpha$. For a small black hole $L \ll \sqrt{\alpha}$ hence the error is parametrically smaller than the scale $L$ and therefore negligible.

In summary, we have shown that, for a small black hole with $d \geq 7$, there is a tensor-polarized graviton trajectory with impact parameter $b \sim L$ that experiences no net deflection and, in the infinite cavity limit, experiences a finite time advance of order $L$.

Finally, we study the cases $d = 5, 6$ for which we cannot consider arbitrarily small black holes because of the failure of hyperbolicity.[37] We want to show that a negative time delay is possible for $d = 5, 6$. To do this, consider the case in which $R = R_0(1+\epsilon)$ with $\epsilon \ll 1$. Under the change of variable $r = R_0(1+x)$ the integral

[37]In $d = 5$ the theory is hyperbolic in the exterior of the black hole for $r_H/\sqrt{\alpha} > \sqrt{1+\sqrt{2}} \approx 1.6$, while in $d = 6$ this happens for $r_H/\sqrt{\alpha} > \left(\sqrt{5\left(5+2\sqrt{6}\right)}-1\right)^{-1/2} \approx 0.4$

for $\Delta t$ becomes:

$$\Delta t = 2R_0 \int_0^{\epsilon} dx \, \left( f(R_0(1+x)) \sqrt{1 - \frac{f(R_0(1+x))c_A(R_0(1+x))b^2}{R_0^2(1+x)^2}} \right)^{-1} . \quad (4.2.26)$$

We can now expand in powers of $x \ll 1$ and integrate:

$$\Delta t = \frac{4R_0\sqrt{\epsilon}}{\sqrt{f(R_0)}\sqrt{2 - \frac{R_0 f'(R_0)}{f(R_0)} - \frac{R_0 c_A'(R_0)}{c_A(R_0)}}} + \mathcal{O}(\epsilon^{3/2}). \quad (4.2.27)$$

Similarly one can compute the deflection angle:

$$\Delta\phi = \frac{4\sqrt{c_A(R_0)}\sqrt{\epsilon}}{\sqrt{f(R_0)}\sqrt{2 - \frac{R_0 f'(R_0)}{f(R_0)} - \frac{R_0 c_A'(R_0)}{c_A(R_0)}}} + \mathcal{O}(\epsilon^{3/2}). \quad (4.2.28)$$

It follows that the time delay is given by:

$$D = \frac{4R_0\sqrt{\epsilon}}{\sqrt{f(R_0)}\sqrt{2 - \frac{R_0 f'(R_0)}{f(R_0)} - \frac{R_0 c_A'(R_0)}{c_A(R_0)}}} \left(1 - \sqrt{c_A(R_0)}\right) + \mathcal{O}(\epsilon^{3/2}). \quad (4.2.29)$$

Since $c_S, c_V < 1$ we see that in this setting the time delay is always positive for scalar and vector modes, as expected. For tensor modes this can be negative. In particular, in $d = 5$ we have $c_T > 1$ everywhere and thus, for a black hole of arbitrary size we have a negative time delay when $R = R_0(1+\epsilon)$. In $d = 6$, for $r \gtrsim L$ we also have $c_T > 1$. Motivated by this, we compute numerically the time delay in $d = 5, 6$ for values of $R$ comparable to $R_0$. The numerics confirm that it is possible to obtain a time advance when $d = 5, 6$ (Figure 4.8). We have also studied the deflection angle, which we find is always greater than $\pi$, so zero net deflection trajectories do not occur for $d = 5, 6$.

## 4.3 Time machines

In this section we will discuss the suggestion that one can exploit the negative Shapiro time delay to construct a causality violating spacetime, i.e., a "time machine", in Einstein–Gauss–Bonnet theories [41]. This argument is closely related to arguments applying to any Lorentz covariant field theory with superluminal propagation, some of which have appeared in Ref. [83], which considers various flat space field theories with superluminal propagation. Such time machine constructions have been criticized by Geroch [84] (see also Ref. [85]). In this section we will discuss how these criticisms apply to the constructions of Refs. [41, 83].

Consider first the case of General Relativity. There are many solutions of the Einstein equation which exhibit causality violation e.g. Minkowski spacetime with

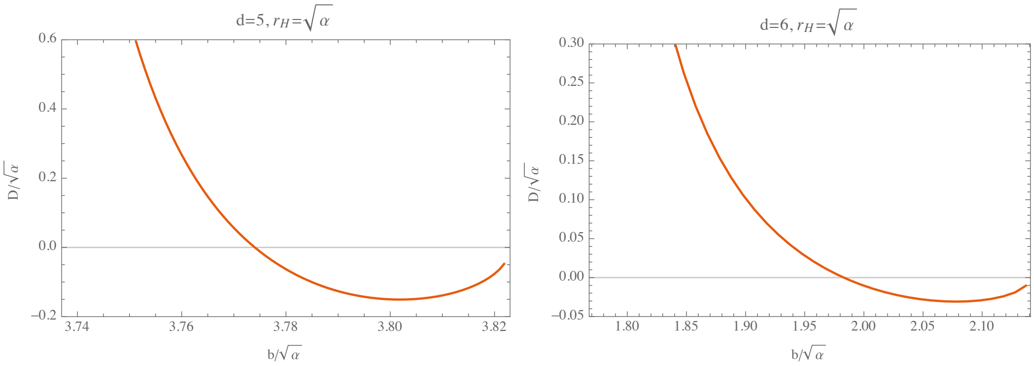


Figure 4.8: Time delay for tensor-polarized gravitons in $d = 5, 6$ dimensions. In $d = 5$ we set $r_H = 2\sqrt{\alpha}$, which gives $\mu = 4.5\alpha$, $L \approx 1.5\sqrt{\alpha}$, and we choose $R = 3\sqrt{\alpha}$. In $d = 6$ we set $r_H = \sqrt{\alpha}$, which gives $\mu = 1.5\alpha^{3/2}$, $L \approx 1.1\sqrt{\alpha}$, and we choose $R = 2\sqrt{\alpha}$.

a periodically identified time direction. We do not reject GR as a physical theory because it admits such solutions. This is because one cannot "make" these time machines starting from initial data. Stated mathematically: such causality-violating solutions are not the Cauchy development of any initial data.[38]

Let's consider now the argument of Ref. [83], which discussed several flat spacetime field theories with superluminal propagation. The simplest example is a scalar field with action

$$S = -\frac{1}{2}\int d^4x \left[\eta^{\mu\nu}\partial_\mu\pi\partial_\nu\pi - \frac{c_3}{\Lambda^4}\left(\eta^{\mu\nu}\partial_\mu\pi\partial_\nu\pi\right)^2\right], \tag{4.3.1}$$

where $\Lambda$ is a mass scale and $c_3$ a dimensionless constant. For this theory, the equation of motion is:

$$E(\pi, \partial\pi, \partial^2\pi) \equiv \left(1 - \frac{2c_3}{\Lambda^4}\partial\pi\cdot\partial\pi\right)\partial^\mu\partial_\mu\pi - \frac{4c_3}{\Lambda^4}\partial^\mu\pi\partial^\nu\pi\partial_\mu\partial_\nu\pi = 0. \tag{4.3.2}$$

For this equation, a hypersurface is characteristic if, and only if, it is null w.r.t. the (inverse) "effective metric":

$$G^{\mu\nu}(\partial\pi) = \left[1 - \frac{2c_3}{\Lambda^4}(\partial\pi\cdot\partial\pi)\right]\eta^{\mu\nu} - \frac{4c_3}{\Lambda^4}\partial^\mu\pi\partial^\nu\pi. \tag{4.3.3}$$

[38]More generally, the class of spacetimes of interest in GR is the class of spacetimes that arises as the maximal Cauchy development of suitable initial data, where "suitable" depends on the physical situation e.g. one would usually require the initial data to be geodesically complete and impose some asymptotic boundary condition e.g. asymptotic flatness. By definition, the maximal development is globally hyperbolic and so it can never violate causality. But the maximal development might be *extendible* beyond a Cauchy horizon into a causality violating region, which would capture the notion of formation of a time machine. The strong cosmic censorship conjecture asserts that, for suitable initial data, the maximal development is generically inextendible. Hence, if correct, strong cosmic censorship excludes time machines because either they can't be formed or they are infinitely fine-tuned (non-generic).

Assume that there exists a fiducial inertial frame for which $|\partial_\mu \pi| \ll \Lambda^2$.[39] Then $G^{\mu\nu}$ has Lorentzian signature, which implies that the equation of motion is hyperbolic. Note that Sylvester's law of inertia guarantees that the signature of $G^{\mu\nu}$ remain invariant under a change of basis, hence the above statement regarding the hyperbolicity of the equations is frame independent. It is $G^{\mu\nu}$ that determines causal properties of this equation. Following the terminology of the Introduction, we say that a covector $\xi$ is timelike iff it is timelike w.r.t. $G^{\mu\nu}$ etc.

Inverting $G^{\mu\nu}$ gives the effective metric. In the fiducial inertial frame we have

$$G_{\mu\nu}(\partial\pi) \approx \eta_{\mu\nu} + \frac{4c_3}{\Lambda^4}\partial_\mu\pi\partial_\nu\pi. \tag{4.3.4}$$

Contracting with a vector $X^\mu$ gives

$$G_{\mu\nu}X^\mu X^\nu \approx \eta_{\mu\nu}X^\mu X^\nu + \frac{4c_3}{\Lambda^4}(X\cdot\partial\pi)^2. \tag{4.3.5}$$

From this it can be seen that the null cones of $G_{\mu\nu}$ and $\eta_{\mu\nu}$ are nested, with the null cone of $G_{\mu\nu}$ inside that of $\eta_{\mu\nu}$ when $c_3 > 0$ and outside that of $\eta_{\mu\nu}$ when $c_3 < 0$, i.e., the theory has superluminal propagation when $c_3 < 0$ [83].

We can now discuss the initial value problem. Given some inertial frame $x^\mu$, we would like to specify initial data $(\pi, \partial_0\pi)$ on $\Sigma = \{x^0 = 0\}$. Of course $\Sigma$ is spacelike w.r.t. $\eta^{\mu\nu}$ but for a well-posed problem it is necessary that $\Sigma$ also be spacelike w.r.t. $G^{\mu\nu}$. For $c_3 > 0$ this is automatic. If $c_3 < 0$ then this appears to restrict our freedom to choose the initial data for $\partial_0\pi$. But this is not a new restriction: we already imposed a restriction on the initial data, i.e., the existence of the fiducial inertial frame. In the fiducial frame, the surface $x^0 = 0$ is obviously spacelike w.r.t. $G^{\mu\nu}$.

If, in some inertial frame, the surface $x^0 = 0$ is not spacelike w.r.t. $G^{\mu\nu}$ then the initial value problem will not be well-posed. In this case, for generic initial data, one would not expect a solution of the equation of motion to exist, even locally near $x^0 = 0$. One might be able to find a solution for very special initial data e.g. if the data is analytic then a solution will exist locally by the Cauchy-Kowalevskaya theorem. But this is infinitely fine-tuned: if one perturbs the initial data in a compact region then the resulting data will be non-analytic and no solution can be expected to exist. More generally, the solution does not depend continuously on the initial data.

Ref. [83] argues heuristically that it is possible to construct a time machine when $c_3 < 0$ by considering two lumps of scalar field, well-separated in the $x^2$ direction, which are highly boosted w.r.t. to each other in the $x^1$-direction. So consider initial data at $x^0 = 0$ consisting of two such lumps with a large relative boost.[40] The

[39] This is probably required for the validity of effective field theory.

[40] Note that such initial data will not satisfy the condition $|\partial_\mu\pi| \ll \Lambda^2$ everywhere.

problem is that such an initial data surface is not everywhere spacelike w.r.t. $G^{\mu\nu}$ [83] so this initial value problem is not well-posed. In general one would not expect any solution of the equations of motion to exist for such initial data. So one cannot build a time machine this way.

One might argue that it is obvious that a time machine could never result from Cauchy evolution of initial data since Cauchy evolution will break down when one is on the threshold of forming a time machine. So the question we should really ask is whether such a breakdown can occur starting from "good" initial data. In more physical terms: if one wishes to employ a large relative boost to build a time machine one must specify how this large relative boost will arise from "good" initial data [84].

For the scalar field theory above, "good" means that the initial data surface $x^0 = 0$ should be spacelike w.r.t. $G^{\mu\nu}$. Cauchy evolution remains well-posed as long as surfaces of constant $x^0$ remain spacelike w.r.t. $G^{\mu\nu}$ so a necessary condition for formation of a time machine would be existence of a time $T > 0$ at which the solution remains smooth but the surface $x^0 = T$ becomes null w.r.t. $G^{\mu\nu}$ at one more more points. This would correspond to the threshold of formation of the time machine.

Can this happen? It is well-known that such behaviour does not occur starting from *small* initial data, i.e., data such that $\pi$ and its first few derivatives are small. The solution arising from small initial data simply disperses in a similar way to a solution of the linear wave equation [90, 91]. So superluminal propagation leads to no pathologies in the behaviour of solutions arising from small initial data. Not much is known about the global behaviour of solutions of non-linear wave equations for *large* initial data. For most non-linear equations, global regularity of solutions is not expected. Solutions can suffer shock formation, i.e., blow-up of the field $\pi$ (or a derivative of $\pi$) at some time $T > 0$. See for example Ref. [92] (albeit not for a Lorentz covariant equation). As far as we know, it is not excluded that the equation discussed above could have large data solutions that evolve to the threshold of formation of a time machine. But, as we have discussed, there is no compelling reason to believe that this is the case.

We now turn to the proposal of Ref. [41] that it is possible to construct a time machine in EGB theory. This is done by exploiting the negative Shapiro time delay experienced by gravitons. The proposed time machine arises from two high energy gravitons, moving in opposite directions with non-zero impact parameter. Each graviton is described by an Aichelburg–Sexl "shock-wave" solutions [79, 80]. It is assumed that the spacetime resulting from the collision is well-approximated by two outgoing Aicheburg–Sexl shock waves. Under these assumptions one can argue that there exist closed causal curves in the spacetime.

One problem with this construction is the use of Aichelburg–Sexl solutions. The curvature of an AS solution is a delta-function localized on a null hypersurface (with the amplitude of the delta-function diverging on a null line within this hypersurface: this is viewed as the worldline of the graviton). Owing to special symmetries of its curvature tensor, this is an exact solution of Einstein–Gauss–Bonnet theory. Now clearly one can superpose two such solutions moving in opposite directions (since the spacetime is flat between them) to obtain a solution valid until the two null hypersurfaces intersect. But when these hypersurfaces intersect, it is far from clear that there is any sense in which the equation of motion can be satisfied. This is because the equation of motion involves products of curvature tensors. Hence along the line of intersection of the hypersurfaces, there will be a product of delta functions that cannot be balanced. Therefore it seems unlikely that the spacetime can be extended to the future of this intersection.[41]

This problem arises from the fact that an AS solution is singular. So maybe we can solve the problem by smoothing out the singularity. As discussed in the Introduction, an AS solution can be obtained by taking a limit in which one boosts a black hole solution and takes the boost to infinity whilst scaling the black hole mass to zero, keeping the total energy fixed. This suggests that we should consider initial data consisting of two small (compared to the GB scale) black holes, moving with high relative boost in opposite directions with large impact parameter. It would be a difficult matter to construct such data explicitly, solving the constraint equations, but there is no reason to doubt that this can be done.

Now the question is whether this is “good” initial data. As discussed above, black holes with arbitrarily small mass don’t exist for $d = 5$. When $d = 6$, small black holes are unphysical because the equation of motion is not hyperbolic. So consider $d \geq 7$. In section 4.1.2 we showed that there is a speed limit for small black holes arising from the condition that the initial data surface be spacelike. Hence we cannot start from initial data describing two black holes with a very large relative boost: such initial data will not be everywhere spacelike and hence this data cannot be evolved (or is infinitely fine-tuned), just as for the scalar field example discussed above.

We can attempt to construct legitimate initial data by requiring that the speed limit is respected. Consider two small black holes, each of mass parameter $\mu$, boosted in opposite directions with speed $v \leq v_{\max}$, separated in the transverse direction by a distance $R$. Assume that the distance is sufficiently large for the gravitational interaction between the black holes to be negligible, i.e. $R \gg L$. Consider a tensor-

[41] Note that a theory of interacting impulsive (i.e. delta-function curvature) gravitational waves does exist for GR [93].

polarized graviton propagating between the holes. Our numerical results suggested that, for a black hole at rest, the maximum time advance $|D|$ that a graviton can experience is of order $L$ and is achieved for $b \sim L$. If we now boost the black hole in a direction transverse to the motion of the graviton ( in order to avoid issues related to the length contraction effect), the time advance gets amplified by $\gamma = (1 - v^2)^{-1/2}$: $|D| \sim \gamma L$. For the argument of [41] to work we need the time advance to "compensate" the time taken by the graviton to travel between the two black holes, i.e. $|D| \sim R$. However for this to hold we would need $\gamma L \sim R \gg L$, that is $\gamma \gg 1$, which cannot be achieved because of the restriction $v < v_{\text{max}}$.

In summary, we have argued that attempting to build a time machine spacetime in EGB theory using the method suggested in Ref. [41] will not work. This is because the initial data required is not everywhere spacelike (in the sense defined in the Introduction) so the initial value problem is not well-posed: either no solution will exist, or it will be infinitely fine-tuned.

# Appendix

## 4.A Time delay: perturbative calculations

We give here more details on the perturbative calculation of the time delay and deflection angle. Recall from section 4.2.3 that we want to compute:

$$\Delta\phi = 2b \int_{R_0}^{R} dr\ c_A(r) r^{-2} h(r), \quad \Delta t = 2 \int_{R_0}^{R} dr\ f(r)^{-1} h(r), \tag{4.A.1}$$

where we have introduced:

$$h(r) = \left(1 - \frac{f(r) c_A(r) b^2}{r^2}\right)^{-1/2}. \tag{4.A.2}$$

We want to calculate the above quantities subject to the assumption that $R_0$ is large compared to the black hole size in the following sense

$$\frac{\mu}{R_0^{d-3}} \ll 1, \quad \frac{L}{R_0} \ll 1 \tag{4.A.3}$$

We will assume that the cavity radius is large:

$$\frac{R}{R_0} \gg 1 \tag{4.A.4}$$

For the time delay we will need to assume that the cavity radius is not *too* large:

$$\frac{R}{R_0} \ll \frac{R_0^{d-3}}{\mu}, \left(\frac{R_0}{L}\right)^{d-1}. \tag{4.A.5}$$

## Approximation for $h(r)$

The impact parameter is related to $R_0$ by equation (4.2.10). Using (4.A.3) we have:

$$b^2 = R_0^2 \left(1 + \frac{\mu}{R_0^{d-3}} \left(1 - \frac{2\alpha\beta_A}{R_0^2}\right)\right) + \dots \tag{4.A.6}$$

similarly, introducing $z = R_0/r$ (so $0 < z \leq 1$):

$$\frac{f(r)c_A(r)}{r^2} = \frac{z^2}{R_0^2} \left(1 - \frac{\mu}{R_0^{d-3}} \left(z^{d-3} - \frac{2\alpha\beta_A}{R_0^2} z^{d-1}\right)\right) + \dots \tag{4.A.7}$$

And hence:

$$h(r) = (1-z^2)^{-1/2} \left(1 + \frac{1}{2} \frac{\mu}{R_0^{d-3}} \frac{z^2}{1-z^2} \left((1-z^{d-3}) - \frac{2\alpha\beta_A}{R_0^2}(1-z^{d-1})\right)\right) + \dots \tag{4.A.8}$$

and the ellipsis denotes terms of order $\mathcal{O}\left(\left(\frac{\mu}{R_0^{d-3}} + \frac{L^{d-1}}{R_0^{d-1}}\right)^2\right)$.

## Approximation for the deflection angle

Changing the integration variable to $z$:

$$\Delta\phi = 2\frac{b}{R_0} \int_{R_0/R}^{1} dz\, c_A h \tag{4.A.9}$$

In our approximation, denoting by $\dots$ terms of order $\mathcal{O}\left(\left(\frac{\mu}{R_0^{d-3}} + \frac{L^{d-1}}{R_0^{d-1}}\right)^2\right)$, we obtain:

$$c_A h = (1-z^2)^{-1/2} \left(1 + \frac{\mu}{R_0^{d-3}} \left(\frac{1}{2} z^2 \frac{(1-z^{d-3})}{(1-z^2)} - \frac{\alpha}{R_0^2}\beta_A \left(\frac{z^2 + z^{d+1} - 2z^{d-1}}{1-z^2}\right)\right)\right) + \dots \tag{4.A.10}$$

Which yields:

$$\Delta\phi = 2\arccos\left(\frac{R_0}{R}\right) + \frac{\mu}{R_0^{d-3}} \left[2\left(\frac{1}{2} - \frac{\alpha\beta_A}{R_0^2}\right)\arccos\left(\frac{R_0}{R}\right) + 2J\right] + \dots \tag{4.A.11}$$

where we have defined:

$$J = \int_{R_0/R}^{1} dz\, (1-z^2)^{-1/2} \left(\frac{1}{2} z^2 \frac{(1-z^{d-3})}{(1-z^2)} - \frac{\alpha\beta_A}{R_0^2} \left(\frac{z^2 + z^{d+1} - 2z^{d-1}}{1-z^2}\right)\right). \tag{4.A.12}$$

Using the large cavity radius assumption (4.A.4) we have $J = J_0 + \mathcal{O}(R_0^2/R^2)$, where

$$J_0 = \sqrt{\pi}(d-1)\left(1 - \frac{2\alpha\beta_A}{R_0^2}\frac{(d-2)}{(d-1)}\right) \frac{\Gamma\left(\frac{d}{2}\right)}{2\Gamma\left(\frac{d+1}{2}\right)} - \frac{\pi}{2}\left(\frac{1}{2} - \frac{\alpha\beta_A}{R_0^2}\right). \tag{4.A.13}$$

Moreover we have that

$$2\arccos\left(\frac{R_0}{R}\right) = \pi - 2\frac{R_0}{R} + \mathcal{O}\left(\frac{R_0^2}{R^2}\right) \tag{4.A.14}$$

and hence

$$2\left(\frac{1}{2}-\frac{\alpha\beta_A}{R_0^2}\right)\arccos\left(\frac{R_0}{R}\right)+2J=\sqrt{\pi}(d-1)\left(1-\frac{2\alpha\beta_A}{R_0^2}\frac{(d-2)}{(d-1)}\right)\frac{\Gamma\left(\frac{d}{2}\right)}{2\Gamma\left(\frac{d+1}{2}\right)}+\mathcal{O}\left(\frac{R_0}{R}\right). \tag{4.A.15}$$

We can then conclude that

$$\Delta\phi=\pi-2\frac{R_0}{R}+\frac{\mu}{R_0^{d-3}}\left[\sqrt{\pi}(d-1)\left(1-\frac{2\alpha\beta_A}{R_0^2}\frac{(d-2)}{(d-1)}\right)\frac{\Gamma\left(\frac{d}{2}\right)}{2\Gamma\left(\frac{d+1}{2}\right)}\right]$$
$$+\mathcal{O}\left(\left[\frac{R_0}{R}+\left(\frac{\mu}{R_0^{d-3}}+\frac{L^{d-1}}{R_0^{d-1}}\right)\right]^2\right). \tag{4.A.16}$$

## Approximation for the proper time

Denoting again by ... terms of order $\mathcal{O}\left(\left(\frac{\mu}{R_0^{d-3}}+\frac{L^{d-1}}{R_0^{d-1}}\right)^2\right)$, we have:

$$f(z)^{-1}h(z)=(1-z^2)^{-1/2}\left(1+\frac{\mu}{R_0^{d-3}}z^2\left(z^{d-5}+\frac{1}{2}\frac{(1-z^{d-3})}{1-z^2}-\frac{2\alpha\beta_A}{R_0^2}\frac{(1-z^{d-1})}{1-z^2}\right)\right)+\ldots \tag{4.A.17}$$

from which:

$$\Delta t=2\sqrt{R^2-R_0^2}+2\frac{\mu}{R_0^{d-4}}I+\ldots \tag{4.A.18}$$

where:

$$I=\int_{R_0/R}^{1}dz\,(1-z^2)^{-1/2}\left(z^{d-5}+\frac{1}{2}\frac{(1-z^{d-3})}{(1-z^2)}-\frac{\alpha\beta_A}{R_0^2}\frac{(1-z^{d-1})}{1-z^2}\right). \tag{4.A.19}$$

For large cavity radius, i.e., (4.A.4), we have:

$$\Delta t=2R-\frac{R_0^2}{R}+2\frac{\mu}{R_0^{d-4}}I_0+\mathcal{O}\left(R_0\left[\frac{R_0}{R}+\left(\frac{\mu}{R_0^{d-3}}+\frac{L^{d-1}}{R_0^{d-1}}\right)\right]^2\right), \tag{4.A.20}$$

where

$$I_0=\sqrt{\pi}\left(1-\frac{2\alpha\beta_A}{R_0^2}\frac{(d-4)}{(d-3)}\right)\frac{(d-1)(d-3)}{(d-4)}\frac{\Gamma\left(\frac{d}{2}\right)}{4\Gamma\left(\frac{d+1}{2}\right)}. \tag{4.A.21}$$

Plugging back in the above we obtain

$$\Delta t=2R+\frac{\mu}{R_0^{d-4}}\left[\sqrt{\pi}\left(1-\frac{2\alpha\beta_A}{R_0^2}\frac{(d-4)}{(d-3)}\right)\frac{(d-1)(d-3)}{(d-4)}\frac{\Gamma\left(\frac{d}{2}\right)}{2\Gamma\left(\frac{d+1}{2}\right)}\right]$$
$$+\mathcal{O}\left(R_0\left[\frac{R_0}{R}+\left(\frac{\mu}{R_0^{d-3}}+\frac{L^{d-1}}{R_0^{d-1}}\right)\right]^2+\frac{R_0^2}{R}\right). \tag{4.A.22}$$

Finally, note that in this approximation we have

$$f(R)=1+\mathcal{O}\left(\frac{\mu}{R^{d-3}}\right)\quad\text{and}\quad\frac{\mu}{R^{d-3}}=\frac{\mu}{R_0^{d-3}}\left(\frac{R_0}{R}\right)^{d-3} \tag{4.A.23}$$

which is negligible to the order of approximation used above, hence

$$\Delta\tau\approx\Delta t. \tag{4.A.24}$$

## Approximation for the time in Minkowski

We have:

$$\begin{aligned}2R\sin\left(\Delta\phi/2\right) &= 2R\sin\left[\arccos(R_0/R) + \frac{\mu}{R_0^{d-3}}\left(\left(\frac{1}{2} - \frac{\alpha\beta_A}{R_0^2}\right)\arccos\left(R_0/R\right) + J\right) + \ldots\right] \\ &= 2\sqrt{R^2 - R_0^2} + \frac{\mu}{R_0^{d-4}}\left(\left(\frac{1}{2} - \frac{\alpha\beta_A}{R_0^2}\right)\arccos\left(R_0/R\right) + J\right) \\ &\quad + \mathcal{O}\left(R\left(\frac{\mu}{R_0^{d-3}} + \frac{L^{d-1}}{R_0^{d-1}}\right)^2\right). \qquad (4.A.25)\end{aligned}$$

We now use the condition (4.A.5) that the cavity is not too large. This ensures that the last term above is small and we obtain:

$$\Delta t_{\text{Mink}} = 2R + \frac{\mu}{R_0^{d-4}}\left(\sqrt{\pi}(d-1)\left(1 - \frac{2\alpha\beta_A}{R_0^2}\frac{(d-2)}{(d-1)}\right)\frac{\Gamma\left(\frac{d}{2}\right)}{2\Gamma\left(\frac{d+1}{2}\right)}\right) \qquad (4.A.26)$$

$$+\,\mathcal{O}\left(R\left[\frac{R_0}{R} + \left(\frac{\mu}{R_0^{d-3}} + \frac{L^{d-1}}{R_0^{d-1}}\right)\right]^2\right). \qquad (4.A.27)$$

## Approximation for the time delay

Finally, putting the above results together, we have determined the time delay under the conditions (4.A.3), (4.A.4), (4.A.5):

$$D = \frac{\mu}{R_0^{d-4}}\left[\sqrt{\pi}\left(1 - \frac{2\alpha\beta_A}{R_0^2}\frac{(d-4)}{(d-1)}\right)\frac{(d-1)}{(d-4)}\frac{\Gamma\left(\frac{d}{2}\right)}{2\Gamma\left(\frac{d+1}{2}\right)}\right] \qquad (4.A.28)$$

$$+\,\mathcal{O}\left(R\left[\frac{R_0}{R} + \left(\frac{\mu}{R_0^{d-3}} + \frac{L^{d-1}}{R_0^{d-1}}\right)\right]^2\right). \qquad (4.A.29)$$

# Chapter 5

# On the local well-posedness of Lovelock theories

In this chapter we will discuss the initial value problem for Lovelock theories. The contents of this Chapter are the results of original research conducted in collaboration with my supervisor, Harvey Reall, and have been published in [44].

## 5.1 The initial value problem in Lovelock gravity

In this section we will discuss the initial value problem for Lovelock theories of gravity in vacuum. Similar considerations will apply to the Horndeski case analysed in Chapter 6.

In Chapter 2 we highlighted the importance of the initial value problem in physics and we reviewed how one may obtain a well-posed formulation for the Einstein equations. As discussed in the Introduction, a corresponding result is still missing for most modified gravity theories, in particular for Lovelock and Horndeski theories. Recall that in order to consider a theory physical we must prove that the initial value problem be (at least locally) well-posed. It is therefore important to investigate whether these theories admit a well-posed initial value problem, for a negative result would mean that they do not constitute viable alternatives to Einstein's General Relativity.

In order to discuss the initial value problem for a diffeomorphism-covariant theory, we will first need to address the issue of gauge fixing. For the Einstein equations it is well-known that the choice of harmonic coordinates leads to a well-posed initial value formulation [42] (see Section 2.5). We will use the same gauge in our study of Lovelock gravity.

The question of well-posedness for Lovelock gravity was first addressed by

Choquet-Bruhat in Ref. [77], where she dealt with the harmonic gauge Lovelock equations. In harmonic gauge, similarly to what happens for the Einstein equations (cf. Sec 2.1), the vacuum Lovelock equations separate in a set of constraints for the initial data and an evolution equation for the metric

$$g^{\alpha\beta}\partial_\alpha\partial_\beta g_{\mu\nu} = \mathcal{N}(g, \partial g, \partial^2 g)_{\mu\nu}. \tag{5.1.1}$$

In contrast with the Einstein equations $\mathcal{N}$ depends on second derivatives of the metric and, moreover, it does so in a non-linear manner. It can be shown, however, that these equations are actually *linear* in the "time" derivatives $\partial_0^2 g$ [76, 77, 70].[42] Thanks to this property it is possible to use the Cauchy–Kovalevskaya Theorem to establish local existence and uniqueness for *analytic* data [77]. However, this is far from ideal. In this setting, in general, the solution will not depend continuously on the initial data. In particular, if we smoothly perturb the initial data in a compact region, then generically there will exist no solution arising from the perturbed data. Furthermore, analytic data is somewhat unphysical as, in a certain sense, it does away with causality: the solution can be described in terms of a power series whose coefficients are constants which can be entirely determined from the initial data by solving a set of algebraic equations at a point; in other words the solution is determined everywhere in an open neighbourhood by its behaviour at a point. Moreover, in the class of analytic solutions, it is not possible to make sense of the finite speed of propagation of signals, a fundamental feature of systems described by hyperbolic equations. Roughly speaking, finite speed of propagation means that if a solution is compactly supported at some instant of time then it must be compactly supported at all later times, with the support "moving" at a finite speed. Analytic solutions do not have this property, since they can only be compactly supported if they vanish identically. These reasons highlight that the results of [77] are unsatisfactory and that one must consider a more general class of data.

The main issue with establishing the well-posedness of the initial value problem for Lovelock theories is that the equations are not necessarily *hyperbolic*. Recall from Chapter 2 that for the initial value problem to be well-posed it is necessary that the equations of motion be *hyperbolic*. In particular we distinguished two notions of hyperbolicity: *weak* and *strong* hyperbolicity, with the latter implying the former (see Definitions 5 and 6). Roughly speaking, *weak hyperbolicity* means that there are no solutions which grow exponentially fast in time and frequency, while *strong hyperbolicity* corresponds to the existence of an "energy estimate" bounding the energy of a solution at a time $t$ in terms of the energy of the initial data. The existence of

[42] This property holds in any gauge.

such "energy estimate" is sufficient to establish, via a standard technique, the local well-posedness of the linearised problem. Moreover, energy estimates constitute a standard tool used in establishing the well-posedness of the non-linear problem.

In Chapter 2 we argued that for a non-linear problem to be well-posed, it is *necessary* that all the initial value problems obtained by *linearising* the original problem around any background solution in a given open neighbourhood in solution space be well posed. Since the well-posedness of the linearised problems is related to the hyperbolicity of the equations of motion, we infer that a *necessary* condition for the well-posedness of these theories is that their equations of motion be *strongly hyperbolic*.

We have seen in Section 2.5 that this condition is always verified in GR, independently of the background. In Lovelock theories, however, it is know that even weak hyperbolicity can fail when the background curvature becomes too large. As discussed in Chapter 4, it was shown in Ref. [70] that weak hyperbolicity fails (in any gauge) for linear perturbations of "small" black hole solutions of Lovelock theories. Here "small" refers to the length scale set by the dimensionful coupling constants of such theory. More generally, one expects that weak hyperbolicity will fail in a large class of backgrounds with large curvature. These examples show that the equations of motion of Lovelock theory are *not* always weakly hyperbolic. Hence, for general initial data one cannot expect local well-posedness. However, one might hope that restricting the initial data so that the equations of motion *are* hyperbolic will lead to a well-posed initial value problem. In particular, one might expect that this issue could be resolved by restricting to "weak curvature" backgrounds.

The equations of motion for Lovelock gravity can be split in an Einstein part and Lovelock "correction". The "size" of this correction is determined by the magnitude of the coupling constants and the background curvature. When these are small, the equations can be considered as "small perturbations" of the Einstein equations. Naively, one would expect that if the background curvature and the couplings are sufficiently small then these corrections should not spoil the strong hyperbolicity of the Einstein equations. This, however, is not necessarily true; the reason being that, albeit small, these corrections affect the highest order derivatives and hence the structure of the principal symbol. Let us illustrate this problem with an example. Consider the following system of scalar fields in two-dimensional Minkowski space:

$$\Box\varphi = k\,\epsilon\,\partial_0\partial_1\psi \qquad \Box\psi = -k\,\epsilon\,\partial_0\partial_1\varphi. \tag{5.1.2}$$

We can consider this as a toy model for the equations governing linear perturbations around a "weak field" background solutions in Lovelock or Horndeski theory. The parameter $k$ will play the role of the coupling constant of the theory, while $\epsilon$ will be a measure of the strength of the background fields. Using the formalism of Chapter 2 we can study the hyperbolicity of this system. For small $\epsilon$ (i.e. "weak fields"), the matrix $M$ has the following eigenvalues

$$\xi_0^{\pm} = \pm\xi_1 \left(1 + \frac{i}{2}k\epsilon + \mathcal{O}(\epsilon^2)\right) \qquad \tilde{\xi}_0^{\pm} = \pm\xi_1 \left(1 - \frac{i}{2}k\epsilon + \mathcal{O}(\epsilon^2)\right) \tag{5.1.3}$$

From this we can deduce that when the coupling vanishes, $k = 0$, the eigenvalues are all real and hence the system is (strongly) hyperbolic. However, for $k \neq 0$, the eigenvalues will fail to be real, even if $\epsilon \ll 1$. In particular they are all complex and hence the perturbation has made an hyperbolic system into an *elliptic* one, which does not admit a well-posed initial value problem. In other words, this example shows that if we perturb a strongly hyperbolic system at the level of the highest order derivatives, then the resulting system might fail to be hyperbolic, no matter how small the perturbation is.[43]

One of the main results of this dissertation will be to show that, in fact, Lovelock and Horndeski theories do not suffer from such problem. We will prove in Sec. 5.4 and 6.4 that both these theories are *weakly hyperbolic* when the background fields are small.

As discussed in Chapter 2, weak hyperbolicity is not sufficient to establish the well-posedness of the theory. In order for the initial value problem to be well-posed we will need the equations to be *strongly hyperbolic*.

Our most important results concern strong hyperbolicity of Lovelock and Horndeski theories. As discussed above, strong hyperbolicity is needed in order to establish local well-posedness of the initial value problem, and in numerical applications. However, we will prove that for Lovelock theories, in harmonic gauge, the linearised equation of motion is *not* strongly hyperbolic in a generic weakly curved background. The word "generic" is important here: there certainly exist particular backgrounds for which the linearised equation of motion *is* strongly hyperbolic (e.g. Minkowski spacetime [86]) so the equation of motion for linear perturbations around such backgrounds is locally well-posed. However, such backgrounds are non-generic e.g. they always have symmetries. In order to have any hope of establishing local well-posedness for the *non-linear* theory for weak fields, one would need strong hyperbolicity for *any* weakly curved background. This is not the case, at least not in harmonic gauge. Hence the most straightforward approach to establishing local well-posedness for Lovelock theories does not work.[44]

[43] There is, however, an important difference between the system (5.1.2) and a Lovelock (or Horndeski) theory, which is that (5.1.2) is not obtained from an action principle.

[44] Note that the recent discussion of local well-posedness in Ref. [94] simply *assumes* that the

Finally, one could argue that these results suggest that in order to have a well-posed problem one may need to further restrict initial data so that the equations are strongly hyperbolic. Besides being extremely restrictive in the first place, this argument does not work. We will show that even restricting to non-generic initial data for which the equations are strongly hyperbolic does not guarantee that hyperbolicity will be preserved throughout evolution. In particular, we will consider non-generic initial data with small curvature such that the equations of motion are strongly hyperbolic. The solution develops large curvature over time causing even *weak* hyperbolicity to fail dynamically.

The rest of the Chapter is organised as follows. We begin by deriving the linearised Lovelock equations in harmonic gauge (Section 5.2). In Section 5.4 we present the proof of weak hyperbolicity in a low curvature background, while in Section 5.5 we show that, in this setting, strong hyperbolicity does not hold. We conclude the Chapter by discussing how strong hyperbolicity may be violated dynamically (Section 5.6).

## 5.2 Equations of motion in harmonic gauge

We begin our study by deriving the linearised Lovelock equations in harmonic gauge and their principal symbol.

The vacuum equations of motion of Lovelock gravity in $d > 4$ spacetime dimensions are given by (cf. Section 3.2)

$$A^a{}_b[g] \equiv G^a{}_b + \Lambda\delta^a_b + \sum_{p\geq 2} k_p \delta^{ac_1c_2\dots c_{2p}}_{bd_1d_2\dots d_{2p}} R_{c_1c_2}{}^{d_1d_2}\cdots R_{c_{2p-1}c_{2p}}{}^{d_{2p-1}d_{2p}} = 0, \qquad (5.2.1)$$

where we assumed that the coefficient of the Einstein term does not vanish and is normalised to one. To investigate hyperbolicity we linearise around a background solution $g_{ab}$, i.e. we consider the metric perturbation $g_{ab} + h_{ab}$ and linearise in the perturbation $h_{ab}$

$$A_{ab}[g+h] = A_{ab}[g] + A^{(1)}_{ab}[h] + \dots \qquad (5.2.2)$$

so that the linearised equation of motion is

$$A^{(1)}_{ab}[h] = 0. \qquad (5.2.3)$$

The diffeomorphism covariance of the theory implies that we will need to choose an appropriate gauge in order to conclude whether the theory is strongly hyperbolic or not.

harmonic gauge equation of motion is suitably hyperbolic. Our result shows that this assumption is incorrect.

For the *non-linear* equation, one can choose *harmonic coordinates*:

$$0 = g^{\nu\rho}\nabla_\nu\nabla_\rho x^\mu = \frac{1}{\sqrt{-g}}\partial_\nu\left(\sqrt{-g}g^{\mu\nu}\right). \tag{5.2.4}$$

Upon linearisation this reduces to the Lorenz gauge condition for the linearised metric perturbation:

$$H_b \equiv \nabla^a h_{ab} - \frac{1}{2}\nabla_b h^a{}_a = 0. \tag{5.2.5}$$

Actually, linearising the harmonic gauge condition around a non-trivial background gives a *generalized* Lorenz gauge condition with a non-vanishing RHS. But this RHS does not depend on derivatives of $h_{ab}$ which implies that it does not affect the hyperbolicity analysis. Therefore we will just use the standard Lorenz gauge.

Although most properly referred to as Lorenz gauge, henceforth we will refer to (5.2.5) as harmonic gauge because it is inconvenient to use different words for the linear and non-linear gauge conditions. Of course, it is well-known that the gauge condition (5.2.5) can always be achieved by a suitable gauge transformation in the linearised theory.

In harmonic gauge, the Einstein equation is strongly hyperbolic (and in fact locally [42] and globally [58] well-posed). We will investigate whether the same is true for Lovelock theory. We will do this by studying the hyperbolicity of the linearised theory. The harmonic gauge linearised equation of motion is

$$\tilde{A}^{(1)}_{ab}[h] = 0, \tag{5.2.6}$$

where

$$\tilde{A}^{(1)}_{ab}[h] \equiv A^{(1)}_{ab}[h] - \nabla_{(a}H_{b)} + \frac{1}{2}g_{ab}\nabla^c H_c \tag{5.2.7}$$

This is the equation of motion whose hyperbolicity we will study.

A standard argument shows that the harmonic gauge condition is propagated by the harmonic gauge equation of motion [77]. The argument is based on the fact that the tensor $A_{ab}$ arises from a diffeomorphism covariant action and therefore satisfies a contracted Bianchi identity $\nabla^b A_{ab} = 0$. Linearising around a background solution gives, for any $h_{ab}$,

$$\nabla^b A^{(1)}_{ab}[h] = 0 \tag{5.2.8}$$

so, when (5.2.6) is satisfied, the divergence of (5.2.7) gives

$$\nabla^b\nabla_b H_a + R_{ab}H^b = 0. \tag{5.2.9}$$

This is a standard linear wave equation so provided the initial data is chosen such that $H_a$ and its first time derivative vanish then the solution will have $H_a \equiv 0$. (As for the Einstein equation, vanishing of the first time derivative of $H_a$ is equivalent, via the equation of motion, to the condition that the initial data satisfies the constraint equations [77].) This proves that the gauge condition (5.2.5) is propagated by the equation of motion (5.2.6). Hence the resulting solution will satisfy the original equation of motion (5.2.3).

The linearised harmonic gauge equation of motion (5.2.6) takes the form

$$P^{abcdef}\nabla_e\nabla_f h_{cd} + \ldots = 0 \tag{5.2.10}$$

where the ellipsis denotes terms involving fewer than two derivatives of $h_{ab}$. The coefficient here defines the principal symbol (3.2.11)

$$P(\xi)^{abcd} \equiv P^{abcdef}\xi_e\xi_f \tag{5.2.11}$$

for an arbitrary covector $\xi$. The coefficient is symmetric in $ab$ and in $cd$. It can be split into the terms coming from the (harmonic gauge) Einstein tensor, and those coming from the extra Lovelock terms:

$$P(\xi)^{abcd} = P_{\text{Einstein}}(\xi)^{abcd} + \delta P^{abcd}(\xi) \tag{5.2.12}$$

The Einstein term was computed in Section 2.5 and takes the form (2.5.16). For a symmetric tensor $t_{ab}$ we have

$$P_{\text{Einstein}}(\xi)^{abcd}t_{cd} = -\frac{1}{2}\xi^2 G^{abcd}t_{cd}$$

with $\xi^2 = g^{ab}\xi_a\xi_b$ and

$$G^{abcd} = \frac{1}{2}\left(g^{ac}g^{bd} + g^{ad}g^{bc} - g^{ab}g^{cd}\right). \tag{5.2.13}$$

Viewed as a quadratic form on symmetric tensors, $G^{abcd}$ has signature $(d, d(d-1)/2)$, i.e., $d$ negative eigenvalues and $d(d-1)/2$ positive eigenvalues.

The Lovelock contribution is given by [70] (see also Section 3.2, Eq. (3.2.14))

$$\delta P(\xi)^a{}_b{}^{cd}t_{cd} = \sum_{p\geq 2} k_p \delta^{ac_1\ldots c_{2p}}_{bd_1\ldots d_{2p}}\xi_{c_1}\xi^{d_1}t_{c_2}{}^{d_2}R_{c_3c_4}{}^{d_3d_4}\cdots R_{c_{2p-1}c_{2p}}{}^{d_{2p-1}d_{2p}}. \tag{5.2.14}$$

Recall from our discussion in Section 3.2 that the principal symbol satisfies the following identities

$$\delta P^{abcdef} = \delta P^{cdabef} \tag{5.2.15}$$

and

$$\delta P^{(a|bcd|ef)} = \delta P^{a(bc|de|f)} = 0, \tag{5.2.16}$$

from which it follows that

$$\xi_a\delta P^{abcd}(\xi) = \xi_b\xi_c\xi_f\delta P^{abcdef} = 0. \tag{5.2.17}$$

We will discuss these properties in more detail in Section 6.3.

## 5.3 Setting up the problem

We will investigate whether the harmonic gauge linearised Lovelock equation of motion is hyperbolic when the curvature of the background spacetime is small. Here, "small" means small compared to any of the scales defined by the dimensionful coupling constants $k_p$, so one expects the Lovelock terms in the equation of motion to be small compared to the Einstein term.

To relate to the discussion of Section 2.3 we need to introduce coordinates $x^\mu = (t, x^i)$. We assume that these are chosen so that surfaces of constant $t$ are spacelike, i.e. $g^{00} < 0$, which ensures that the initial value problem for the harmonic gauge linearised Einstein equation is well-posed. We want to ask whether the same is true for the harmonic gauge linearised Lovelock equation when the background curvature is small. Here, by *small*, we mean that there exists an orthonormal basis $\{e_\mu\}$ with $e_0$ orthogonal to surfaces of constant $t$, for which the magnitude of the largest component of the Riemann tensor is $L^{-2}$ where

$$|k_p| L^{-2p} \ll 1, \qquad \forall\, p \geq 2. \tag{5.3.1}$$

This ensures that the Lovelock terms in the principal symbol are small compared to the Einstein term.

The principal symbol $P(\xi)$ maps symmetric tensors to symmetric tensors so we regard it as a $N \times N$ matrix where $N = d(d+1)/2$. We define $N \times N$ matrices $A(x)$, $B(x, \xi_i)$ and $C(x, \xi_i)$ using equation (2.3.4), i.e.,

$$A = P^{00} \qquad B(\xi_i) = 2\xi_i P^{0i} \qquad C(\xi_i) = \xi_i \xi_j P^{ij}. \tag{5.3.2}$$

Here $\xi_i$ is real with $\xi_i \xi_i = 1$ (since this is what we need in the definitions of strong and weak hyperbolicity).[45] Throughout this section we will not write explicitly the dependence on the spacetime coordinates $x^\mu$. Note that these matrices are real and symmetric: the latter property arises because the equation of motion can be obtained from a Lagrangian (see Section 6.3).

Our assumption that the surfaces of constant $t$ are spacelike ensures that $A$ is invertible when the Lovelock terms are absent. Hence, by continuity, $A$ is also invertible when the background curvature is small. We can therefore define $M(\xi_i)$ as above, i.e.,

$$M(\xi_i) = \begin{pmatrix} 0 & I \\ -A^{-1}C(\xi_i) & -A^{-1}B(\xi_i) \end{pmatrix}. \tag{5.3.3}$$

[45] We denote $\xi_i \xi_i = \delta^{ij} \xi_i \xi_j$.

Recall that weak hyperbolicity is the requirement that the eigenvalues of this matrix are real. For strong hyperbolicity it is necessary that the eigenvalues are real and the matrix is diagonalizable.

From the discussion of section (2.3) we know that $\xi_0$ is an eigenvalue of $M(\xi_i)$ if, and only if, the corresponding eigenvector $v$ has the form

$$v = \begin{pmatrix} t \\ \xi_0 t \end{pmatrix} \tag{5.3.4}$$

for some non-zero symmetric $t_{\mu\nu}$ such that

$$P(\xi) \cdot t = 0 \tag{5.3.5}$$

where $\xi_\mu \equiv (\xi_0, \xi_i)$ in the argument of $P$.

Consider first the case of the linearised Einstein equation. Since $G^{abcd}$ is non-degenerate, equation (2.5.16) implies that $\xi_\mu$ is characteristic if, and only if, it is null:

$$P_{\text{Einstein}}(\xi) \cdot t = 0, \qquad t \neq 0 \qquad \Leftrightarrow \qquad g^{\mu\nu}\xi_\mu\xi_\nu = 0. \tag{5.3.6}$$

Let $\xi_0^\pm$ denote the two solutions of $g^{\mu\nu}\xi_\mu\xi_\nu = 0$ for the given $\xi_i$. Of course these solutions are real, so the (harmonic gauge) Einstein equation is weakly hyperbolic. We define the null covectors

$$\xi_\mu^\pm = (\xi_0^\pm, \xi_i). \tag{5.3.7}$$

These covectors will play an important role throughout this Chapter. By solving explicitly one finds that

$$\xi_0^+ + \xi_0^- = -2\frac{g^{0i}\xi_i}{g^{00}} \qquad \Rightarrow \qquad \xi^{0+} + \xi^{0-} = 0. \tag{5.3.8}$$

Hence we can adopt the convention $\xi^{0+} < 0$, $\xi^{0-} > 0$.[46]

We have $P_{\text{Einstein}}(\xi^\pm)t = 0$ for *any* $t_{ab}$. Hence for the Einstein equation, the matrix $M$ has two real eigenvalues $\xi_0^\pm$ and the associated eigenvectors are $(t, \xi_0^\pm t)^T$. Each eigenvalue has $N$ eigenvectors associated to it. It follows that $M$ has $2N$ linearly independent eigenvectors and hence $M$ is diagonalizable, as required by strong hyperbolicity.

We now return to the general case of Lovelock theory. Define a $2N \times 2N$ real symmetric (and hence hermitian) matrix $H(\xi_i)$ by

$$H(\xi_i) = \begin{pmatrix} B(\xi_i) & A \\ A & 0 \end{pmatrix}. \tag{5.3.9}$$

[46]We cannot have $\xi^{0\pm} = 0$ because that would violate the facts that $\xi_\mu^\pm$ is null and $e^0$ is timelike.

We then have

$$H(\xi_i)M(\xi_i)H(\xi_i)^{-1} = M(\xi_i)^T \tag{5.3.10}$$

so $M$ is real symmetric (and hence hermitian) w.r.t. $H$. It is easy to see that $H$ is non-degenerate: if $v = (t, t')^T$ then $Hv = 0$ implies $t = t' = 0$ using the fact that $A$ is invertible.[47] $H$ is hermitian and non-degenerate so its eigenvalues are real and non-zero. We can determine the signature of $H$ by writing the Lovelock couplings as

$$k_p = \epsilon \tilde{k}_p \qquad p \geq 2. \tag{5.3.11}$$

Since the eigenvalues of $H$ are real, non-vanishing, and depend continuously on $\epsilon$ (with $\tilde{k}_p$ and the background curvature fixed), the signature of $H$ cannot depend on $\epsilon$. Hence it can be evaluated at $\epsilon = 0$, i.e. for the linearised Einstein equation. The result is that $H$ has $N$ positive eigenvalues and $N$ negative eigenvalues, even for strong background fields. Thus, although $H$ and $M$ satisfy the condition (2.3.16), this does not imply strong hyperbolicity because $H$ is not positive definite.[48]

## 5.4 Proof of weak hyperbolicity in a low curvature background

To proceed, we will use a continuity argument involving the parameter $\epsilon$ defined in (5.3.11). Note that taking $\epsilon$ small at fixed $\tilde{k}_p$ and fixed background curvature is equivalent to assuming the background curvature to be small at fixed $k_p$. We will establish weak hyperbolicity for small $\epsilon$, which is equivalent to establishing it for small background curvature. In what follows we will suppress the dependence of $M$ and $H$ on $\xi_i$ and write simply $M(\epsilon)$ and $H(\epsilon)$.

For $\epsilon = 0$ we showed above that $\xi_0^\pm$ are the only eigenvalues of $M(\epsilon)$, each with degeneracy $N$. The eigenvalues of $M(\epsilon)$ depend continuously on $\epsilon$ [96]. Hence, for small $\epsilon$, they can be split unambiguously into two sets according to whether they approach $\xi_0^+$ or $\xi_0^-$ as $\epsilon \to 0$. We will follow [96] and refer to these sets as the $\xi_0^+$*-group* and the $\xi_0^-$*-group*. Each group contains $N$ eigenvalues.

Since we do not know whether or not the eigenvalues and eigenvectors of $M(\epsilon)$ are real, we will regard $M(\epsilon)$ and $H(\epsilon)$ as acting on a *complex* vector space $V$ of dimension $2N$.

[47] The matrix $H$ is closely related to the symplectic current density $\omega^\mu$ defined in [95]. Roughly speaking, $H$ is the high spatial frequency part of the Fourier space analogue of $-i\omega^0$.

[48] This is the case even for the Einstein equation ($\epsilon = 0$). However, for the Einstein equation we have shown that we can diagonalize $M$ so we can construct a positive definite matrix $K$ as explained above equation (2.2.19).

For $\epsilon = 0$, the eigenvalues $\xi_0^\pm$ are degenerate but "semi-simple", i.e., $M(0)$ is diagonalizable. However, there is no reason for this to remain true when $\epsilon \neq 0$: the Jordan canonical form of $M(\epsilon)$ may involve non-trivial Jordan blocks. For any eigenvalue $\xi_0$, one can define a *generalized eigenspace* as

$$\{v : \exists r \text{ such that } (M - \xi_0 I)^r v = 0\}\,. \tag{5.4.1}$$

This is the sum of the vector spaces associated with the Jordan blocks corresponding to that eigenvalue. We define the "total generalized eigenspace for the $\xi_0^\pm$-group" $V^\pm(\epsilon)$ as the sum over generalized eigenspaces of the eigenvalues in the $\xi_0^\pm$-group. Since any eigenvalue belongs to one of these groups we have

$$V = V^+(\epsilon) \oplus V^-(\epsilon). \tag{5.4.2}$$

We denote the projection onto $V^\pm(\epsilon)$ as $\Pi^\pm(\epsilon)$, i.e.,

$$V^\pm(\epsilon) = \Pi^\pm(\epsilon)V. \tag{5.4.3}$$

These projection matrices are holomorphic in $\epsilon$ for small $\epsilon$, in fact there is an explicit formula [96]

$$\Pi^\pm(\epsilon) = -\frac{1}{2\pi i}\int_{\Gamma^\pm}(M(\epsilon) - zI)^{-1}dz \tag{5.4.4}$$

where $\Gamma^\pm$ is a simple closed curve in the complex plane such that $\xi_0^\pm$ lies inside $\Gamma^\pm$ but $\xi_0^\mp$ lies outside $\Gamma^\pm$. Note that $\Gamma^\pm$ does not depend on $\epsilon$. For small non-zero $\epsilon$, the integrand has poles at the eigenvalues of $M(\epsilon)$ but only the eigenvalues that belong to the $\xi_0^\pm$-group lie inside $\Gamma^\pm$.

It can be shown that $M(\epsilon)$ and $H(\epsilon)$ satisfying (5.3.10) can be brought simultaneously to a block-diagonal canonical form, where $M(\epsilon)$ is in Jordan canonical form and $M(\epsilon)$ and $H(\epsilon)$ have the same block structure [97]. Since $V^+(\epsilon)$ and $V^-(\epsilon)$ contain different Jordan blocks of $M(\epsilon)$ it follows that these subspaces are orthogonal w.r.t. $H(\epsilon)$. Consider the restriction of $H(\epsilon)$ to these subspaces. Define the projection of $H(\epsilon)$ onto $V^\pm(\epsilon)$:

$$H^\pm(\epsilon) = \Pi^\pm(\epsilon)^\dagger H(\epsilon)\Pi^\pm(\epsilon). \tag{5.4.5}$$

This is a hermitian matrix which depends holomorphically on $\epsilon$. We will need to determine its signature. Any vector in $V^\mp(\epsilon)$ is an eigenvector with eigenvalue 0 hence $H^\pm(\epsilon)$ has at least $N$ vanishing eigenvalues. The remaining eigenvalues are associated to eigenvectors living in $V^\pm(\epsilon)$. Since the restriction of $H^\pm(\epsilon)$ to $V^\pm$ is the same as the restriction of $H(\epsilon)$ to $V^\pm$, it follows that this restriction is non-degenerate, i.e., these remaining eigenvalues are all non-zero. Therefore we can determine the signs of these eigenvalues by looking at the signs of the eigenvalues when $\epsilon = 0$, and using continuity. For $\epsilon = 0$, we know that $V^{\pm}(0)$ consists of vectors of the form $v = (t, \xi_0^{\pm} t)^T$. Taking the inner product of two such vectors w.r.t. $H^{\pm}(0)$ gives

$$\begin{aligned} v_1^\dagger H^{\pm}(0) v_2 &= t_1^\dagger B(0) t_2 + 2\xi_0^{\pm} t_1^\dagger A(0) t_2 \\ &= 2\xi_\mu^{\pm} t_1^\dagger P^{0\mu}_{\text{Einstein}} t_2 = -\xi^{0\pm} t_1^\dagger G t_2 \end{aligned} \tag{5.4.6}$$

where $G$ is defined in (5.2.13). Hence the signature of $H^{\pm}(0)$ restricted to $V^{\pm}(0)$ is the same as the signature of $-\xi^{0\pm} G$. Recall that $\xi^{0+} < 0$, $\xi^{0-} > 0$. It follows that within $V^{\pm}(0)$, $H^{\pm}(0)$ has the same signature as $\pm G$, i.e., $d$ negative eigenvalues and $d(d-1)/2$ positive eigenvalues for $H^{+}(0)$ and vice-versa for $H^{-}(0)$. Hence, by continuity, it follows that $H^{+}(\epsilon)$ has $d$ negative eigenvalues and $d(d-1)/2$ positive eigenvalues, with eigenvectors in $V^{+}(\epsilon)$, as well as $N = d(d+1)/2$ vanishing eigenvalues with eigenvectors in $V^{-}(\epsilon)$. Similarly for $H^{-}(\epsilon)$ with positive and negative interchanged.

We can identify an important subset of eigenvectors of $M(\epsilon)$ explicitly, for any $\epsilon$. They are associated to a residual gauge freedom. These "pure gauge" eigenvectors have $v$ of the form (5.3.4) with (cf. Def. 8, page 25)

$$\xi_0 = \xi_0^{\pm} \qquad t_{\mu\nu} = \xi^{\pm}_{(\mu} X_{\nu)} \tag{5.4.7}$$

for arbitrary complex $X_\mu$. Of course a pure gauge eigenvector with eigenvalue $\xi_0^{\pm}$ belongs to $V^{\pm}(\epsilon)$. It is interesting to calculate the inner product of two pure gauge eigenvectors so let $t'_{\mu\nu} = \xi^{\pm}_{(\mu} X'_{\nu)}$ and consider the associated vector $v'$ defined by (5.3.4). Since $v, v'$ are elements of $V^{\pm}(\epsilon)$, their inner product w.r.t. $H^{\pm}(\epsilon)$ is the same as their inner product w.r.t. $H(\epsilon)$:

$$\begin{aligned} v'^\dagger H(\epsilon) v &= t'^\dagger B(\epsilon) t + 2\xi_0^{\pm} t'^\dagger A(\epsilon) t \\ &= 2\xi_\mu^{\pm} t'^\dagger P^{0\mu}(\epsilon) t \\ &= 2\xi_\mu^{\pm} \xi_\nu^{\pm} \xi_\rho^{\pm} \bar{X}'_\sigma X_\tau P^{\nu\sigma\rho\tau 0\mu}(\epsilon) = 0 \end{aligned} \tag{5.4.8}$$

where in the final step we used the second equation in (3.2.17), and the fact that two such "pure gauge" vectors $t, t'$ are orthogonal w.r.t. $G^{\mu\nu\rho\sigma}$. This result shows that the pure gauge eigenvectors with eigenvalue $\xi_0^{\pm}$ define a $d$-dimensional subspace $N^{\pm}$ of $V^{\pm}(\epsilon)$ that is null w.r.t. $H^{\pm}(\epsilon)$.

We can now prove that the harmonic gauge linearised equation of motion of Lovelock theory is weakly hyperbolic in a small curvature background. Consider the possibility of an eigenvalue $\xi_0$ that is complex, with eigenvector $v$. For concreteness, assume that $\xi_0$ belongs to the $\xi_0^{+}$-group, so $v \in V^{+}(\epsilon)$. Equation (5.3.10) implies that a pair of eigenvectors whose eigenvalues are not complex conjugates of each other must be orthogonal w.r.t. $H(\epsilon)$. This implies that $v$ is orthogonal, w.r.t. $H^+(\epsilon)$, to the "pure gauge" eigenvectors in $V^+(\epsilon)$. Furthermore, since $\xi_0$ is complex, the $H(\epsilon)$-norm of $v$ must vanish, which implies that $v$ is null w.r.t. $H^+(\epsilon)$. The linear span of $v$ and $N^+$ now gives a $(d+1)$-dimensional subspace of $V^+(\epsilon)$ that is null w.r.t. $H^+(\epsilon)$. However, this is impossible because we showed above that for small $\epsilon$, $H^+(\epsilon)$ has $d$ negative eigenvalues and $d(d-1)/2$ positive eigenvalues which implies the maximal dimension of a null subspace of $V^+(\epsilon)$ is given by $\min(d, d(d-1)/2) = d$ [97]. This proves that complex $\xi_0$ is not possible for small $\epsilon$.

The final step is to note that the above argument assumed fixed $\xi_i$, i.e., for given $\xi_i$ then complex $\xi_0$ is not possible for small enough $\epsilon$. But we need our final result to be *uniform* in $\xi_i$, i.e., we need to show that the upper bound on $\epsilon$ does not depend on $\xi_i$. To do this we recall that our definition of weak hyperbolicity refers only to $\xi_i$ satisfying the condition $\xi_i\xi_i = 1$, i.e., $\xi_i$ belonging to a compact set. The spectrum of a matrix $M$ has uniformly continuous dependence on $M$ when $M$ is restricted to a bounded set [96]. It follows that the spectrum of $M(\epsilon)$ and $H(\epsilon)$ has uniformly continuous dependence on $\epsilon$ and $\xi_i$ when $\epsilon$ is restricted to a bounded set and $\xi_i\xi_i = 1$. Using this it can be shown that our results above are indeed uniform in $\xi_i$. The same argument establishes that our result is uniform in the spacetime point $x^\mu$ provided we restrict to a compact region of spacetime.

The above argument is restricted to a weakly curved background spacetime. If the curvature is not weak then the argument can fail. Imagine increasing $\epsilon$ to arbitrarily large values. There are two things that could go wrong. First, our assumption that $A$ is invertible may fail, i.e., we might reach a value of $\epsilon$ for which a surface of constant $t$ becomes characteristic somewhere. Second, it might not be possible to separate the eigenvalues into the $\xi_0^+$ group and the $\xi_0^-$ group as we did above. For example, as we increase $\epsilon$, an eigenvalue from one group might coincide with an eigenvalue from the other group. At larger $\epsilon$, this eigenvalue could then split into a pair of complex conjugate eigenvalues, violating weak hyperbolicity.

## 5.5 Failure of strong hyperbolicity in a generic low curvature background

For strong hyperbolicity, $M$ must be diagonalizable. We will now demonstrate that this is not the case for a *generic* weakly curved background spacetime.[49] We showed above that eigenvalues $\xi_0$ are all real in a weakly curved background. Therefore in

[49] In this section, we will not write explicitly the dependence on the parameter $\epsilon$ e.g. we write $M$ instead of $M(\epsilon)$.

this section we will assume that all vector spaces $V^\pm$, $N^\pm$, etc., are real. Note that the assumption that the background is weakly curved is required to define these spaces.

As discussed above, $M$ and $H$ satisfying (5.3.10) can be brought simultaneously via a change of basis to a certain canonical form [97]. We need to discuss this canonical form in more detail. In the canonical basis, $M$ has Jordan normal form and $H$ is block diagonal, with the same block structure as $M$. By this we mean that a $n \times n$ Jordan block in $M$ corresponds to a $n \times n$ block in $H$. Such a block of $H$ consists of zeros everywhere except on the diagonal running from top right to bottom left. Along this diagonal, the elements are all equal to 1 or all equal to $-1$. For example, if $M$ has a $3 \times 3$ Jordan block then the corresponding $3 \times 3$ block in $H$ has the form

$$\begin{pmatrix} 0 & 0 & \pm 1 \\ 0 & \pm 1 & 0 \\ \pm 1 & 0 & 0 \end{pmatrix}. \tag{5.5.1}$$

Each $n \times n$ block in $H$ is non-degenerate and has signature either $+1$ or $-1$ (if $n$ is odd) or 0 (if $n$ is even).

Recall the definition (5.4.1) of a generalized eigenspace. Note that a generalized eigenspace corresponds to a sum of all Jordan blocks associated to the given eigenvalue. Hence $V^\pm$ is a direct sum of the basis vectors associated to Jordan blocks of eigenvalues in the $\xi_0^\pm$-group. Hence any Jordan block is associated either to $V^+$ or to $V^-$. The canonical form then implies that $V^+$ and $V^-$ are orthogonal w.r.t. $H$, as stated above.

Let $E^\pm \subset V^\pm$ denote the generalized eigenspace of the eigenvalue $\xi_0^\pm$. We have shown that $N^\pm \subset E^\pm$. Hence, when restricted to $E^\pm$, $H^\pm$ must admit a $d$-dimensional null subspace. Consider $H^+$. From the canonical form we know that $H^+$ is non-degenerate when restricted to $E^+$. If this restriction has signature $(r, s)$ then the dimension of a maximal null subspace of $E^+$ is $\min(r, s)$ [97] hence we have $r, s \geq d$. However we already know that $H^+$ has signature $(d, d(d-1)/2)$ within $V^+$. The canonical form for $H$ shows that the signature is equal to the union of the signatures of each block. Therefore $H^+$ can have at most $d$ negative eigenvalues within $E^+$, i.e., we must have $r \leq d$. Combining these inequalities we see that $r = d$ and $s \geq d$. Hence $E^+$ has dimension $r + s \geq 2d$. Similarly $E^-$ has dimension at least $2d$.

A necessary condition for strong hyperbolicity is that $M$ is diagonalizable, i.e. there should be no non-trivial Jordan blocks. In other words, strong hyperbolicity requires that all generalized eigenspaces are simply eigenspaces. Hence if the theory is strongly hyperbolic then $E^\pm$ must be an eigenspace. Hence strong hyperbolicity requires that $M$ admits at least $2d$ eigenvectors with eigenvalue $\xi_0^\pm$. We already know that there are $d$ such eigenvectors in $N^\pm$. But for strong hyperbolicity there must exist at least an extra $d$ eigenvectors beyond these "pure gauge" ones. In terms of the principal symbol, this condition is that there exist at least $2d$ solutions $t_{ab}$ of $P(\xi^\pm)t = 0$ or equivalently (since $P_{\text{Einstein}}(\xi^\pm) = 0$) $\delta P(\xi^\pm)t = 0$. In other words $\ker \delta P(\xi^\pm)$ should have dimension at least $2d$. Furthermore, for strong hyperbolicity, this must be true for any $\xi_i$ and hence for any null $\xi^\pm$. In other words:

**Condition 1.** *A necessary condition for strong hyperbolicity is that, for any null* $\xi$, $\ker \delta P(\xi)$ *has dimension at least* $2d$.

There are certainly examples of background spacetimes for which this condition is satisfied. An extreme example is a flat background, for which $\delta P = 0$. In this case $M$ is diagonalizable and the equation of motion is strongly hyperbolic. A less trivial example is supplied by the class of Ricci flat spacetimes with Weyl tensor of type N, which are solutions of Lovelock theory with $\Lambda = 0$. In this case, the results of Ref. [70] imply that $M$ is diagonalizable so the equation of motion is strongly hyperbolic in such a background (even for large curvature). For this class of spacetimes, in addition to the pure gauge eigenvectors, generically there exist $d$ additional eigenvectors in $E^\pm$. This implies that $\ker \delta P(\xi^\pm)$ generically has dimension $2d$ for these spacetimes, in agreement with the above argument.

These background spacetimes are clearly very special because they have symmetries. In a *generic* weakly curved background, with null $\xi$, there is no reason to expect that $\ker \delta P(\xi)$ contains any non-gauge elements. To explain this, first note that if we are interested in non-gauge elements of $\ker \delta P(\xi^\pm)$ then we can regard $\delta P(\xi^\pm)$ as a map from the quotient space $V^\pm/N^\pm$, which has dimension $d(d-1)/2$, to the space of symmetric tensors which have vanishing contraction with $\xi^\pm$ (because of (3.2.17)). The latter space also has dimension $d(d-1)/2$. There is no reason to expect this map to have non-trivial kernel.

Perhaps we are overlooking some hidden symmetry of $\delta P$ that would guarantee that its kernel is larger than we expect. To exclude this possibility, we have calculated $\ker \delta P(\xi)$ for null $\xi$ in a generic background using computer algebra as follows.[50] We fix a point in spacetime and work at that point. Note that $\delta P$ is determined by the Riemann tensor of the background. For given null $\xi$ we can introduce a null basis for which $\xi$ is one of the basis vectors. In this basis, we can generate a random Riemann tensor satisfying the background equation of motion. To do this, we generate a random (small) Weyl tensor then use the background equation

[50] We use the `xTensor` and `xCoba` packages for `Wolfram Mathematica` [98].

of motion to determine the Ricci tensor and hence the Riemann tensor. Since the equation of motion is non-linear in curvature, there can be multiple solutions for the Ricci tensor but typically only one of these has small components, so this is the one we use. We then calculate $\ker \delta P(\xi)$ for this background Riemann tensor. The result is that, generically, this kernel has dimension $d$, i.e., it consists only of the "pure gauge" elements.

In summary, we have proven Condition 1 and we argued that it fails for a generic background, testing our claim numerically. Our argument suggests that $M$ is not diagonalizable for a generic weak field background. Therefore *the harmonic gauge linearised Lovelock equation of motion is not strongly hyperbolic in a generic weak field background.*

It is interesting to consider the canonical form of $M$ in more detail. Let's examine the condition for $M$ to have a $n \times n$ Jordan block with $n \geq 2$. From the canonical form, it is clear that the eigenvector associated to such a block must be null.[51] Assume that this eigenvector lives in $V^+$. If the eigenvalue is not $\xi_0^+$ then this eigenvector must be $H^+$-orthogonal to $N^+$, which implies that we could add this eigenvector to $N^+$ to construct a null subspace of dimension $d+1$, contradicting the fact that $N^+$ is a maximal null subspace. Hence the eigenvalue must be $\xi_0^+$. Similarly if the eigenvector lives in $V^-$ then the eigenvalue is $\xi_0^-$. We conclude that a non-trivial Jordan block must have eigenvalue $\xi_0^\pm$, so the basis vectors associated to the block must lie in $E^\pm$.

Any such Jordan block admits a vector $v \in E^\pm$ such that $(M - \xi_0^\pm)^2 \cdot v = 0$ but $(M - \xi_0^\pm) \cdot v \neq 0$ ($v$ is simply the second basis vector associated to the block) hence $(M - \xi_0^\pm) \cdot v$ is an eigenvector of $M$ with eigenvalue $\xi_0^\pm$. So we must have

$$\left(M - \xi_0^\pm\right) \cdot v = \begin{pmatrix} s \\ \xi_0^\pm s \end{pmatrix} \tag{5.5.2}$$

for some non-zero $s_{\mu\nu}$ such that (using $P_{\text{Einstein}}(\xi^\pm) = 0$)

$$\delta P(\xi^\pm) \cdot s = 0\,. \tag{5.5.3}$$

To examine whether such a block is possible, we need to determine whether (5.5.2) admits a solution $v$ for some $s_{\mu\nu} \neq 0$. If such a solution exists then $M$ is not diagonalizable.

Writing $v = (t, t')^T$ we find that (5.5.2) reduces to

$$t' = \xi_0^\pm t + s \tag{5.5.4}$$

[51] For example, for a $3 \times 3$ block, in the canonical basis, the eigenvector is $(1,0,0)^T$ and evaluating the norm of this using (5.5.1) gives 0.

and

$$\delta P(\xi^{\pm}) \cdot t = -(2\xi_0^{\pm} A + B) \cdot s. \tag{5.5.5}$$

The necessary and sufficient condition for this equation to admit a solution $t$ is for the RHS to have vanishing contraction with any element of $\ker \delta P(\xi^{\pm})$. We know this kernel always contains the "pure gauge" eigenvectors, i.e., it contains $N^{\pm}$. So contract with a "pure gauge" vector of the form $r_{\mu\nu} = \xi^{\pm}_{(\mu} Y_{\nu)}$. The LHS vanishes and we can rewrite the RHS in terms of $H$ to obtain

$$0 = \begin{pmatrix} r & \xi_0^{\pm} r \end{pmatrix} \cdot H \cdot \begin{pmatrix} s \\ \xi_0^{\pm} s \end{pmatrix}. \tag{5.5.6}$$

Hence $(s, \xi_0^{\pm} s)^T$ must be orthogonal (w.r.t. $H$) to all pure gauge eigenvectors in $E^{\pm}$, i.e., orthogonal to $N^{\pm}$. Furthermore, equation (5.5.3) shows that $s$ belongs to the kernel of $\delta P(\xi^{\pm})$ so we also need the contraction of $s$ with the RHS of (5.5.5) to vanish. This implies that $(s, \xi_0^{\pm} s)^T$ is null w.r.t. $H$. Therefore if this vector is not pure gauge, we could add it to $N^{\pm}$ to enlarge our null subspace, contradicting maximality of this null subspace. This proves that $s$ must be pure gauge, i.e.,

$$s_{\mu\nu} = \xi^{\pm}_{(\mu} X_{\nu)} \tag{5.5.7}$$

for some $X_\nu \neq 0$. Hence, *non-trivial Jordan blocks can arise only from pure gauge eigenvectors.* For $s_{\mu\nu}$ of this form, the RHS of (5.5.5) has vanishing contraction with any element of $N^{\pm}$.

We argued above that, in a generic weakly curved background, *all* elements of $\ker \delta P(\xi^{\pm})$ are "pure gauge", i.e., $\ker \delta P(\xi^{\pm}) = N^{\pm}$. It follows that in such a background, (5.5.5) can be solved for any pure gauge $s_{\mu\nu}$, i.e., *all* pure gauge eigenvectors belong to non-trivial Jordan blocks of $M$. So generically there are $d$ non-trivial Jordan blocks in each of $E^{\pm}$ and $M$ has $2d$ non-trivial blocks in total. In non-generic backgrounds, $\ker \delta P(\xi^{\pm})$ may contain non-gauge elements in which case $M$ may have fewer than $2d$ non-trivial blocks.

We have shown that, in a generic weak field background, every pure gauge eigenvector is associated to a $n \times n$ Jordan block of $M$ with $n \geq 2$. It is interesting to ask whether we could have $n \geq 3$. If $n \geq 3$ then there is a vector $v \in E^{\pm}$ such that $(M - \xi_0^{\pm})^3 v = 0$ with $(M - \xi_0^{\pm})^2 v \neq 0$. Let $(M - \xi_0^{\pm}) v \equiv (t, t')^T$, then $(t, t')$ must obey the equations (5.5.4), (5.5.5). Writing $v = (u, u')^T$ then gives

$$u' = \xi_0^{\pm} u + t \tag{5.5.8}$$

$$\delta P(\xi^{\pm}) \cdot u = -(2\xi_0^{\pm} A + B) \cdot t - A \cdot s. \tag{5.5.9}$$

As with (5.5.5), the necessary and sufficient condition for this equation to admit a solution is that the RHS has vanishing contraction with any element of $\ker \delta P(\xi^\pm)$. Generically we have $\ker \delta P(\xi^\pm) = N^\pm$ so we need the RHS to have vanishing contraction with any pure gauge vector $r_{\mu\nu} = \xi^\pm_{(\mu} Y_{\nu)}$. This contraction is just the $H$-inner product of $(t, t')$ with $(r, \xi_0^\pm r)$, so these vectors must be $H$-orthogonal for any pure gauge vector $r$. But there is no reason why this should be true. So generically we do not expect the above equations to admit a solution, i.e., the generic situation is $n = 2$.

To summarize: we have shown that, in a generic weak field background, every pure gauge eigenvector of $M$ belongs to a Jordan block of size $2 \times 2$.[52] Since non-trivial Jordan blocks can arise only from pure gauge eigenvectors, it follows that, generically, $V^\pm$ consists of $d$ $2 \times 2$ Jordan blocks, one for each pure gauge eigenvector, and $d(d-3)/2$ additional non-gauge eigenvectors. For a generic Ricci flat type N spacetime, it has been shown that these $d(d-3)/2$ additional eigenvectors have eigenvalues distinct from $\xi_0^\pm$ [70] so they do not belong to $E^\pm$ hence we expect this to be the behaviour in a generic spacetime. Therefore, generically, $E^\pm$ will have dimension $2d$.

Note that the $d(d-3)/2$ eigenvectors in $V^\pm$ that do not belong to $E^\pm$ can be regarded as the "physical graviton polarizations" [70]. To understand why, note that these eigenvectors have the form (5.3.4) where $t_{\mu\nu}$ satisfies the harmonic gauge condition. To prove the latter statement, simply contract the equation

$$P(\xi)^{\mu\nu\rho\sigma} t_{\rho\sigma} = 0 \tag{5.5.10}$$

with $\xi_\nu$ and use (3.2.17) to obtain

$$\xi^2 \left( \xi^\nu t_{\mu\nu} - \frac{1}{2} \xi_\mu t^\rho_\rho \right) = 0 \qquad \Rightarrow \qquad \xi^\nu t_{\mu\nu} - \frac{1}{2} \xi_\mu t^\rho_\rho = 0 \tag{5.5.11}$$

where we used the fact that $\xi^2 \neq 0$ because the eigenvector is not in $E^\pm$. Here the LHS is the "high frequency part" of the harmonic gauge condition. It is easy to check that the "pure gauge" eigenvectors in $N^\pm$ also satisfy this condition. However, there is no reason to expect that the vectors $t_{\mu\nu}$ obtained by solving (5.5.5) will satisfy this condition. Hence, generically, the $d$ "non-gauge" vectors in $E^\pm$ are associated to $t_{\mu\nu}$ which violate the harmonic gauge condition. So generically $E^\pm$ consists only of "pure gauge" and "gauge violating" vectors, which is why the $d(d-3)/2$ elements of $V^\pm$ that do not belong to $E^\pm$ can be regarded as the "physical polarizations".

[52] More precisely, this is true for a generic point and for generic $\xi_i$, in a generic weakly curved background.

## 5.6 Dynamical violation of weak hyperbolicity

We have shown that the linearised harmonic gauge equation of motion of Lovelock theory is not strongly hyperbolic in a *generic* weak curvature background. However, as mentioned above, it can be strongly hyperbolic in a non-generic weak curvature background. In this section, we will discuss a class of such backgrounds, namely homogeneous, isotropic, cosmological solutions of Lovelock theory. The aim is to demonstrate that weak (and hence also strong) hyperbolicity can be violated *dynamically*: there are "collapsing universe" solutions that start with small curvature but develop large curvature over time, in such a way that weak hyperbolicity is violated. Once this happens, local well-posedness of the equation of motion is lost, which implies that generic linear perturbations of the solution can no longer be evolved.

Lovelock theories admit FLRW-type solutions [99, 100]

$$g = -dt^2 + a(t)^2\gamma \tag{5.6.1}$$

where $\gamma$ is the metric of a $(d-1)$-dimensional submanifold of constant curvature $K$. We denote by $D$ the Levi-Civita connection associated to $\gamma$. The non-vanishing components of the Riemann tensor associated to $g$ are

$$R_{ij}{}^{kl} = \alpha(t)\delta_{ij}^{kl} \qquad R_{0i}{}^{0j} = \beta(t)\delta_i^j \tag{5.6.2}$$

where, in terms of the Hubble parameter $H = \dot{a}/a$,

$$\alpha = \frac{K}{2a^2} + H^2 \qquad \beta = H^2 + \dot{H}. \tag{5.6.3}$$

The non-vanishing components of the Lovelock tensor (3.2.1) are

$$A^0{}_0 = \sum_p k'_p\alpha^p \tag{5.6.4}$$

$$A^i{}_j = \delta^i_j \sum_p \frac{k'_p}{(d-1)}\alpha^{p-1}(2p\beta + (d-2p-1)\alpha) \tag{5.6.5}$$

where, for convenience, we have rescaled the coupling constants

$$k'_p = 2^p\frac{(d-1)!}{(d-2p-1)!}k_p \qquad k_0 = \Lambda, \quad k_1 = -1/4\,. \tag{5.6.6}$$

Taking our matter source to be a perfect fluid with equation of state $P = \omega\rho$, the equations of motion read

$$\sum_p k'_p\alpha^p = -\rho \tag{5.6.7}$$

$$\beta = -\frac{\sum_p k'_p\alpha^p\left[(d-1)(\omega+1) - 2p\right]}{\sum_p 2pk'_p\alpha^{p-1}}. \tag{5.6.8}$$

To observe how weak hyperbolicity can be violated dynamically in this setting, it is sufficient to look at the linearised equations for *transverse-traceless tensor* perturbations $g \to g + \delta g$:

$$\delta g_{0\mu} = 0 \quad \delta g_{ij} = 2a^2 h_{ij} \quad h_{ij} = h_{ji} \quad \gamma^{ij} h_{ij} = 0 \quad D^i h_{ij} = 0. \tag{5.6.9}$$

These are governed by the equation

$$- F_1(t)\ddot{h}_{ij} + F_2(t)a^{-2}(t)\, D_k D^k h_{ij} + \ldots = 0 \tag{5.6.10}$$

where the ellipsis denotes terms with fewer than 2 derivatives and we have defined

$$F_1(t) = \sum_p (d-3)pk'_p \alpha^{p-1} \tag{5.6.11}$$

$$F_2(t) = \sum_p pk'_p[2(p-1)\alpha^{p-2}\beta + (d-2p-1)\alpha^{p-1}]. \tag{5.6.12}$$

From this we can read off the principal symbol (restricted to tensor perturbations) and construct the matrices $A, B$ and $C$ described in Section 2.3

$$A^{ijkl} = -\gamma^{i(k}\gamma^{l)j} F_1(t) \tag{5.6.13}$$

$$B^{ijkl} = 0 \tag{5.6.14}$$

$$C^{ijkl} = \gamma^{i(k}\gamma^{l)j} a^{-2}(t)\gamma^{mn}\xi_m\xi_n F_2(t). \tag{5.6.15}$$

We can then compute the eigenvalues of $M$, or equivalently find the $\xi_0$ that solves $(\xi_0^2 A + C) \cdot t = 0$. For $F_1(t) \neq 0$ we find

$$\xi_0 = \tilde{\xi}_0^{\pm} \equiv \pm \frac{1}{a(t)} \sqrt{\gamma^{ij}\xi_i\xi_j \frac{F_2(t)}{F_1(t)}}\,. \tag{5.6.16}$$

Since $\gamma$ is a Riemannian metric (hence it is positive definite), the hyperbolicity of the theory is determined by the sign of $F_2(t)/F_1(t)$. If the background is weakly curved then the Einstein term dominates $F_1$ and $F_2$ and both of these quantities are negative so $\tilde{\xi}_0^{\pm}$ are real and the theory is weakly hyperbolic. However, if the curvature becomes large, e.g. in a collapsing universe solution, then one of these quantities might become positive, which makes $F_2/F_1$ negative so the theory is no longer weakly hyperbolic.

In agreement with the comments at the end of Section 5.4, we see that weak hyperbolicity can fail either when $F_1$ vanishes, i.e., the matrix $A$ becomes singular, or when $F_2$ vanishes, in which case an eigenvalue from the $\xi_0^+$ group becomes equal to an eigenvalue from the $\xi_0^-$ group, i.e., it is no longer possible to distinguish these two groups.

If $F_1$ or $F_2$ becomes positive then $\xi_0$ is imaginary and there exist linearised solutions which grow exponentially with time. For this reason, in the cosmology literature, a change in sign of $F_1$ or $F_2$ is usually referred to as an "instability" of the background solution. More specifically, if $F_1$ becomes negative then the background is said to suffer a "ghost instability" and if $F_2$ becomes negative it is said to suffer a "gradient instability".[53] However, this nomenclature is misleading. For the concept of stability to make sense, one needs the initial value problem for perturbations to be locally well-posed so that one can ask what happens when a generic initial perturbation is evolved in time. But when $F_1/F_2$ becomes negative then the equation for linear perturbations is not weakly hyperbolic which implies that the initial value problem is not well-posed: a generic linear perturbation cannot be evolved in time so dynamics no longer makes sense.

Further examples of dynamical violation of weak hyperbolicity can be obtained by considering the interior of a static, spherically symmetric black hole solution of a Lovelock theory [86, 99]. For a large black hole, the equations for linear perturbations are weakly hyperbolic outside the event horizon [70].[54] However, one can show that in the interior of such a black hole, the equations of motion fail to be weakly hyperbolic in a region $0 < r < r_*$. Here $r$ is the area-radius of the $(d-2)$-spheres, orbits of the symmetry group. Inside the black hole, surfaces of constant $r$ are spacelike and $-\partial/\partial r$ provides a time orientation. One can impose initial data for linear perturbations on a surface $r = r_0 > r_*$ inside the black hole. For large enough $r_0$, the curvature will be small on such a surface. Evolving this data then leads to a violation of weak hyperbolicity at time $r = r_*$. Generic linear perturbations cannot be evolved beyond this time.

[53]This behaviour was first discussed in the context of cosmological solutions of Horndeski theories [101, 102, 31].

[54]We expect them to be also strongly hyperbolic although we have not checked this.

# Chapter 6

# On the local well-posedness of Horndeski theories

In this Chapter we will discuss the initial value problem for Horndeski theories, using the machinery developed in Chapters 2 and 5. The contents of this Chapter are the result of original research. Sections 6.2 to 6.5 contain results obtained in collaboration with my supervisor, Harvey Reall, published in [44]. The rest of the Chapter consists of research conducted on my own, published in [45].

## 6.1 The initial value problem in Horndeski gravity

We begin by briefly introducing the initial value problem for Horndeski theories. Most of the issues with these theories are analogous to those encountered in the Lovelock case. We refer the reader to Section 5.1 for a more detailed discussion.

As discussed in Chapter 3, Horndeski theories are the most general four-dimensional diffeomorphism-covariant theories involving a metric tensor and a scalar field, with second order equations of motion [21].

Again, the diffeomorphism-covariant nature of the equations requires us to choose an appropriate gauge in order to study the initial value problem. Our arguments in the previous Chapter relied on considering Lovelock's equations as deformations of the Einstein equations. We will follow an analogous approach in this case. Given the presence of the scalar field, we will consider the Horndeski equations as a deformation of the *Einstein–scalar field* equations. Intuitively, we may need to deform the gauge condition to account for this deviation from Einstein–scalar field theory. While Einstein–scalar field theory is well-posed in harmonic gauge, since at the level of the principal symbol it decouples into the Einstein equation together with a scalar wave equation, the Horndeski contribution to the equations of motion introduces terms involving second derivatives of the scalar field which affect the structure of the principal symbol. One may expect that terms involving derivatives of the scalar field must be included in the gauge condition in order to cancel the problematic terms whilst retaining those that yield a strongly hyperbolic system. A natural way to achieve this is by considering a *generalised harmonic gauge* — i.e., by adding a source term on the RHS of the harmonic gauge condition

$$\Box x^\mu = J^\mu(g, \Phi, \partial\Phi). \tag{6.1.1}$$

Note that we are considering a more general version of what is usually referred to as "generalised harmonic gauge", as we are taking the source function to depend on the metric, the scalar field and its gradient.[55] The Einstein–scalar field equations are strongly hyperbolic in *any* (i.e., for any choice of $J^\mu$) generalised harmonic gauge and for any background. Horndeski theories, instead, exhibit pathologies similar to Lovelock theories. It has been shown, in fact, that cosmological solutions suffer from "ghost and gradient instabilities" when the background fields become large [101, 102, 31]. As in Lovelock gravity, these "instabilities" are not dynamical instabilities but rather indicate a failure of weak hyperbolicity in such backgrounds. We will therefore restrict our analysis to "weak background fields".

Recall that even restricting to "weak fields" does not guarantee a priori that the equations will be weakly hyperbolic (cf. Eq. 5.1.2). We will prove that Horndeski theories are weakly hyperbolic around a weak field background, in *any* generalised harmonic gauge.

To prove the well-posedness of the system it is necessary that the equations of motion be, not only weakly, but strongly hyperbolic. We will prove that for a particular class of Horndeski theories, namely $\partial_X \mathcal{G}_4 = \mathcal{G}_5 = 0$, there exists a generalised harmonic gauge for which the linearised equation of motion *is* strongly hyperbolic for arbitrary weak background fields. This class of theories involves no coupling between derivatives of the scalar field and curvature tensors in the action. This class includes various models of interest, e.g. "k-essence" theories or scalar-tensor theories such as Brans-Dicke theory [23]. However, for more general Horndeski theories, we find that the situation is analogous to the Lovelock case: there exists no generalised harmonic gauge for which the linearised theory is strongly hyperbolic in a generic weak field background. We will precede the proof of the failure of hyperbolicity in the most general case ($\mathcal{G}_5 = \mathcal{G}_5(\Phi, X) \neq 0$) by a detailed study of Einstein–dilaton–Gauss–Bonnet (EdGB) theory — a special case of the

[55] If we considered instead the standard generalised harmonic gauge, where $J^\mu = J^\mu(x)$, then its presence would not alter the structure of the principal symbol.

general Horndeski theory. This, besides being a theory of particular interest on its own, will provide a "pedagogical" introduction to the proof of the general result.

The results obtained can be strengthened considerably as follows. Consider a Horndeski theory for which there exists a generalised harmonic gauge such that the linearised equation of motion is strongly hyperbolic in a generic weak field background. We can now ask: does this extend to the non-linear theory? In particular, does there exist a generalised harmonic gauge for the non-linear theory such that the non-linear equation of motion is strongly hyperbolic in a generic weak-field background? For this to be the case, the generalised harmonic gauge condition for the non-linear theory must, upon linearisation, reduce to the generalised harmonic gauge condition for the linearised theory. However, this implies that the source function appearing in the gauge condition of the linearised theory must satisfy a certain integrability condition. This condition is not satisfied in general. Using this condition we find that the class of Horndeski theories for which there exists a generalised harmonic gauge for which the non-linear theory is strongly hyperbolic in a generic weak-field background is simply the class of "k-essence" type theories coupled to Einstein gravity (i.e. $\mathcal{G}_3 = \partial_X \mathcal{G}_4 = \mathcal{G}_5 = 0$).

The rest of the Chapter is organised as follows. In the next section we will derive the linearised Horndeski equations in generalised harmonic gauge. In Section 6.3 we will prove that the identities (6.3.5) — which played a crucial role in the proof of weak hyperbolicity for Lovelock theories — are a consequence of the gauge symmetry of the theory and that they hold for Horndeski theories as well. In Section 6.4 we establish that, in generalised harmonic gauge, all Horndeski theories are weakly hyperbolic around a weak field background. The remaining is devoted to the study of strong hyperbolicity: we introduce the set up in Section 6.5. Section 6.6 considers the $\mathcal{G}_5 = 0$ case; Section 6.7 deals with Einstein–dilaton–Gauss–Bonnet gravity, an example of the general theory considered in Section 6.8. We conclude the chapter by summarising the results and by showing how they can be strengthened by requiring the linearised gauge condition to arise as the linearisation of a corresponding non-linear condition.

## 6.2 Equations of motion in generalised harmonic gauge

We will begin our study of the hyperbolicity of Horndeski theories by deriving the generalised harmonic gauge equations of motion.

Recall from our discussion in Chapter 3 that the equations of motion for Horndeski theories can be obtained by varying the action (3.3.1):

$$E^{ab}[g,\Phi] \equiv -\frac{1}{\sqrt{-g}}\frac{\delta S}{\delta g_{ab}} = 0 \tag{6.2.1}$$

$$E_\Phi[g,\Phi] \equiv -\frac{1}{\sqrt{-g}}\frac{\delta S}{\delta \Phi} = 0. \tag{6.2.2}$$

We refer the reader to Eqs. (3.3.9) and (3.3.10) for the explicit form of the equations of motion. To study the hyperbolicity of these equations, we linearise around a background solution $(g,\Phi)$, i.e. we consider $(g+h,\Phi+\psi)$ and linearise in $h$ and $\psi$

$$E_{ab}[g+h,\Phi+\psi] = E_{ab}[g,\Phi] + E^{(1)}_{ab}[h,\psi] + \dots \tag{6.2.3}$$

$$E_\Phi[g+h,\Phi+\psi] = E_\Phi[g,\Phi] + E^{(1)}_{\Phi}[h,\psi] + \dots \tag{6.2.4}$$

so the linearised equations of motion are

$$E^{(1)}_{ab}[h,\psi] = E^{(1)}_{\Phi}[h,\psi] = 0. \tag{6.2.5}$$

Recall that the equations of motion resulting from the Einstein–scalar field theory are strongly hyperbolic if we impose the usual harmonic gauge condition which is[56]

$$G^{abcd}\nabla_b h_{cd} \equiv \nabla_b h^{ab} - \frac{1}{2}\nabla^a h^b{}_b = 0, \tag{6.2.6}$$

where $G^{abcd}$ is defined by (5.2.13). Motivated by this, we will attempt to obtain hyperbolic equations of motion for Horndeski theory by imposing a generalised harmonic gauge condition

$$H_a \equiv (1+f)G_a{}^{bcd}\nabla_b h_{cd} - \mathcal{H}_a{}^b\nabla_b\psi = 0, \tag{6.2.7}$$

where the scalar $f$ and the tensor $\mathcal{H}_a{}^b$ depend only on background quantities. The idea is that when we deform the theory away from the Einstein-scalar field theory we may need to deform the gauge condition in order to preserve hyperbolicity. The quantities $f$ and $\mathcal{H}$ describe such a deformation.[57] This gauge condition could be generalised further by including terms that do not involve derivatives of $h_{ab}$ or $\psi$. However such terms do not affect the principal symbol and therefore do not influence hyperbolicity.

To see that we can impose such a gauge condition, let $Y^a$ be a vector field and consider the infinitesimal diffeomorphism generated by $Y^a$:

$$h_{ab} \to h_{ab} + \nabla_{(a}Y_{b)} \qquad \psi \to \psi + Y\cdot\nabla\Phi. \tag{6.2.8}$$

[56]More properly we should call this a Lorenz gauge condition, but we will refer to it as a harmonic gauge condition for the reasons discussed below equation (5.2.5).

[57]Of course we could divide through by $(1+f)$ to absorb $f$ into $\mathcal{H}$. The reason for including $f$ here is that it leads to a more general class of gauge-fixed equations of motion when we perform the gauge-fixing procedure described below.

Under such transformation $H_a$ will change as

$$H_a \to H_a + \frac{1}{2}(1+f)(\nabla^b\nabla_b Y_a + R_{ab}Y^b) - \mathcal{H}_a{}^b\nabla_b(Y\cdot\nabla\Phi) \tag{6.2.9}$$

and can then be set to zero by choosing $Y_a$ to solve

$$\nabla^b\nabla_b Y_a - \frac{2}{1+f}\mathcal{H}_a{}^b\nabla_b(Y\cdot\nabla\Phi) + R_{ab}Y^b = -\frac{2}{1+f}H_a\,. \tag{6.2.10}$$

This is a linear wave equation of a standard type, which guarantees the existence of such $Y_a$. Note that if we changed the tensor structure of the first derivatives of $h_{ab}$ in (6.2.7) then this argument would no longer work.

To obtain the equations of motion in the generalised harmonic gauge, consider expanding the action to quadratic order in $(h,\psi)$ to obtain an action governing the linearised perturbation. Now to this action we add the gauge-fixing term[58]

$$S_{\text{gauge}} = -\frac{1}{2}\int\sqrt{-g}\,H_aH^a. \tag{6.2.11}$$

This will contribute to the equations of motion for the metric and the scalar field via terms

$$\frac{1}{\sqrt{-g}}\frac{\delta S_{\text{gauge}}}{\delta h_{ab}} = G^{abcd}\nabla_c((1+f)H_d) \tag{6.2.12}$$

$$\frac{1}{\sqrt{-g}}\frac{\delta S_{\text{gauge}}}{\delta\psi} = -\nabla_b(H^a\mathcal{H}_a{}^b) \tag{6.2.13}$$

respectively. We can now write the generalised harmonic gauge linearised equations as

$$\tilde{E}^{(1)}_{ab} = 0, \qquad \tilde{E}^{(1)}_{\Phi} = 0, \tag{6.2.14}$$

where

$$\tilde{E}^{(1)}_{ab} = E^{(1)}_{ab} - G_{ab}{}^{cd}\nabla_c((1+f)H_d) \tag{6.2.15}$$

$$\tilde{E}^{(1)}_{\Phi} = E^{(1)}_{\Phi} + \nabla_b(H^a\mathcal{H}_a{}^b). \tag{6.2.16}$$

It remains to show that the generalised harmonic gauge condition is propagated by the equations of motion. To see this, recall that the action for Horndeski is diffeomorphism invariant, thus for the non-linear theory we have

$$0 = \int d^4x\left(\frac{\delta S}{\delta g_{ab}}\nabla_a Y_b + \frac{\delta S}{\delta\Phi}Y^b\nabla_b\Phi\right) = \int d^4x\sqrt{-g}\left(\nabla^a E_{ab} - E_\Phi\nabla_b\Phi\right)Y^b. \tag{6.2.17}$$

This holds for arbitrary $Y^a$ hence, independent of any equation of motion,

$$\nabla^a E_{ab} - E_\Phi\nabla_b\Phi = 0 \tag{6.2.18}$$

[58]The reason for implementing the gauge-fixing this way is because obtaining the equation of motion from an action guarantees symmetry of the principal symbol, see section 6.3.

and so linearising around a background solution gives

$$\nabla^a E^{(1)}_{ab} - E^{(1)}_{\Phi} \nabla_b \Phi = 0. \tag{6.2.19}$$

Taking the divergence of (6.2.15) when (6.2.14) holds and using the above we obtain

$$\begin{aligned} 0 &= \nabla^a E^{(1)}_{ab} + G_{ab}{}^{cd} \nabla^b \nabla_c ((1+f) H_d) \\ &= E^{(1)}_{\Phi} \nabla_b \Phi - \frac{1}{2}(1+f)(\nabla^c \nabla_c H_b + R_{bc} H^c) - \nabla^b f \nabla_b H_a - \frac{1}{2} H_a \nabla^b \nabla_b f, \end{aligned} \tag{6.2.20}$$

that is

$$(1+f)\nabla_b \nabla^b H_a + 2\nabla^b f \nabla_b H_a + 2\nabla_c(\mathcal{H}^{cd} H_d)\nabla_a \Phi + (1+f) R_{ab} H^b + H_a \nabla^b \nabla_b f = 0. \tag{6.2.21}$$

This is a linear wave equation of a standard type for $H_a$, thus, provided that $H_a$ and its time derivative both vanish initially, they will continue to vanish throughout the evolution, i.e. the gauge condition (6.2.7) is propagated by the equations of motion (6.2.14). It then follows that a solution of the generalised harmonic gauge equations (6.2.14) is also a solution of the original linearised Horndeski equations of motion (6.2.5).

We will now compute the principal symbol in generalised harmonic gauge. The procedure is analogous to the one employed in Section 3.3, with the difference that now we will have contributions arising from the gauge fixing terms in the equations of motion. The linearised generalised harmonic gauge equations of motion (6.2.14) take the following form

$$P_{gg}^{abcdef} \nabla_e \nabla_f h_{cd} + P_{g\Phi}^{abef} \nabla_e \nabla_f \psi + \ldots = 0, \tag{6.2.22}$$

$$P_{\Phi g}^{cdef} \nabla_e \nabla_f h_{cd} + P_{\Phi\Phi}^{ef} \nabla_e \nabla_f \psi + \ldots = 0, \tag{6.2.23}$$

where the ellipses denotes terms with fewer than 2 derivatives. We can then define the principal symbol for this system

$$P(\xi) = \begin{pmatrix} P_{gg}^{abcdef} \xi_e \xi_f & P_{g\Phi}^{abef} \xi_e \xi_f \\ P_{\Phi g}^{cdef} \xi_e \xi_f & P_{\Phi\Phi}^{ef} \xi_e \xi_f \end{pmatrix} \tag{6.2.24}$$

and we think of it as acting on vectors of the form $(t_{cd}, \alpha)^T$, where $t_{cd}$ is a symmetric 2-tensor and $\alpha$ is a number.

It is convenient to split the principal symbol in its Einstein-scalar field and Horndeski parts

$$P(\xi) = P_{\mathrm{Esf}}(\xi) + \delta P(\xi) \tag{6.2.25}$$

where

$$P_{\mathrm{Esf}}(\xi) = \begin{pmatrix} -\frac{1}{2}\xi^2 G^{abcd} & 0 \\ 0 & -\xi^2 \end{pmatrix} \tag{6.2.26}$$

is the principal symbol for the generalised harmonic gauge Einstein-scalar field equations of motion. We write

$$\delta P(\xi) = \delta\tilde{P}(\xi) + \delta Q(\xi) \tag{6.2.27}$$

where $\delta\tilde{P}$ denotes the terms arising from the Horndeski terms, already computed in Section 3.3 (cf. Eqs. (3.3.16)); $\delta Q$ denotes the $f$- and $\mathcal{H}$-dependent parts of the gauge-fixing terms. Explicitly we have

$$\delta Q(\xi) = \begin{pmatrix} -f(f+2)G^{abeh}G_h{}^{fcd}\xi_e\xi_f & (1+f)\xi^e G^{fhab}\xi_h\mathcal{H}_{ef} \\ (1+f)\xi^e G^{fhcd}\xi_h\mathcal{H}_{ef} & -\mathcal{H}_h{}^e\mathcal{H}^{hf}\xi_e\xi_f \end{pmatrix}. \tag{6.2.28}$$

From the form of $P_{\text{Esf}}$ it is clear that all characteristics of the harmonic gauge Einstein-scalar field system are null.

### Weak background fields

We conclude this section by making precise the notion of "weak background fields" in the Horndeski setting. We follow a similar approach to the one used for Lovelock theories (cf. Section 5.3). Consider an orthonormal basis $\{e_\mu\}$, such that $e_0$ is orthogonal to constant $t$ surfaces. Denote by $L_R^{-2}$, $L_1^{-1}$ and $L_2^{-2}$ the magnitude of the largest components in such a basis of the Riemann tensor, $\nabla\Phi$ and $\nabla\nabla\Phi$ respectively and define

$$L^{-2} = \max\{L_R^{-2}, L_1^{-2}, L_2^{-2}\}. \tag{6.2.29}$$

We want our definition of "weak fields" to ensure that the Horndeski terms in the principal symbol be small compared to the Einstein–scalar field terms, i.e., $\delta P$ (Eq. (3.3.16)) must be small compared to $P_{\text{Esf}}$. This is achieved by requiring the background fields to satisfy[59]

$$|\partial_X^k(X\partial_X\mathcal{G}_5)|\,L^{-2(1+k)} \ll 1, \qquad k=0,1,2 \tag{6.2.30a}$$

$$|\partial_X^k(\partial_X\mathcal{G}_4 - \partial_\Phi\mathcal{G}_5)|\,L^{-2(1+k)} \ll 1, \qquad k=0,1,2 \tag{6.2.30b}$$

$$|\partial_X^k(\mathcal{G}_4 - 2X\partial_X\mathcal{G}_4 + X\partial_\Phi\mathcal{G}_5)|\,L^{-2k} \ll 1, \qquad k=0,1,2 \tag{6.2.30c}$$

$$|\partial_X^k(X\partial_X\mathcal{G}_3 + \partial_\Phi\mathcal{G}_4 + 2X\partial^2_{X\Phi}\mathcal{G}_4)|\,L^{-2k} \ll 1, \qquad k=0,1 \tag{6.2.30d}$$

$$|\partial_X^k(\partial_X\mathcal{G}_3 + 4\partial^2_{X\Phi}\mathcal{G}_4 - \partial^2_\Phi\mathcal{G}_5)|\,L^{-2(1+k)} \ll 1, \qquad k=0,1 \tag{6.2.30e}$$

$$|\partial_X\mathcal{G}_2 + 2X\partial_X^2\mathcal{G}_2 + 2\partial_\Phi\mathcal{G}_3 + 2X\partial^2_{X\Phi}\mathcal{G}_3| \ll 1, \tag{6.2.30f}$$

$$|\partial_X^2\mathcal{G}_2 + 2\partial^2_{X\Phi}\mathcal{G}_3 + 2\partial^3_{X\Phi\Phi}\mathcal{G}_4|\,L^{-2} \ll 1. \tag{6.2.30g}$$

[59]Note that these conditions are weaker than those expressed in [44]. If the conditions in [44] hold, then these hold as well.

We will also require smallness of the functions appearing in the gauge condition:

$$|f| \ll 1, \qquad |\mathcal{H}_\mu{}^\nu| \ll 1. \tag{6.2.31}$$

In practice, we will see that strong hyperbolicity will force us to take $f$ and $\mathcal{H}_a{}^b$ to be particular functions of the background fields, and (6.2.31) then follows from weakness of the background fields.

## 6.3 Symmetries of the principal symbol

For Lovelock theories, our argument for weak hyperbolicity exploited equations (3.2.17) following from the identities (3.2.16). Therefore we will need to determine the analogous identities for Horndeski theories. This could be done by explicit computation. Instead we will derive the identities as a consequence of the gauge symmetry of the theory. We will appeal to results of Lee and Wald [95] to do this.

Consider some diffeomorphism covariant theory of gravity, possibly coupled to additional fields, and expand the action to second order around a background solution:

$$S = \int d^d x \sqrt{-g} \left( -\frac{1}{2} K^{IJab} \nabla_a u_I \nabla_b u_J + \dots \right), \tag{6.3.1}$$

where $u_I$ denotes the perturbation to the fields (including the metric perturbation), the ellipsis denotes terms with fewer than two derivatives, and

$$K^{IJab}(x) = K^{JIba}(x). \tag{6.3.2}$$

Varying the action gives the (linearised) equation of motion

$$K^{IJab} \nabla_a \nabla_b u_J + \dots = 0, \tag{6.3.3}$$

where the ellipsis denotes terms with fewer than two derivatives of $u_I$. From this we read off the principal symbol

$$P^{IJab} = K^{IJ(ab)}, \tag{6.3.4}$$

so from (6.3.2) we have

$$P^{IJab} = P^{JIab}. \tag{6.3.5}$$

Hence, the symmetry of the principal symbol is a consequence of the variational principle. Varying the action also gives a total derivative term $\nabla_a \theta^a$, where

$$\theta^a = -K^{IJab} \delta u_I \nabla_b u_J + \dots \tag{6.3.6}$$

where the ellipsis denotes terms without derivatives. We then define the *symplectic current* for two independent variations $\delta_1 u_I$ and $\delta_2 u_I$ [95]

$$\omega^a = \delta_1 \theta_2^a - \delta_2 \theta_1^a = K^{IJab} \delta_1 u_I \nabla_b \delta_2 u_J - (1 \leftrightarrow 2) + \dots \tag{6.3.7}$$

Given coordinates $(t, x^i)$ where $t$ is a time function, we define the symplectic form as an integral over a surface $\Sigma$ of constant $t$ with unit normal $n_a$

$$\omega(\delta_1 u, \delta_2 u) = \int_\Sigma \omega^\mu n_\mu = \int_\Sigma d^{d-1}x\sqrt{-g}\,\omega^0. \tag{6.3.8}$$

For a theory with a gauge symmetry, Ref. [95] proves that this vanishes if $\delta_2 u$ is taken to be an infinitesimal gauge transformation and $\delta_1 u$ satisfies the (linearised) equation of motion. In particular, it will vanish if $\delta_1 u$ and $\delta_2 u$ are *both* infinitesimal gauge transformations. Taking them to be compactly supported gauge transformations we can integrate w.r.t. $t$ to obtain

$$0 = \int d^d x\sqrt{-g}\left[K^{IJ0\nu}\delta_1 u_I \nabla_\nu \delta_2 u_J - (1 \leftrightarrow 2) + \ldots\right]. \tag{6.3.9}$$

As before, the ellipsis denotes terms without derivatives of $\delta_1 u$ or $\delta_2 u$.

Consider first the case of Lovelock theory (without any gauge-fixing), for which $u_I = h_{ab}$ and we have the symmetries

$$K^{abcdef} = K^{bacdef} = K^{abdcef}. \tag{6.3.10}$$

The gauge transformations are infinitesimal diffeomorphisms:

$$\delta h_{ab} = \nabla_{(a}X_{b)}, \tag{6.3.11}$$

where $X_a$ is an arbitrary vector field, assumed compactly supported. Gauge invariance of the action implies, via integration by parts,

$$0 = \int d^d x\sqrt{-g}X_b\left(-K^{abcdef}\nabla_a\nabla_e\nabla_f h_{cd} + \ldots\right) \tag{6.3.12}$$

where the ellipsis denotes terms with fewer than 3 derivatives of $h_{\mu\nu}$. Since $X_a$ is arbitrary, this implies

$$0 = K^{abcdef}\nabla_a\nabla_e\nabla_f h_{cd} + \ldots \tag{6.3.13}$$

and since $h_{ab}$ is arbitrary, terms with different numbers of derivatives must vanish independently. From the 3-derivative term we obtain

$$0 = K^{(a|bcd|ef)}, \tag{6.3.14}$$

which implies

$$P^{(a|bcd|ef)} = 0. \tag{6.3.15}$$

Now we consider the implications of (6.3.9). Take the two gauge transformations to be

$$\delta_1 h_{\mu\nu} = \nabla_{(\mu}X_{\nu)}, \qquad \delta_2 h_{\mu\nu} = \nabla_{(\mu}Y_{\nu)}, \tag{6.3.16}$$

for arbitrary compactly supported vector fields $X^\mu$, $Y^\mu$. Compact support lets us integrate by parts in (6.3.9):

$$\begin{aligned} 0 &= \int d^dx\sqrt{-g}\left[\nabla_\mu X_\nu K^{\mu\nu\rho\sigma 0\alpha}\nabla_\alpha\nabla_\rho Y_\sigma - (1\leftrightarrow 2) + \ldots\right] \\ &= \int d^dx\sqrt{-g}X_\nu\left[-K^{\mu\nu\rho\sigma 0\alpha}\nabla_\mu\nabla_\alpha\nabla_\rho Y_\sigma - K^{\mu\sigma\rho\nu 0\alpha}\nabla_\alpha\nabla_\rho\nabla_\mu Y_\sigma + \ldots\right], \end{aligned} \quad (6.3.17)$$

where the ellipsis denotes terms with fewer than 3 derivatives of $Y^\mu$. Since $X_\nu$ is arbitrary we must have

$$\begin{aligned} 0 &= K^{\mu\nu\rho\sigma 0\alpha}\nabla_\mu\nabla_\alpha\nabla_\rho Y_\sigma + K^{\mu\sigma\rho\nu 0\alpha}\nabla_\alpha\nabla_\rho\nabla_\mu Y_\sigma + \ldots \\ &= \left(K^{\mu\nu\rho\sigma 0\alpha} + K^{\mu\sigma\rho\nu 0\alpha}\right)\partial_\mu\partial_\rho\partial_\alpha Y_\sigma + \ldots \\ &= \left(K^{\mu\nu\rho\sigma 0\alpha} + K^{\rho\nu\mu\sigma\alpha 0}\right)\partial_\mu\partial_\rho\partial_\alpha Y_\sigma + \ldots \end{aligned} \quad (6.3.18)$$

Since $Y_\mu$ is arbitrary, the terms with different numbers of derivatives of $Y_\mu$ must vanish independently. Vanishing of the 3-derivative term requires

$$0 = K^{\nu(\mu\rho|\sigma 0|\alpha)} + K^{\nu(\rho\mu|\sigma|\alpha)0} = 2P^{\nu(\mu\rho|\sigma 0|\alpha)}. \quad (6.3.19)$$

Since the 0 index refers to an *arbitrary* time function $t$, this equation implies

$$P^{a(bc|de|f)} = 0. \quad (6.3.20)$$

The above argument applies to the theory *before* fixing the gauge. Of course we can do the same for the Einstein equation. Subtracting the Einstein results from the Lovelock results gives

$$\delta P^{(a|bcd|ef)} = \delta P^{a(bc|de|f)} = 0. \quad (6.3.21)$$

We can now apply this to the harmonic gauge Lovelock equation of motion because the harmonic gauge condition does not affect $\delta P$. In particular we have

$$\delta P^{abcdef}\xi_a\xi_e\xi_f = \delta P^{abcdef}\xi_b\xi_c\xi_f = 0. \quad (6.3.22)$$

Hence we see that the identities (3.2.17) are a consequence of the gauge symmetry.

For a Horndeski theory (before any gauge fixing) we have $u_I = (h_{ab}, \psi)$. A gauge transformation is

$$\delta h_{ab} = \nabla_{(a}X_{b)} \qquad \delta\psi = X^a\nabla_a\Phi. \quad (6.3.23)$$

Repeating the above argument for gauge invariance of the action gives

$$P_{gg}^{(a|bcd|ef)} = P_{g\Phi}^{(a|b|cd)} = 0. \quad (6.3.24)$$

The symmetry of the principal symbol (6.3.5) then implies that

$$P_{\Phi g}^{(a|b|cd)} = P_{g\Phi}^{(a|b|cd)} = 0. \quad (6.3.25)$$

Repeating the argument based on (6.3.9), the highest derivatives of the gauge transformation parameters $X_\mu$ and $Y_\mu$ arise only from the transformation of $h_{\mu\nu}$ so the result is essentially the same as for Lovelock theory:

$$P_{gg}^{a(bc|de|f)} = 0. \tag{6.3.26}$$

These results apply also to the Einstein-scalar field theory (before gauge fixing). So subtracting the principal symbols for these two cases gives

$$0 = \delta\tilde{P}_{gg}^{(a|bcd|ef)} = \delta\tilde{P}_{g\Phi}^{(a|b|cd)} = \delta\tilde{P}_{\Phi g}^{(a|b|cd)} = \delta\tilde{P}_{gg}^{a(bc|de|f)}. \tag{6.3.27}$$

Finally, we note that the gauge fixing terms do not affect $\delta\tilde{P}$ so these results apply also to the generalised harmonic gauge equation of motion.

## 6.4 Weak hyperbolicity for weak field background

We will now begin our study of the hyperbolicity of the linearised Horndeski equations in a generalised harmonic gauge. In this Section we will establish weak hyperbolicity of these equations in a weak field background for any generalised harmonic gauge. Much of the analysis is similar to the analysis of the weak hyperbolicity of harmonic gauge Lovelock theories performed above so we will be briefer here.

As in Section 5.3 we introduce coordinates $x^\mu = (t, x^i)$ such that $dt$ is timelike so surfaces of constant $t$ are non-characteristic for the Einstein–scalar field theory. Again we will denote by $\xi_0^\pm$ the two solutions of $g^{\mu\nu}\xi_\mu\xi_\nu = 0$ for fixed real $\xi_i$, and we define the null covectors $\xi_\mu^\pm = (\xi_0^\pm, \xi_i)$.

The principal symbol can be regarded as a quadratic form acting on vectors of the form $(t_{\mu\nu}, \chi)^T$, with $t_{\mu\nu}$ symmetric. Such vectors form an 11-dimensional space. Hence $A$, $B(\xi_i)$ and $C(\xi_i)$ (defined in Sec. 5.3) are $11 \times 11$ matrices. Explicitly we have

$$A = \begin{pmatrix} P_{gg}^{\mu\nu\rho\sigma 00} & P_{g\Phi}^{\mu\nu 00} \\ P_{\Phi g}^{\rho\sigma 00} & P_{\Phi\Phi}^{00} \end{pmatrix} \qquad B(\xi_i) = \begin{pmatrix} 2P_{gg}^{\mu\nu\rho\sigma(0i)}\xi_i & 2P_{g\Phi}^{\mu\nu(0i)}\xi_i \\ 2P_{\Phi g}^{\rho\sigma(0i)}\xi_i & 2P_{\Phi\Phi}^{(0i)}\xi_i \end{pmatrix}$$

$$C(\xi_i) = \begin{pmatrix} P_{gg}^{\mu\nu\rho\sigma ij}\xi_i\xi_j & P_{g\Phi}^{\mu\nu ij}\xi_i\xi_j \\ P_{\Phi g}^{\rho\sigma ij}\xi_i\xi_j & P_{\Phi\Phi}^{ij}\xi_i\xi_j \end{pmatrix} \tag{6.4.1}$$

where, again, $\xi_i$ is real and $\xi_i\xi_i = 1$. These matrices are all real and symmetric: the latter property follows from the fact that the gauge-fixed equations of motion can be derived from an action so (6.3.5) holds.

For the harmonic gauge Einstein–scalar field equations, since surfaces of constant $t$ are spacelike, the matrix $A$ is invertible. By continuity, this will continue to hold for sufficiently weak background fields, once we include the Horndeski terms. Hence we can define real $M(\xi_i)$ as in (5.3.3) and real symmetric $H(\xi_i)$ as in (5.3.9). These are $22 \times 22$ matrices. As for Lovelock, the matrix $H$ is non-degenerate so its signature can be determined by continuity, i.e., by its signature for the Einstein-scalar field equations. The result is that it has signature $(11, 11)$, i.e., 11 positive eigenvalues and 11 negative eigenvalues. As for Lovelock, $M$ is symmetric w.r.t. $H$, i.e., equation (5.3.10) holds here.

We consider these matrices as acting on a complex vector space $V$ of dimension 22. For the Einstein-scalar field theory we know that $M$ is diagonalizable with eigenvalues $\xi_0^{\pm}$, each with degeneracy 11. So, for linearised Horndeski theory in a weak field background we can proceed as in Sec. 5.4 and define the 11-dimensional subspaces $V^{\pm}$ as the sum over the generalised eigenspaces of the eigenvectors (of $M$) belonging to the $\xi_0^{\pm}$-group, respectively. The restriction of $H$ to $V^{\pm}$ is denoted by $H^{\pm}$.

Let us summarize the proof of weak hyperbolicity that we used for Lovelock theories. First we showed that there exist "pure gauge" eigenvectors of $M$, with eigenvalue $\xi_0^{\pm}$. We then showed that such eigenvectors are null and orthogonal w.r.t. $H$ so they form null subspaces $N^{\pm}$ of $V^{\pm}$, and that these null subspaces have the maximum dimension consistent with the signature of $H^{\pm}$. This then excludes the possibility of $M$ possessing a complex eigenvalue $\xi_0$ in, say, the $\xi_0^{+}$-group, for the corresponding eigenvector would have to be null and orthogonal to $N^{+}$ so we could add it to $N^{+}$ to produce a larger null subspace of $V^{+}$, thereby violating maximality of $N^{+}$. Hence $M$ cannot have a complex eigenvalue, which establishes weak hyperbolicity.

All of this extends straightforwardly to Horndeski theories. First note that, as in section 2.3, an eigenvector $v$ of $M$ with eigenvalue $\xi_0$ must have the form

$$v = \begin{pmatrix} T \\ \xi_0 T \end{pmatrix}, \tag{6.4.2}$$

where the 11-vector $T$ must satisfy

$$P(\xi) \cdot T = 0 \tag{6.4.3}$$

with $\xi_\mu = (\xi_0, \xi_i)$. We can identify a set of "pure gauge" eigenvectors, with eigenvalue $\xi_0^{\pm}$, given by[60]

$$T = \begin{pmatrix} \xi^{\pm}_{(\mu} X_{\nu)} \\ 0 \end{pmatrix} \tag{6.4.4}$$

[60]The vanishing of the final component of this vector is related to the fact that under the gauge transformation (6.2.8), the transformation of $\psi$ does not involve a derivative of $Y_a$.

for some $X_\mu$. That this satisfies (6.4.3) (with $\xi = \xi^\pm$) can be seen as follows. First $P_{\rm Esf}(\xi^\pm) = 0$ because $\xi^\pm_\mu$ is null. Second, the results in (6.3.27) imply

$$\delta\tilde{P}(\xi^\pm)\cdot T = 0. \tag{6.4.5}$$

Finally, it can be checked explicitly that $\delta Q(\xi^\pm)\cdot T = 0$.

We define $N^\pm$ to be the 4-dimensional subspace of $V^\pm$ defined by these pure gauge eigenvectors. We now want to prove that $N^\pm$ is null w.r.t. $H^\pm$. Consider two pure gauge eigenvectors $v, v' \in N^\pm$ with corresponding $T = (\xi^\pm_{(\mu}X_{\nu)}, 0)^T$ and $T' = (\xi^\pm_{(\mu}X'_{\nu)}, 0)^T$. Their inner product w.r.t. $H^\pm$ is the same as their inner product w.r.t. $H$, i.e., as in (5.4.8), we have

$$v'^\dagger H v = 2\xi^\pm_\mu T'^\dagger P^{0\mu} T = 2\xi^\pm_\mu \xi^\pm_\nu \xi^\pm_\rho \bar{X}'_\sigma X_\tau P_{gg}^{\nu\sigma\rho\tau 0\mu} = 0, \tag{6.4.6}$$

where the final equality follows from $P_{\rm Esf}(\xi^\pm) = 0$, the final symmetry in (6.3.27), and the fact that

$$\xi^\pm_\mu \xi^\pm_\nu \xi^\pm_\rho \delta Q_{gg}^{\nu\sigma\rho\tau\lambda\mu} = 0. \tag{6.4.7}$$

It follows that any two elements of $N^\pm$ are orthogonal w.r.t. $H^\pm$ hence $N^\pm$ defines a 4-dimensional $H^\pm$-null subspace of $V^\pm$.

Since $H^\pm$ is the restriction of $H$ to $V^\pm$, it follows that $H^\pm$ is non-degenerate when restricted to $V^\pm$. Hence its signature can be determined by continuity, as we did for Lovelock. In other words, its signature can be determined using the Einstein–scalar field theory. For this theory, consider two vectors $v_1$ and $v_2$ in $V^\pm$, and hence of the form (6.4.2) with $\xi_0 = \xi_0^\pm$. Let the corresponding 11-vectors be $T_1 = (t_{1ab}, \chi_1)^T$ and $T_2 = (t_{2ab}, \chi_2)^T$. The inner product of $v_1$ and $v_2$ w.r.t. $H^\pm$ is the same as the inner product w.r.t. $H$:

$$v_1^\dagger H v_2 = T_1^\dagger B T_2 + 2\xi_0^\pm T_1^\dagger A T_2 = 2\xi_\mu T_1^\dagger P^{0\mu} T_2 = -\xi^{0\pm}\left(t_1^\dagger G t_2 + \bar{\chi}_1\chi_2\right). \tag{6.4.8}$$

The argument following (5.4.6) now shows that, when restricted to $V^+$, $H^+$ has 4 negative eigenvalues and $6+1=7$ positive eigenvalues (the $+1$ coming from $\bar{\chi}_1\chi_2$). Similarly for $H^-$ when restricted to $V^-$, with positive and negative interchanged. Hence the dimension of a maximal null subspace of $V^\pm$ is 4 so $N^\pm$ are maximal null subspaces of $V^\pm$. The proof of weak hyperbolicity follows as explained above.

## 6.5 Strong hyperbolicity: setting up the problem

We have shown that, in any generalised harmonic gauge, linearised Horndeski theory is weakly hyperbolic in a weak field background. We will now investigate whether it is also strongly hyperbolic. In particular, strong hyperbolicity requires that $M$ be diagonalizable, i.e., it has no non-trivial Jordan blocks. We can investigate whether or not this is true using the method of Section 5.5.

As in Section 5.5 we define $E^{\pm}$ to be the generalised eigenspace of the eigenvalue $\xi_0^{\pm}$. Since $N^{\pm} \subset E^{\pm}$ it follows as in Section 5.5 that $E^{\pm}$ must have dimension at least 8. If $M$ is diagonalizable, then $E^{\pm}$ are genuine eigenspaces and hence there must exist at least 8 eigenvectors with eigenvalue $\xi_0^{\pm}$. So using (6.4.3) and $P_{\text{Esf}}(\xi^{\pm}) = 0$ we must have 8 vectors $T$ satisfying $\delta P(\xi^{\pm}) \cdot T = 0$. So we have the following

**Condition 2.** *A necessary condition for strong hyperbolicity is that, for any null $\xi$, $\ker \delta P(\xi)$ has dimension at least 8.*

Hence, strong hyperbolicity implies that, for any null $\xi$, $\ker \delta P(\xi)$ must contain at least 4 linearly independent "non-gauge" elements.

Let us now look at the condition for a non-trivial Jordan block to exist. As in Section 5.5, one can show that the corresponding eigenvalue must be $\xi_0^{\pm}$ so the block must lie in $E^{\pm}$. For any such block, there exists a vector $v \in E^{\pm}$ such that $(M - \xi_0^{\pm})v$ is an eigenvector of $M$ with eigenvalue $\xi_0^{\pm}$, so we must have

$$\left(M - \xi_0^{\pm}\right) \cdot v = \begin{pmatrix} S \\ \xi_0^{\pm} S \end{pmatrix} \tag{6.5.1}$$

for some non-zero $S = (s_{\mu\nu}, \omega)^T$ such that (using $P_{\text{Esf}}(\xi^{\pm}) = 0$)

$$\delta P(\xi^{\pm}) \cdot S = 0. \tag{6.5.2}$$

Writing $v = (T, T')^T$ we find that (6.5.1) reduces to equations analogous to (5.5.4) and (5.5.5):

$$T' = \xi_0^{\pm} T + S \tag{6.5.3}$$

and

$$\delta P(\xi^{\pm}) \cdot T = -(2\xi_0^{\pm} A + B) \cdot S. \tag{6.5.4}$$

As in Section 5.5 we contract this with an arbitrary "pure gauge" vector

$$R = (\xi^{\pm}_{(\mu} X_{\nu)}, 0)^T. \tag{6.5.5}$$

The LHS vanishes and the RHS gives the $H^{\pm}$-inner product of $(R, \xi_0^{\pm} R)^T$ with $(S, \xi_0^{\pm} S)^T$. It follows that $(S, \xi_0^{\pm} S)^T$ must be $H^{\pm}$ orthogonal to any pure gauge eigenvector. Similarly, contracting this equation with $S$ and using ( 6.5.2) shows that $(S, \xi_0^\pm S)^T$ is null w.r.t. $H^\pm$. Hence if this vector were not pure gauge then we could add it to $N^\pm$ and violate maximality of this null subspace. Therefore this vector must be pure gauge, i.e., we have $S = (\xi^\pm_{(\mu} Y_{\nu)}, 0)^T$ for some $Y_\mu \neq 0$. So, writing $T = (t_{\mu\nu}, \chi)^T$, (6.5.4) takes the form

$$\delta P(\xi^\pm) \cdot \begin{pmatrix} t_{\rho\sigma} \\ \chi \end{pmatrix} = -(2\xi_0^\pm A + B) \cdot \begin{pmatrix} \xi^\pm_{(\rho} Y_{\sigma)} \\ 0 \end{pmatrix}. \tag{6.5.6}$$

If this equation admits a solution for some $Y_\mu \neq 0$ then $M$ has a non-trivial Jordan block. So strong hyperbolicity requires that this equation admits no solution $(t_{\mu\nu}, \chi)^T$ for any $Y_\mu \neq 0$.

### Outline of the proofs

In the following sections we will study in detail the strong hyperbolicity of Horndeski theories. The main idea will be to verify for which subclasses of Horndeski theories Condition 2 is satisfied. For those theories that satisfy this condition we then check whether Eq. (6.5.6) admits a solution for non-vanishing $Y^\mu$. Note that requiring these conditions to be satisfied will force $f$ and $\mathcal{H}$ in the gauge condition to take specific forms.

Since various Horndeski theories exhibit different qualitative features — in particular, the tensorial structure of the principal symbol is not the same for all theories — we will not tackle the problem of establishing the strong hyperbolicity of the most general Horndeski theory directly. Instead, we will begin from the simplest non-trivial case (i.e. the simplest theory with non-minimal coupling between the metric and the scalar field) and then we will progressively include more general terms. More precisely, in Section 6.6 we consider all the Horndeski theories with $\mathcal{G}_5 = 0$: we start from the $\mathcal{G}_4 = \mathcal{G}_5 = 0$ case, we then include a non-vanishing $\mathcal{G}_4$ term, distinguishing between two cases according to whether it depends on $X$ or not.

Before moving onto the general $\mathcal{G}_5 \neq 0$ we will study in Section 6.7 a special case of this theory: Einstein–dilaton–Gauss–Bonnet gravity. This, besides being a theory of particular interest on its own, will provide a "pedagogical" introduction to the proof of the general result. Finally, in Section 6.8 we will build upon the results obtained to study the hyperbolicity of the most general Horndeski theory.

## 6.6 Strong hyperbolicity: the $\mathcal{G}_5 = 0$ case

### Proof of strong hyperbolicity for $\mathcal{G}_4 = \mathcal{G}_5 = 0$

Let us begin by considering the theory with Lagrangian

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3. \tag{6.6.1}$$

The non-linear equations of motion for this theory are

$$E_{ab} \equiv G_{ab} + \partial_X \mathcal{G}_3 \left[ -\frac{1}{2} \Box\Phi \nabla_a \Phi \nabla_b \Phi + G_{ab}{}^{ed} \nabla^c \nabla_e \Phi \nabla_c \Phi \nabla_d \Phi \right] + \ldots = 0 \tag{6.6.2}$$

$$\begin{aligned} E_\Phi \equiv & -(1 + \partial_X \mathcal{G}_2 + 2\partial_\Phi \mathcal{G}_3 + 2X\partial^2_{X\Phi}\mathcal{G}_3)\Box\Phi \\ & + \partial_X^2 \mathcal{G}_2 \nabla^a \Phi \nabla^b \Phi \nabla_a \nabla_b \Phi \\ & - (\partial_X \mathcal{G}_3 + X\partial_X^2 \mathcal{G}_3)\delta^{c_1 c_2}_{d_1 d_2} \nabla_{c_1} \nabla^{d_1} \Phi \nabla_{c_2} \nabla^{d_2} \Phi \\ & - \frac{1}{2} \partial_X^2 \mathcal{G}_3 \delta^{c_1 c_2 c_3}_{d_1 d_2 d_3} \nabla_{c_1} \nabla^{d_1} \Phi \nabla_{c_2} \nabla^{d_2} \Phi \nabla_{c_3} \Phi \nabla^{d_3} \Phi \\ & - 2\partial^2_{X\Phi} \mathcal{G}_3 \delta^{c_1 c_2}_{d_1 d_2} \nabla_{c_1} \nabla^{d_1} \Phi \nabla_{c_2} \Phi \nabla^{d_2} \Phi \\ & + \partial_X \mathcal{G}_3 R_{ab} \nabla^a \Phi \nabla^b \Phi + \ldots = 0 \end{aligned} \tag{6.6.3}$$

where again the ellipsis denotes terms not involving second derivatives. After determining the linearised equations in generalised harmonic gauge (6.2.14), we compute the principal symbol and we find that

$$\delta P_{gg}(\xi)^{abcd} = \delta Q_{gg}(\xi)^{abcd} \tag{6.6.4}$$

$$\delta P_{g\Phi}(\xi)^{ab} = \delta P_{\Phi g}(\xi)^{ab} = -\frac{1}{2} \partial_X \mathcal{G}_3 \nabla^a \Phi \nabla^b \Phi \, \xi^2 + \xi^c G^{deab} \xi_e \mathcal{K}_{cd} \tag{6.6.5}$$

$$\begin{aligned} \delta P_{\Phi\Phi}(\xi) = & \left(-\partial_X \mathcal{G}_2 - 2\partial_\Phi \mathcal{G}_3 + 2X\partial^2_{X\Phi}\mathcal{G}_3 - 2\partial_X \mathcal{G}_3 \Box\Phi - 2X\partial_X^2 \mathcal{G}_3 \Box\Phi\right) \xi^2 \\ & + 2\left(\partial_X \mathcal{G}_3 + X\partial_X^2 \mathcal{G}_3\right) \xi^c \xi^d \nabla_c \nabla_d \Phi + (2\partial^2_{X\Phi}\mathcal{G}_3 + \partial_X^2 \mathcal{G}_2)(\xi \cdot \nabla\Phi)^2 \\ & - \partial_X^2 \mathcal{G}_3 \delta^{c_1 c_2 c_3}_{d_1 d_2 d_3} \xi_{c_1} \xi^{d_1} \nabla_{c_2} \nabla^{d_2} \Phi \nabla_{c_3} \Phi \nabla^{d_3} \Phi + \delta Q_{\Phi\Phi}(\xi) \end{aligned} \tag{6.6.6}$$

where

$$\mathcal{K}_{ab} \equiv (1 + f)\mathcal{H}_{ab} + \partial_X \mathcal{G}_3 \nabla_a \Phi \nabla_b \Phi. \tag{6.6.7}$$

For strong hyperbolicity to hold, equation (6.5.6) must admit no solution $(t_{\mu\nu}, \chi)^T$ when $Y_\mu \neq 0$. Writing out this equation gives

$$\text{LHS} = \begin{pmatrix} G^{\mu\nu\rho\sigma}[-f(f+2)G_\rho{}^{\lambda\alpha\beta}\xi^\pm_\sigma\xi^\pm_\lambda t_{\alpha\beta} + \xi^{\pm\lambda}\mathcal{K}_{\lambda\rho}\xi^\pm_\sigma\chi] \\ \\ \xi^{\pm\mu}G^{\nu\lambda\rho\sigma}\xi^\pm_\lambda t_{\rho\sigma}\mathcal{K}_{\mu\nu} + \delta P_{\Phi\Phi}(\xi^\pm)\chi \end{pmatrix} \tag{6.6.8}$$

$$\text{RHS} = \begin{pmatrix} \xi^{0\pm}G^{\mu\nu\rho\sigma}\xi^\pm_\rho Y_\sigma \\ \\ (\partial_X\mathcal{G}_3)\xi^{0\pm}(\xi^\pm\cdot\nabla\Phi)(Y\cdot\nabla\Phi) - \mathcal{K}_{\lambda\sigma}\xi^{\pm\lambda}G^{\mu\nu\sigma 0}\xi^\pm_\mu Y_\nu \end{pmatrix}. \tag{6.6.9}$$

Looking at the first row of this equation, the non-degeneracy of $G^{\mu\nu\rho\sigma}$ implies that if $f \neq 0$ then we can solve for the "non-transverse" part of $t_{\mu\nu}$:[61]

$$G_\mu{}^{\nu\rho\sigma}\xi^\pm_\nu t_{\rho\sigma} = \frac{1}{f(f+2)}(\xi^{\pm\rho}\mathcal{K}_{\rho\mu}\chi - \xi^{0\pm}Y_\mu). \tag{6.6.10}$$

This can then be substituted into the second row to obtain an equation that determines $\chi$. Hence if $f \neq 0$ then a solution of (6.5.6) exists for any $Y_\mu \neq 0$. Therefore strong hyperbolicity requires that $f = 0$. With $f = 0$, the first row of (6.6.8) implies

$$\xi^{0\pm}Y_\mu = \xi^{\pm\rho}\mathcal{K}_{\rho\mu}\chi. \tag{6.6.11}$$

Plugging this into the second row of (6.6.8) now gives a linear homogeneous scalar equation for $\chi$ and $t_{\mu\nu}$. Since this is only one equation for 11 unknowns, there exist non-trivial solutions. We see that we can solve (6.5.6) for $Y_\mu$ of the form (6.6.11). Hence if this $Y_\mu$ is non-vanishing then the equation is not strongly hyperbolic. Therefore strong hyperbolicity, requires (6.6.11) to vanish for any (null) $\xi^\pm_\mu$ which implies (since generically $\chi \neq 0$) $\mathcal{K}_{\mu\nu} = 0$. Hence strong hyperbolicity selects a unique generalised harmonic gauge:

$$f = 0 \qquad \mathcal{H}_{ab} = -\partial_X\mathcal{G}_3\nabla_a\Phi\nabla_b\Phi. \tag{6.6.12}$$

Note that this guarantees that the smallness condition (6.2.31) is satisfied. This follows from (6.2.30e), which in the $\mathcal{G}_4 = \mathcal{G}_5 = 0$ case reduces to:

$$|\partial^k_X(\partial_X\mathcal{G}_3)|L^{-2(1+k)} \ll 1, \qquad k = 0, 1. \tag{6.6.13}$$

If our gauge functions $f$ and $\mathcal{H}_{ab}$ satisfy this equation then $M$ is diagonalizable, as required by strong hyperbolicity. As explained above (2.2.19), diagonalizability ensures that there exists a positive definite symmetrizer $K$ satisfying (2.3.16). To complete the proof of strong hyperbolicity we need to check that $K$ depends smoothly on $\xi_i$. We will do this in a more general setting later in this Chapter (see the discussion below Eq. (6.6.52)).

[61] Note that our smallness assumption (6.2.31) implies that $f \neq -2$.

### Failure of strong hyperbolicity for $\partial_X \mathcal{G}_4 \neq 0$, $\mathcal{G}_5 = 0$

The situation is different if we include $\mathcal{L}_4$ i.e. we work with the theory

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4 \tag{6.6.14}$$

We will show that if $\partial_X \mathcal{G}_4 \neq 0$ then there is no generalised harmonic gauge for which this theory is strongly hyperbolic.

The terms in the equations of motion $E^a{}_b$ and $E_\Phi$ arising from $\mathcal{L}_4$ are [31, 103]

$$\begin{aligned}
E^{a(4)}_{\ b} = &(\mathcal{G}_4 - 2X\partial_X \mathcal{G}_4)G^a{}_b + \frac{1}{4}\partial_X \mathcal{G}_4 \delta^{ac_1c_2c_3}_{bd_1d_2d_3} R_{c_1c_2}{}^{d_1d_2}\nabla_{c_3}\Phi\nabla^{d_3}\Phi \\
&+ \frac{1}{2}(\partial_X \mathcal{G}_4 + 2X\partial^2_X \mathcal{G}_4)\delta^{ac_1c_2}_{bd_1d_2}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi \\
&+ \frac{1}{2}\partial^2_X \mathcal{G}_4 \delta^{ac_1c_2c_3}_{bd_1d_2d_3}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\Phi\nabla^{d_3}\Phi \\
&+ 2\partial^2_{X\Phi}\mathcal{G}_4\delta^{ac_1c_2}_{bd_1d_2}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\Phi\nabla^{d_2}\Phi \\
&+ (\partial_\Phi \mathcal{G}_4 + 2X\partial^2_{X\Phi}\mathcal{G}_4)\delta^{ac_1}_{bd_1}\nabla_{c_1}\nabla^{d_1}\Phi
\end{aligned} \tag{6.6.15}$$

$$\begin{aligned}
E^{(4)}_\Phi = &-(\partial_\Phi \mathcal{G}_4 + 2X\partial^2_{X\Phi}\mathcal{G}_4)R \\
&- \frac{1}{2}(\partial_X \mathcal{G}_4 + 2X\partial^2_X \mathcal{G}_4)\delta^{c_1c_2c_3}_{d_1d_2d_3}\nabla_{c_1}\nabla^{d_1}\Phi R_{c_2c_3}{}^{d_2d_3} \\
&- \frac{1}{2}\partial^2_X \mathcal{G}_4 \delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\Phi\nabla^{d_2}\Phi R_{c_3c_4}{}^{d_3d_4} \\
&- \partial^2_{X\Phi}\mathcal{G}_4\delta^{c_1c_2c_3}_{d_1d_2d_3}\nabla_{c_1}\Phi\nabla^{d_1}\Phi R_{c_2c_3}{}^{d_2d_3} \\
&- \left(\partial^2_X \mathcal{G}_4 + \frac{2}{3}X\partial^3_X \mathcal{G}_4\right)\delta^{c_1c_2c_3}_{d_1d_2d_3}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\nabla^{d_3}\Phi \\
&- 2\partial^3_{X\Phi\Phi}\mathcal{G}_4\delta^{c_1c_2}_{d_1d_2}\nabla_{c_1}\Phi\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi \\
&- (2X\partial^3_{XX\Phi}\mathcal{G}_4 + 3\partial^2_{X\Phi}\mathcal{G}_4)\delta^{c_1c_2}_{d_1d_2}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi \\
&- 2\partial^3_{XX\Phi}\mathcal{G}_4\delta^{c_1c_2c_3}_{d_1d_2d_3}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\Phi\nabla^{d_3}\Phi \\
&- \frac{1}{3}\partial^3_X \mathcal{G}_4\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\nabla_{c_1}\nabla^{d_1}\Phi\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\nabla^{d_3}\Phi\nabla_{c_4}\Phi\nabla^{d_4}\Phi.
\end{aligned} \tag{6.6.16}$$

Linearising these equations, and including the gauge-fixing terms, one can then compute $\delta\tilde{P}^{(4)}$, the contribution to $\delta\tilde{P}$ arising from $\mathcal{L}_4$. It takes the following form

$$\begin{aligned}
\delta\tilde{P}^{(4)}_{gg}(\xi)^{a}{}_{b}{}^{cd}t_{cd} =& -\frac{1}{2}(\mathcal{G}_4 - 2X\partial_X\mathcal{G}_4)\delta^{ac_1c_2}_{bd_1d_2}\xi_{c_1}\xi^{d_1}t_{c_2}{}^{d_2} \\
&- \frac{1}{2}\partial_X\mathcal{G}_4\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\xi_{c_1}\xi^{d_1}t_{c_2}{}^{d_2}\nabla_{c_3}\Phi\nabla^{d_3}\Phi \qquad (6.6.17)\\
\delta\tilde{P}^{(4)}_{g\Phi}(\xi)^{a}{}_{b} = \delta\tilde{P}^{(4)}_{\Phi g}(\xi)^{a}{}_{b} =& (\partial_X\mathcal{G}_4 + 2X\partial_X^2\mathcal{G}_4)\delta^{ac_1c_2}_{bd_1d_2}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\nabla^{d_2}\Phi \\
&+ \partial_X^2\mathcal{G}_4\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\Phi\nabla^{d_3}\Phi \\
&+ 2\partial^2_{X\Phi}\mathcal{G}_4\delta^{ac_1c_2}_{bd_1d_2}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\Phi\nabla^{d_2}\Phi \\
&+ (\partial_\Phi\mathcal{G}_4 + 2X\partial^2_{X\Phi}\mathcal{G}_4)\delta^{ac_1}_{bd_1}\xi_{c_1}\xi^{d_1} \qquad (6.6.18)\\
\delta\tilde{P}^{(4)}_{\Phi\Phi}(\xi) =& -\frac{1}{2}(\partial_X\mathcal{G}_4 + 2X\partial_X^2\mathcal{G}_4)\delta^{c_1c_2c_3}_{d_1d_2d_3}\xi_{c_1}\xi^{d_1}R_{c_2c_3}{}^{d_2d_3} \\
&- \frac{1}{2}\partial_X^2\mathcal{G}_4\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\Phi\nabla^{d_2}\Phi R_{c_3c_4}{}^{d_3d_4} \\
&- (3\partial_X^2\mathcal{G}_4 + 2X\partial_X^3\mathcal{G}_4)\delta^{c_1c_2c_3}_{d_1d_2d_3}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\nabla^{d_3}\Phi \\
&- \partial_X^3\mathcal{G}_4\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\nabla^{d_3}\Phi\nabla_{c_4}\Phi\nabla^{d_4}\Phi \\
&- 4\partial^3_{XX\Phi}\mathcal{G}_4\delta^{c_1c_2c_3}_{d_1d_2d_3}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\Phi\nabla^{d_3}\Phi \\
&- 2(2X\partial^3_{XX\Phi}\mathcal{G}_4 + 3\partial^2_{X\Phi}\mathcal{G}_4)\delta^{c_1c_2}_{d_1d_2}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\nabla^{d_2}\Phi \\
&- 2\partial^3_{X\Phi\Phi}\mathcal{G}_4\delta^{c_1c_2}_{d_1d_2}\xi_{c_1}\xi^{d_1}\nabla_{c_2}\Phi\nabla^{d_2}\Phi. \qquad (6.6.19)
\end{aligned}$$

As discussed above, for the equations to be strongly hyperbolic it is necessary that the kernel of $\delta P(\xi^\pm)$ has dimension 8 or greater. We will now study whether this condition is satisfied. A vector $(t_{ab},\chi)^T$ is in $\ker\delta P(\xi^\pm)$ if, and only if,

$$\begin{pmatrix} \delta P_{gg}(\xi^\pm)^{abcd}t_{cd} + \delta P_{g\Phi}(\xi^\pm)^{ab}\chi \\ \delta P_{\Phi g}(\xi^\pm)^{cd}t_{cd} + \delta P_{\Phi\Phi}(\xi^\pm)\chi \end{pmatrix} = 0. \qquad (6.6.20)$$

We now assume that $\partial_X\mathcal{G}_4 \neq 0$. In this case we can separate out a term proportional to $t_{ab}$ in the first row of (6.6.20) and write this equation as

$$\begin{aligned}
(\xi^\pm\cdot\nabla\Phi)^2 t_{ab} =& -\xi^\pm_a\xi^\pm_b(\nabla^c\Phi\nabla^d\Phi t_{cd} + \mathcal{G}_4 t^c{}_c) \\
&+ 2\xi^\pm_{(a}t_{b)c}\left(\frac{\mathcal{G}_4}{\partial_X\mathcal{G}_4}\xi^{\pm c} + \nabla^c\Phi(\xi^\pm\cdot\nabla\Phi)\right) \\
&- 2\xi^\pm_{(a}\nabla_{b)}\Phi(t^c{}_c(\xi^\pm\cdot\nabla\Phi) - \xi^{\pm c}\nabla^d\Phi t_{cd}) \\
&- g_{ab}\left(2\xi^{\pm c}\nabla^d\Phi t_{cd}(\xi^\pm\cdot\nabla\Phi) + \frac{\mathcal{G}_4}{\partial_X\mathcal{G}_4}\xi^{\pm c}\xi^{\pm d}t_{cd} - t^c{}_c(\xi^\pm\cdot\nabla\Phi)^2\right) \\
&- \nabla_{(a}\Phi\nabla_{b)}\Phi\xi^{\pm c}\xi^{\pm d}t_{cd} - \nabla_{(a}\Phi t_{b)c}\xi^{\pm c}(\xi^\pm\cdot\nabla\Phi) \\
&+ \frac{2}{\partial_X\mathcal{G}_4}(\delta Q_{gg}(\xi^\pm)_{ab}{}^{cd}t_{cd} + \delta P_{g\Phi}(\xi^\pm)_{ab}\chi). \qquad (6.6.21)
\end{aligned}$$

Note that for a *generic* weak-field background, and *generic* $\xi^\pm$, we have

$$\xi^\pm\cdot\nabla\Phi \neq 0. \qquad (6.6.22)$$

From the tensor structure of this equation, we deduce that $t_{ab}$ must take the form

$$t_{ab} = \xi^{\pm}_{(a} Y_{b)} + \lambda g_{ab} + Z_{(a} \nabla_{b)} \Phi + \mu \nabla_a \nabla_b \Phi \tag{6.6.23}$$

for some $Y_a$, $\lambda$, $Z_a$ and $\mu$. The last term in this expression comes from the fact that $\delta P_{g\Phi}(\xi^{\pm})_{ab}$ contains terms proportional to $\nabla_a \nabla_b \Phi$ as well as terms of the other three types. There is some degeneracy in this expression, e.g. degeneracy between the first and third terms implies that $Z_a$ is defined only up to addition of a multiple of $\xi^{\pm}_a$, i.e., the part of $Z_a$ parallel to $\xi^{\pm}_a$ is "pure gauge". For strong hyperbolicity we need there to exist at least 4 linearly independent "non-gauge" elements of $\ker \delta P(\xi^{\pm})$. The first term in (6.6.23) is pure gauge. The "non-gauge part" is determined by $\chi, \lambda, \mu$ and the non-gauge part of $Z_a$.

Plugging (6.6.23) back into the first row of (6.6.20) we get

$$\begin{aligned}
0 &= \delta P_{gg}(\xi^{\pm})^a{}_b{}^{cd} t_{cd} + \delta P_{g\Phi}(\xi^{\pm})^a{}_b \chi \\
&= \delta^{ac_1c_2c_3}_{bd_1d_2d_3} \xi^{\pm}_{c_1} \xi^{\pm d_1} \nabla_{c_2} \nabla^{d_2} \Phi \nabla_{c_3} \Phi \nabla^{d_3} \Phi \left( -\frac{1}{2} \partial_X \mathcal{G}_4 \mu + \partial_X^2 \mathcal{G}_4 \chi \right) \\
&+ \delta^{ac_1c_2}_{bd_1d_2} \xi^{\pm}_{c_1} \xi^{\pm d_1} \Bigg[ \nabla_{c_2} \Phi \nabla^{d_2} \Phi \left( -\frac{1}{2} \partial_X \mathcal{G}_4 \lambda + 2 \partial^2_{X\Phi} \mathcal{G}_4 \chi \right) \\
&\qquad\qquad - \frac{1}{4} (\mathcal{G}_4 - 2X \partial_X \mathcal{G}_4 - f(f+2))(Z_{c_2} \nabla^{d_2} \Phi + \nabla_{c_2} \Phi Z^{d_2}) \Bigg] \\
&+ \delta^{ac_1c_2}_{bd_1d_2} \xi^{\pm}_{c_1} \xi^{\pm d_1} \nabla_{c_2} \nabla^{d_2} \Phi \Bigg[ (\partial_X \mathcal{G}_4 + 2X \partial_X^2 \mathcal{G}_4) \chi \\
&\qquad\qquad - \frac{1}{2} \mu (\mathcal{G}_4 - 2X \partial_X \mathcal{G}_4 - f(f+2)) \Bigg] \\
&- \xi^{\pm a} \xi^{\pm}_b \Big[ (\partial_\Phi \mathcal{G}_4 + 2X \partial^2_{X\Phi} \mathcal{G}_4) \chi - \lambda (\mathcal{G}_4 - 2X \partial_X \mathcal{G}_4 - f(f+2)) \Big] \\
&+ \xi^{\pm c} G^{dea}{}_b \xi^{\pm}_e \mathcal{K}_{cd} \chi,
\end{aligned} \tag{6.6.24}$$

where $\mathcal{K}_{ab}$ is defined in (6.6.7). We will now show that the requirement of strong hyperbolicity fixes our choice of gauge. Consider first the case

$$\mathcal{G}_4 - 2X \partial_X \mathcal{G}_4 - f(f+2) \neq 0. \tag{6.6.25}$$

In this case, (6.6.24) contains $Z_a$-dependent terms proportional to

$$\delta^{ac_1c_2}_{bd_1d_2} \xi^{\pm}_{c_1} \xi^{\pm d_1} (Z_{c_2} \nabla^{d_2} \Phi + \nabla_{c_2} \Phi Z^{d_2}) = 4 G^a{}_b{}^{ce} \xi^{\pm d} \xi^{\pm}_e G_{cd}{}^{fh} Z_f \nabla_h \Phi. \tag{6.6.26}$$

View the RHS as an operator $\mathcal{O}$ acting on $Z_a$. Let's determine the kernel of this operator. Since $G^{abcd}$ is non-degenerate, vectors in the kernel must satisfy

$$\xi^{\pm d} \xi^{\pm}_{(e} G_{c)d}{}^{fh} Z_f \nabla_h \Phi = 0 \qquad \Rightarrow \qquad \xi^{\pm d} G_{cd}{}^{fh} Z_f \nabla_h \Phi = 0. \tag{6.6.27}$$

However, for generic $\nabla_a\Phi$, it is easy to show that all solutions of this equation have $Z_a$ proportional to $\xi^\pm_a$. Hence the kernel of $\mathcal{O}$ generically contains only vectors proportional to $\xi^\pm_a$. This implies that, generically, if equation (6.6.24) admits a solution then $Z_a$ is determined up to a multiple of $\xi^\pm_a$, in terms of $\chi, \lambda, \mu$. In other words, the non-gauge part of $Z_a$ is fixed uniquely by the 3 quantities $\chi, \lambda, \mu$. Therefore, there exist at most 3 linearly independent non-gauge elements of $\ker \delta P(\xi^\pm)$, whereas strong hyperbolicity requires at least 4 such elements. So if our gauge condition satisfies (6.6.25) then the equation is not strongly hyperbolic.

We have shown that strong hyperbolicity requires that our gauge function $f$ obeys

$$\mathcal{G}_4 - 2X\partial_X\mathcal{G}_4 - f(f+2) = 0. \tag{6.6.28}$$

We can solve this quadratic equation and choose the root that satisfies the smallness condition (6.2.31) when the conditions (6.2.30d) are satisfied:

$$f = -1 + \sqrt{1 + \mathcal{G}_4 - 2X\partial_X\mathcal{G}_4}. \tag{6.6.29}$$

The contraction of (6.6.24) with $\nabla^b\Phi$ gives

$$0 = \xi^{\pm c}\xi^\pm_e \nabla_b\Phi G^{deab}\tilde{\mathcal{K}}_{cd}\chi, \tag{6.6.30}$$

where

$$\tilde{\mathcal{K}}_{cd} \equiv \mathcal{K}_{cd} - (\alpha g_{cd} + \beta\nabla_c\nabla_d\Phi), \tag{6.6.31}$$

with

$$\alpha = \partial_\Phi\mathcal{G}_4 + 2X\partial^2_{X\Phi}\mathcal{G}_4 + \nabla^a\nabla_a\Phi(\partial_X\mathcal{G}_4 + 2X\partial^2_X\mathcal{G}_4), \tag{6.6.32}$$

$$\beta = -2(\partial_X\mathcal{G}_4 + 2X\partial^2_X\mathcal{G}_4). \tag{6.6.33}$$

Consider first the case in which our gauge condition is such that, generically,

$$\xi^{\pm c}\xi^\pm_e \nabla_b\Phi G^{deab}\tilde{\mathcal{K}}_{cd} \neq 0. \tag{6.6.34}$$

Then, in a generic background, for generic null $\xi^\pm_a$, (6.6.30) implies that we must have $\chi = 0$ and equation (6.6.24) then reduces to

$$0 = -\frac{1}{2}\partial_X\mathcal{G}_4\mu\, \delta^{ac_1c_2c_3}_{bd_1d_2d_3}\xi^\pm_{c_1}\xi^{\pm d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\Phi\nabla^{d_3}\Phi - \partial_X\mathcal{G}_4\lambda\, G^a{}_b{}^{ec}\xi^{\pm d}\xi^\pm_e G_{cd}{}^{fh}\nabla_f\Phi\nabla_h\Phi. \tag{6.6.35}$$

In a generic background this implies $\lambda = \mu = 0$ (using $\partial_X\mathcal{G}_4 \neq 0$). But with $\chi = \lambda = \mu = 0$, the "non-gauge" part of the vector $(t_{ab}, \chi)^T$ is determined entirely by $Z_a$ which has at most 3 independent non-gauge components. So in this case we do not have enough non-gauge elements of $\ker \delta P(\xi^\pm)$ for strong hyperbolicity.

We have shown that strong hyperbolicity requires that, generically,

$$\xi^{\pm c}\xi^{\pm}_{e}\nabla_b\Phi G^{deab}\tilde{\mathcal{K}}_{cd}=0. \tag{6.6.36}$$

For this to be satisfied for generic null $\xi^{\pm}$ we must have

$$G^{abde}\nabla_b\Phi\tilde{\mathcal{K}}_{cd}=\rho^a\delta^e_c \tag{6.6.37}$$

for some vector $\rho^a$. Contracting with $\nabla_a\Phi$ we see that

$$(\nabla\Phi)^2\tilde{\mathcal{K}}_a{}^b=2\tilde{\mathcal{K}}_{ac}\nabla^c\Phi\nabla^b\Phi-2(\rho\cdot\nabla\Phi)\delta^b_a \tag{6.6.38}$$

from which we deduce that the most general form $\tilde{\mathcal{K}}$ can take is

$$\tilde{\mathcal{K}}_{ab}=\kappa g_{ab}+W_a\nabla_b\Phi. \tag{6.6.39}$$

for some scalar $\kappa$ and vector $W_a$. Note that we can determine $\rho^a$ in terms of these quantities by taking the trace over the $e$ and $c$ indices in (6.6.37)

$$\rho^a=\frac{1}{4}\left(-\kappa\nabla^a\Phi+G^{abcd}W_c\nabla_d\Phi\nabla_b\Phi\right). \tag{6.6.40}$$

Plugging these back into (6.6.37) we find that the only solution is given by $\kappa=0$ and $W_a=0$, that is

$$\tilde{\mathcal{K}}_{ab}=0. \tag{6.6.41}$$

Hence strong hyperbolicity for a generic weak-field background forces us to make the gauge choice

$$\begin{aligned} f&=-1+\sqrt{1+\mathcal{G}_4-2X\partial_X\mathcal{G}_4}\\ (1+f)\mathcal{H}_{ab}&=\alpha g_{ab}+\beta\nabla_a\nabla_b\Phi-\partial_X\mathcal{G}_3\nabla_a\Phi\nabla_b\Phi. \end{aligned} \tag{6.6.42}$$

With this choice of gauge, (6.6.24) reduces to

$$\begin{aligned} 0=&\,\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\xi^{\pm}_{c_1}\xi^{\pm d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\Phi\nabla^{d_3}\Phi\left(-\frac{1}{2}\partial_X\mathcal{G}_4\mu+\partial_X^2\mathcal{G}_4\chi\right)\\ &+2G^a{}_b{}^{ec}\xi^{\pm d}\xi^{\pm}_e G_{cd}{}^{fh}\nabla_f\Phi\nabla_h\Phi\left(-\frac{1}{2}\partial_X\mathcal{G}_4\lambda+2\partial^2_{X\Phi}\mathcal{G}_4\chi\right). \end{aligned} \tag{6.6.43}$$

For a generic background, this fixes $\lambda$ and $\mu$ in terms of $\chi$:

$$\lambda=4\frac{\partial^2_{X\Phi}\mathcal{G}_4}{\partial_X\mathcal{G}_4}\chi\qquad\mu=2\frac{\partial^2_X\mathcal{G}_4}{\partial_X\mathcal{G}_4}\chi. \tag{6.6.44}$$

We now consider the second row of (6.6.20), which takes the form

$$\mathbb{A}\chi=0 \tag{6.6.45}$$

where

$$
\begin{aligned}
\mathbb{A} = &-\frac{1}{2}[\partial_X\mathcal{G}_4 + 2X\partial_X^2\mathcal{G}_4]\delta^{c_1c_2c_3}_{d_1d_2d_3}\xi^{\pm}_{c_1}\xi^{\pm d_1}R_{c_2c_3}{}^{d_2d_3}\\
&-\frac{1}{2}\partial_X^2\mathcal{G}_4\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\xi^{\pm}_{c_1}\xi^{\pm d_1}\nabla_{c_2}\Phi\nabla^{d_2}\Phi R_{c_3c_4}{}^{d_3d_4}\\
&-\left[3\partial_X^2\mathcal{G}_4 + 2X\partial_X^3\mathcal{G}_4 + \frac{2(\partial_X\mathcal{G}_4 + 2X\partial_X^2\mathcal{G}_4)^2}{1+\mathcal{G}_4 - 2X\partial_X\mathcal{G}_4}\right]\delta^{c_1c_2c_3}_{d_1d_2d_3}\xi^{\pm}_{c_1}\xi^{\pm d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\nabla^{d_3}\Phi\\
&+[2(\partial_X\mathcal{G}_4)^{-1}(\partial_X^2\mathcal{G}_4)^2 - \partial_X^3\mathcal{G}_4]\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\xi^{\pm}_{c_1}\xi^{\pm d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\nabla^{d_3}\Phi\nabla_{c_4}\Phi\nabla^{d_4}\Phi\\
&+4\Bigg[\partial^3_{XX\Phi}\mathcal{G}_4 + 2\frac{\partial_X^2\mathcal{G}_4\partial^2_{X\Phi}\mathcal{G}_4}{\partial_X\mathcal{G}_4}\\
&\qquad -\frac{\partial_X\mathcal{G}_3(\partial_X\mathcal{G}_4 + 2X\partial_X^2\mathcal{G}_4)}{2(1+\mathcal{G}_4 - 2X\partial_X\mathcal{G}_4)} - \partial_X^2\mathcal{G}_3\Bigg]\delta^{c_1c_2c_3}_{d_1d_2d_3}\xi^{\pm}_{c_1}\xi^{\pm d_1}\nabla_{c_2}\nabla^{d_2}\Phi\nabla_{c_3}\Phi\nabla^{d_3}\Phi\\
&+2\Bigg[-8\frac{(\partial^2_{X\Phi}\mathcal{G}_4)^2}{\partial_X\mathcal{G}_4} + \frac{\partial_X\mathcal{G}_3(\partial_\Phi\mathcal{G}_4 + 2X\partial^2_{X\Phi}\mathcal{G}_4 + X\partial_X\mathcal{G}_3)}{(1+\mathcal{G}_4 - 2X\partial_X\mathcal{G}_4)}\\
&\qquad + \left(\partial^2_{X\Phi}\mathcal{G}_3 + \frac{1}{2}\partial_X^2\mathcal{G}_2 + \partial^3_{X\Phi\Phi}\mathcal{G}_4\right)\Bigg](\xi^{\pm}\cdot\nabla\Phi)^2\\
&+2\Bigg[2\frac{(\partial_X\mathcal{G}_4 + 2X\partial_X^2\mathcal{G}_4)(\partial_\Phi\mathcal{G}_4 + 2X\partial^2_{X\Phi}\mathcal{G}_4 - X\partial_X\mathcal{G}_3)}{1+\mathcal{G}_4 - 2X\partial_X\mathcal{G}_4}\\
&\qquad + (\partial_X\mathcal{G}_3 + X\partial_X^2\mathcal{G}_3 + 2X\partial^3_{XX\Phi}\mathcal{G}_4 + 3\partial^2_{X\Phi}\mathcal{G}_4)\Bigg]\xi^{\pm}_a\xi^{\pm}_b\nabla^a\nabla^b\Phi.
\end{aligned}
$$

If $\mathbb{A} \neq 0$ then we must have $\chi = 0$, and hence $\lambda = \mu = 0$ and $Z_a$ is arbitrary. Hence, in a generic weak-field background, $\ker\delta P(\xi^{\pm})$ consists of vectors of the form $(t_{ab}, 0)^T$ where $t_{ab}$ is given by (6.6.23) with $\lambda = \mu = 0$. Given that one component of $Z_a$ is "pure gauge" (i.e. degenerate with the first term in (6.6.23)), it follows that $\ker\delta P(\xi^{\pm})$ generically has dimension 7 and hence the equation of motion is not strongly hyperbolic.

The only way to escape this conclusion is if the theory is one for which $\mathbb{A} = 0$ for *any* background. For this to happen, terms with different dependence on the Riemann tensor, $\nabla\Phi$ and $\nabla\nabla\Phi$ have to cancel independently in $\mathbb{A}$. However this cannot happen in the case we are considering. To see this, note that vanishing of the terms of the (schematic) form $R\nabla\Phi\nabla\Phi$ in any background requires $\partial_X^2\mathcal{G}_4 = 0$. But then vanishing of the terms proportional to $R$ requires $\partial_X\mathcal{G}_4 = 0$, contradicting our assumption $\partial_X\mathcal{G}_4 \neq 0$. Hence in a generic background we have $\mathbb{A} \neq 0$ and therefore a vector in the kernel must have $\chi = 0$.

In summary, we have shown that when $\partial_X\mathcal{G}_4 \neq 0$, there does not exist a generalised harmonic gauge for which the equations of motion are strongly hyperbolic in a generic weak-field background. The best one can do is to choose the gauge (6.6.42), for which $\ker\delta P(\xi^{\pm})$ has dimension 7 in a generic weak-field background (i.e. 4 pure gauge elements, and 3 non-gauge elements). This implies that, in such a background, the matrix $M$ will have two non-trivial Jordan blocks: one in $V^+$ and one in $V^-$. Generically each of these will be $2 \times 2$.

## Proof of strong hyperbolicity for $\mathcal{G}_4 = \mathcal{G}_4(\Phi)$, $\mathcal{G}_5 = 0$

We continue working with the theory defined by (6.6.14), but now consider the case $\partial_X \mathcal{G}_4 = 0$, i.e., $\mathcal{G}_4 = \mathcal{G}_4(\Phi)$.[62] We will show that such theories are strongly hyperbolic in a suitable generalised harmonic gauge. The proof is analogous to that for the theory with $\mathcal{G}_4 = \mathcal{G}_5 = 0$ so we will be brief. For $\partial_X \mathcal{G}_4 = 0$, (6.5.6) reduces to

$$\mathrm{LHS} = \begin{pmatrix} G^{\mu\nu\rho\sigma}[(\mathcal{G}_4 - f(f+2))\xi^{\pm}_{\sigma} G_{\rho}{}^{\lambda\alpha\beta}\xi^{\pm}_{\lambda} t_{\alpha\beta} + \xi^{\pm\lambda}\xi^{\pm}_{\sigma}(\mathcal{K}_{\lambda\rho} - \partial_\Phi \mathcal{G}_4 g_{\lambda\rho})]\chi \\ \\ G^{\mu\nu\rho\sigma} t_{\mu\nu}\xi^{\pm}_{\rho}\xi^{\pm\lambda}(\mathcal{K}_{\lambda\sigma} - \partial_\Phi \mathcal{G}_4 g_{\lambda\sigma}) + \delta P_{\Phi\Phi}(\xi^{\pm})\chi \end{pmatrix} \tag{6.6.46}$$

$$\mathrm{RHS} = \begin{pmatrix} \xi^{0\pm}(1+\mathcal{G}_4) G^{\mu\nu\rho\sigma}\xi^{\pm}_{\rho} Y_\sigma \\ \\ \xi^{0\pm}[(\partial_X \mathcal{G}_3)(\xi^{\pm}\cdot\nabla\Phi)(Y\cdot\nabla\Phi) + \partial_\Phi \mathcal{G}_4(\xi^{\pm}\cdot Y)] - \mathcal{K}_{\lambda\sigma}\xi^{\pm\lambda} G^{\mu\nu\sigma 0}\xi^{\pm}_{\mu} Y_\nu \end{pmatrix}. \tag{6.6.47}$$

Recall that for strong hyperbolicity to hold, this equation must have no solution $(t_{\mu\nu}, \chi)^T$ when $Y_\mu \neq 0$. By the non-degeneracy of $G^{\mu\nu\rho\sigma}$ we see that if

$$\mathcal{G}_4 - f(f+2) \neq 0 \tag{6.6.48}$$

then we can use the first row of this equation to solve uniquely for $G_\mu{}^{\nu\rho\sigma}\xi^{\pm}_{\nu} t_{\rho\sigma}$ (the non-transverse part of $t$). This can then be substituted into the second row of the equation to give an equation which determines $\chi$. Hence, if $\mathcal{G}_4 - f(f+2) \neq 0$ then, for any non-zero $Y_\mu$, Eq. (6.5.6) has a solution. Therefore for strong hyperbolicity to hold, we need

$$\mathcal{G}_4 - f(f+2) = 0 \qquad \Rightarrow \qquad f = -1 + \sqrt{1+\mathcal{G}_4} \tag{6.6.49}$$

where we have chosen the root that satisfies the smallness condition (6.2.31). With this choice of $f$, the first row of (6.5.6) implies

$$\xi^{0\pm} Y_\mu = \frac{1}{1+\mathcal{G}_4}\xi^{\pm\rho}\tilde{\mathcal{K}}_{\rho\mu}\chi. \tag{6.6.50}$$

where

$$\tilde{\mathcal{K}}_{ab} = \mathcal{K}_{ab} - \partial_\Phi \mathcal{G}_4 g_{ab}. \tag{6.6.51}$$

[62] An example of such a theory is Brans-Dicke theory [23] with positive coupling constant $\omega$. After a redefinition of the scalar field, this has $\mathcal{G}_2 = \mathcal{G}_3 = 0$ and $\mathcal{G}_4 = \Phi/\sqrt{2\omega} + \Phi^2/(8\omega)$.

When we plug this into the second row of (6.5.6) we obtain a linear homogeneous scalar equation for $\chi$ and $t_{ab}$. This equation has 11 unknowns and therefore admits a non-trivial solution, generically with $\chi \neq 0$. It follows that if $Y_\mu$ in (6.6.50) is not vanishing, then strong hyperbolicity fails. This means that strong hyperbolicity requires $\xi^{\pm\rho}\tilde{\mathcal{K}}_{\rho\mu}\chi = 0$ for arbitrary null $\xi^\pm$. Since generically $\chi \neq 0$, this implies that we must choose our gauge such that $\tilde{\mathcal{K}}_{\mu\nu} = 0$. Thus we see that strong hyperbolicity in a generic weak-field background requires us to make the gauge choice

$$\begin{aligned} f =& -1 + \sqrt{1+\mathcal{G}_4} \\ (1+f)\mathcal{H}_{ab} =& \partial_\Phi \mathcal{G}_4 g_{ab} - \partial_X \mathcal{G}_3 \nabla_a \Phi \nabla_b \Phi. \end{aligned} \tag{6.6.52}$$

In this gauge, equation (6.5.6) implies $Y_\mu = 0$ so $M$ has no non-trivial Jordan block, i.e., $M$ is diagonalisable. Note that when $\mathcal{G}_4 = 0$ this reduces to the gauge choice (6.6.12).

**Smoothness of the symmetrizer**

Diagonalizability of $M$ is a necessary condition for strong hyperbolicity to hold. It ensures the existence of a positive definite symmetrizer $K$ satisfying (2.3.16). But we need to check that the remaining conditions in the definition of strong hyperbolicity are satisfied. In particular, we need to prove that $K$ depends smoothly on $\xi_i$. To do this, recall that $K$ is constructed from the matrix $S$ which diagonalizes $M$, as explained above (2.2.19). $S$ is the matrix whose columns are the eigenvectors of $M$. Hence if the eigenvectors of $M$ depend smoothly on $\xi_i$ then so does $K$. We will explicitly construct the eigenvectors of $M$ to demonstrate that they depend smoothly on $\xi_i$.

Recall that the eigenvectors of $M$ have the form (6.4.2) where $T$ satisfies (6.4.3). In the gauge (6.6.52), we have

$$\delta P_{gg}(\xi^\pm) = \delta P_{g\Phi}(\xi^\pm) = \delta P_{\Phi g}(\xi^\pm) = 0 \tag{6.6.53}$$

which implies that any vector of the form $T = (t_{ab}, 0)^T$ satisfies (6.4.3) when $\xi = \xi^\pm$. This proves that the eigenvalues $\xi_0^\pm$ each have degeneracy 10. If we choose a basis of symmetric tensors $t_{ab}$ that is independent of $\xi_i$ then the $\xi_i$-dependence of these eigenvectors arises only through the $\xi_0$ in (6.4.2), which implies that these 20 eigenvectors depend smoothly on $\xi_i$. A calculation reveals that the final two eigenvectors have $T = (t_{ab}, 1)^T$ where

$$t_{ab} = -\frac{\partial_X \mathcal{G}_3}{1+\mathcal{G}_4}\left[\nabla_a \Phi \nabla_b \Phi + g_{ab} X\right] - \partial_\Phi \log(1+\mathcal{G}_4) g_{ab} \tag{6.6.54}$$

and eigenvalues $\xi_0$ determined by

$$0 = f^{\mu\nu}\xi_\mu\xi_\nu \equiv -P_{\Phi\Phi}(\xi) - \frac{1}{(1+\mathcal{G}_4)}\left[X^2(\partial_X\mathcal{G}_3)^2 + 2(\partial_\Phi\mathcal{G}_4)^2\right]\xi^2. \tag{6.6.55}$$

For a weak field background, $f^{\mu\nu}$ is close to $g^{\mu\nu}$ and is therefore a Lorentzian metric with $f^{00} \neq 0$. This ensures that there will be two real eigenvalues $\xi_0$ depending smoothly on $\xi_i$. As before, the eigenvectors depend on $\xi_i$ only through $\xi_0$ and are therefore smooth. Hence all eigenvectors have the required smoothness in $\xi_i$ so the symmetrizer is smooth. This establishes strong hyperbolicity in the gauge (6.6.52).[63]

## 6.7 Failure of strong hyperbolicity for EdGB gravity

As explained at the end of Section 6.5, before investigating the strong hyperbolicity of the $\mathcal{G}_5 \neq 0$ Horndeski theory, we will discuss a special case, namely Einstein–dilaton–Gauss–Bonnet (EdGB) gravity [105]. The action for this theory is given by

$$S = \frac{1}{16\pi}\int \sqrt{-g}\,(R + X + F(\Phi)L_{\text{GB}}) \tag{6.7.1}$$

where $F(\Phi)$ is a smooth function and

$$L_{\text{GB}} = \frac{1}{4}\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}R_{c_1c_2}{}^{d_1d_2}R_{c_3c_4}{}^{d_3d_4}. \tag{6.7.2}$$

The equations of motion take the form[64]

$$E^a{}_b \equiv G^a{}_b + (F''(\Phi)\nabla_c\Phi\nabla^d\Phi + F'(\Phi)\nabla_c\nabla^d\Phi)\delta^{acc_1c_2}_{bdd_1d_2}R_{c_1c_2}{}^{d_1d_2} - \frac{1}{2}T^{(\Phi)}_{ab} = 0 \tag{6.7.3}$$

$$E_\Phi \equiv -\Box\Phi - \frac{1}{4}F'(\Phi)\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}R_{c_1c_2}{}^{d_1d_2}R_{c_3c_4}{}^{d_3d_4} = 0 \tag{6.7.4}$$

where

$$T^{(\Phi)}_{ab} = \nabla_a\Phi\nabla_b\Phi + g_{ab}X. \tag{6.7.5}$$

This theory can be cast in Horndeski form with the following choice of $\mathcal{G}_i$ [31]:

$$\mathcal{G}_2(\Phi, X) = 8X^2F^{(4)}(\Phi)(3 - \log|X|) \tag{6.7.6}$$

$$\mathcal{G}_3(\Phi, X) = -4XF^{(3)}(\Phi)(7 - 3\log|X|) \tag{6.7.7}$$

$$\mathcal{G}_4(\Phi, X) = 4XF''(\Phi)(2 - \log|X|) \tag{6.7.8}$$

$$\mathcal{G}_5(\Phi, X) = -4F'(\Phi)\log|X|. \tag{6.7.9}$$

[63] Actually we should also check the inequality below (2.3.16). This follows trivially if we restrict to a compact region of spacetime. For the $\mathcal{L}_1 + \mathcal{L}_2$ theory, a stronger result can be obtained [104]: this theory is *symmetric* hyperbolic even outside of the "weak field" regime provided that $1 + \partial_X\mathcal{G}_2 > 0$ and $1 + \partial_X\mathcal{G}_2 + 2X(\partial_X^2\mathcal{G}_2) > 0$. In our case, the smallness condition (6.2.30c) implies that these conditions are satisfied.

[64] Note that in the metric equation of motion it is sometimes included a term $-\frac{1}{8}F(\Phi)\delta^{ac_1c_2c_3c_4}_{bd_1d_2d_3d_4}R_{c_1c_2}{}^{d_1d_2}R_{c_3c_4}{}^{d_3d_4}$. However, in $d = 4$ we have that $\delta^{ac_1c_2c_3c_4}_{bd_1d_2d_3d_4} = 0$ and hence this term vanishes identically.

Note that, in this case, while the functions $\mathcal{G}_i$ are not smooth at $X = 0$, the combinations that appear in the equations of motion (and in the principal symbol) are. We can linearise the generalised harmonic gauge EdGB equations of motion around a background $(g, \Phi)$ and compute the principal symbol. The part of the principal symbol which arises from the Horndeski terms reads

$$(\delta\tilde{P}_{gg}(\xi)\cdot t)^a{}_b = -2(F''(\Phi)\nabla_c\Phi\nabla^d\Phi + F'(\Phi)\nabla_c\nabla^d\Phi)\delta^{acc_1c_2}_{bdd_1d_2}\xi_{c_1}\xi^{d_1}t_{c_2}{}^{d_2} \tag{6.7.10}$$

$$\delta\tilde{P}_{g\Phi}(\xi)^a{}_b = \delta\tilde{P}_{\Phi g}(\xi)^a{}_b = F'(\Phi)\delta^{acc_1c_2}_{bdd_1d_2}\xi_c\xi^d R_{c_1c_2}{}^{d_1d_2} \tag{6.7.11}$$

$$\delta\tilde{P}_{\Phi\Phi}(\xi) = 0. \tag{6.7.12}$$

Note that for this theory the "weak field" conditions reduce to

$$|F'(\Phi)|L^{-2} \ll 1 \qquad |F''(\Phi)|L^{-2} \ll 1. \tag{6.7.13}$$

Following the approach outlined in Section 6.5 we will study $\ker\delta P(\xi)$, for null $\xi$, to determine whether this theory can be strongly hyperbolic. Recall that a vector $(t_{ab}, \chi)^T$ is in $\ker\delta P(\xi)$ if, and only if,

$$\begin{pmatrix} \delta P_{gg}(\xi)^{abcd}t_{cd} + \delta P_{g\Phi}(\xi)^{ab}\chi \\ \delta P_{\Phi g}(\xi)^{cd}t_{cd} + \delta P_{\Phi\Phi}(\xi)\chi \end{pmatrix} = 0. \tag{6.7.14}$$

Let $\xi$ be a null covector and look at the first row of (6.7.14),

$$\delta P_{gg}(\xi)^{abcd}t_{cd} + \delta P_{g\Phi}(\xi)^{ab}\chi = 0. \tag{6.7.15}$$

If this equation admitted a solution, then it could be used to fix $t_{ab}$ as a function of $\chi$, up to addition of linear combinations of elements of $\ker\delta P_{gg}(\xi)$. It follows from the symmetries of the principal symbol that "pure gauge" vectors belong to this kernel [44]. If these were the only elements, then the above equation would fix completely the non-gauge part of $t_{ab}$, implying that $\ker\delta P(\xi)$ will contain at most one non-gauge element (and thus five elements at most), violating the necessary condition for strong hyperbolicity. Hence, for strong hyperbolicity to hold it is necessary that $\ker\delta P_{gg}(\xi)$ contains non-gauge elements. We will therefore proceed to calculate $\ker\delta P_{gg}(\xi)$. Defining, for convenience, the tensor

$$\mathcal{F}_{ab} = -2(F''(\Phi)\nabla_a\Phi\nabla_b\Phi + F'(\Phi)\nabla_a\nabla_b\Phi) \tag{6.7.16}$$

we can rewrite the condition for a tensor $r_{ab}$ to be in $\ker\delta P_{gg}(\xi)$ — that is to say $\delta P_{gg}(\xi)^{abcd}r_{cd} = 0$ — as

$$\delta^{acc_1c_2}_{bdd_1d_2}\xi_{c_1}\xi^{d_1}r_{c_2}{}^{d_2}\mathcal{F}_c{}^d - \frac{1}{2}f(f+2)\delta^{ac_1c_2}_{bd_1d_2}\xi_{c_1}\xi^{d_1}r_{c_2}{}^{d_2} = 0. \tag{6.7.17}$$

In order to find solutions to this equation, we fix a point in spacetime and introduce a null basis $\{e_0, e_1, e_i\}$ for the tangent space at that point. We take this basis to be adapted to $\xi$, i.e. $\xi^\mu = e_0^\mu$, $e_0 \cdot e_0 = e_1 \cdot e_1 = 0$, $e_0 \cdot e_1 = 1$, $e_i \cdot e_j = \delta_{ij}$ and $e_0 \cdot e_i = e_1 \cdot e_i = 0$. In this basis, the system $(\delta P_{gg}(\xi) \cdot r)_{\mu\nu} = 0$ reduces to

$$\frac{1}{2} f(f+2)(r_{22} + r_{33}) - \mathcal{F}_{33} r_{22} + 2\mathcal{F}_{23} r_{23} - \mathcal{F}_{22} r_{33} = 0 \tag{6.7.18a}$$

$$-\frac{1}{2} f(f+2) r_{02} + \mathcal{F}_{33} r_{02} - \mathcal{F}_{23} r_{03} - \mathcal{F}_{03} r_{23} + \mathcal{F}_{02} r_{33} = 0 \tag{6.7.18b}$$

$$-\frac{1}{2} f(f+2) r_{03} + \mathcal{F}_{22} r_{03} - \mathcal{F}_{23} r_{02} + \mathcal{F}_{03} r_{22} - \mathcal{F}_{02} r_{23} = 0 \tag{6.7.18c}$$

$$\frac{1}{2} f(f+2) r_{00} - \mathcal{F}_{33} r_{00} + 2\mathcal{F}_{03} r_{03} - \mathcal{F}_{00} r_{33} = 0 \tag{6.7.18d}$$

$$\frac{1}{2} f(f+2) r_{00} - \mathcal{F}_{22} r_{00} + 2\mathcal{F}_{02} r_{02} - \mathcal{F}_{00} r_{22} = 0 \tag{6.7.18e}$$

$$\mathcal{F}_{23} r_{00} - \mathcal{F}_{03} r_{02} - \mathcal{F}_{02} r_{03} + \mathcal{F}_{00} r_{23} = 0. \tag{6.7.18f}$$

Note that the "gauge" components of $r_{ab}$, i.e. $r_{1\mu}$, do not appear in the equations. This is a system of six linear equations for the six non-gauge components of $r_{ab}$. Since the number of unknowns equals the number of equations, this system will have no non-trivial solution unless the determinant of the matrix of coefficients vanishes. We will now show that this determinant does not vanish for any choice of generalised Harmonic gauge.

The matrix of coefficients takes the form

$$C = \begin{pmatrix} 0 & 0 & 0 & \frac{f(f+2)}{2} - \mathcal{F}_{33} & 2\mathcal{F}_{23} & \frac{f(f+2)}{2} - \mathcal{F}_{22} \\ 0 & \mathcal{F}_{33} - \frac{f(f+2)}{2} & -\mathcal{F}_{23} & 0 & -\mathcal{F}_{03} & \mathcal{F}_{02} \\ 0 & -\mathcal{F}_{23} & \mathcal{F}_{22} - \frac{f(f+2)}{2} & \mathcal{F}_{03} & -\mathcal{F}_{02} & 0 \\ \frac{f(f+2)}{2} - \mathcal{F}_{33} & 0 & 2\mathcal{F}_{03} & 0 & 0 & -\mathcal{F}_{00} \\ \mathcal{F}_{23} & -\mathcal{F}_{03} & -\mathcal{F}_{02} & 0 & \mathcal{F}_{00} & 0 \\ \frac{f(f+2)}{2} - \mathcal{F}_{22} & 2\mathcal{F}_{02} & 0 & -\mathcal{F}_{00} & 0 & 0 \end{pmatrix}. \tag{6.7.19}$$

In the null basis, its determinant reads

$$\det C = -\frac{1}{8} \Big[ \mathcal{F}_{00} (f(f+2))^2 + f(f+2) \left( -2\mathcal{F}_{00}\mathcal{F}_{22} - 2\mathcal{F}_{00}\mathcal{F}_{33} + 2\mathcal{F}_{02}^2 + 2\mathcal{F}_{03}^2 \right) \\ + 4\mathcal{F}_{00}\mathcal{F}_{22}\mathcal{F}_{33} - 4\mathcal{F}_{00}\mathcal{F}_{23}^2 - 4\mathcal{F}_{02}^2\mathcal{F}_{33} + 8\mathcal{F}_{02}\mathcal{F}_{03}\mathcal{F}_{23} - 4\mathcal{F}_{03}^2\mathcal{F}_{22} \Big]^2. \tag{6.7.20}$$

The condition $\det C = 0$ can be rewritten covariantly and is equivalent to the following

$$f^2(f+2)^2 \xi^c \xi^d \mathcal{F}_{cd} - 2f(f+2)(\delta^{c_1c_2c_3}_{d_1d_2d_3} \xi_{c_1} \xi^{d_1} \mathcal{F}_{c_2}{}^{d_2} \mathcal{F}_{c_3}{}^{d_3}) \\ - \frac{8}{3} \delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4} \xi_{c_1} \xi^{d_1} \mathcal{F}_{c_2}{}^{d_2} \mathcal{F}_{c_3}{}^{d_3} \mathcal{F}_{c_4}{}^{d_4} = 0. \tag{6.7.21}$$

Solving for $f$ gives $f = f_*$, with[65]

$$f_* = -1 + \sqrt{(1 - 2\mathcal{F}^e{}_e) + \frac{2\mathcal{F}_{ae}\mathcal{F}_b{}^e\xi^a\xi^b \pm \sqrt{A}}{\mathcal{F}_{cd}\xi^c\xi^d}} \tag{6.7.22}$$

where

$$A = (\delta^{c_1c_2c_3}_{d_1d_2d_3}\xi_{c_1}\xi^{d_1}\mathcal{F}_{c_2}{}^{d_2}\mathcal{F}_{c_3}{}^{d_3})^2 + \frac{8}{3}\mathcal{F}_b{}^a\xi_a\xi^b\delta^{c_1c_2c_3c_4}_{d_1d_2d_3d_4}\xi_{c_1}\xi^{d_1}\mathcal{F}_{c_2}{}^{d_2}\mathcal{F}_{c_3}{}^{d_3}\mathcal{F}_{c_4}{}^{d_4}. \tag{6.7.23}$$

Note that $f$ should only depend on background fields, it cannot depend on $\xi$. The function $f_*$ could only be independent of $\xi$ if the second term in the square root in the above expression were independent of $\xi$; that is if

$$2\mathcal{F}_{ae}\mathcal{F}_b{}^e\xi^a\xi^b \pm \sqrt{A} = \lambda\mathcal{F}_{ab}\xi^a\xi^b \tag{6.7.24}$$

for some scalar $\lambda$ independent of $\xi$. By rearranging the terms and squaring them, we see that we can equivalently look for a $\lambda$ which solves (expanding $A$ and using $\xi^2 = 0$)

$$\begin{aligned}\mathcal{F}_{ab}\xi^a\xi^b[&\mathcal{F}_{cd}\xi^c\xi^d(\lambda^2 - 8\mathcal{F}_{ef}\mathcal{F}^{ef} + 4\mathcal{F}^e{}_e\mathcal{F}^f{}_f) \\ &- 4\mathcal{F}_c{}^e\mathcal{F}_{de}\xi^c\xi^d(\lambda + 2\mathcal{F}^e{}_e) + 16\mathcal{F}^{ef}\mathcal{F}_{ce}\mathcal{F}_{df}\xi^c\xi^d] = 0.\end{aligned} \tag{6.7.25}$$

However, since the three terms in square parenthesis have different dependence on $\xi$, they would have to vanish independently and there is no choice of $\lambda$ for which this happens in a generic background. Note that there is no special choice of $F(\Phi)$ for which this result would be different. To see this, we substitute the explicit form of $\mathcal{F}_{ab}$ in Eq. (6.7.25); the last term will give rise to a term of the form

$$(F'(\Phi))^3\nabla^e\nabla^f\Phi\nabla_c\nabla_e\Phi\nabla_d\nabla_f\Phi\xi^c\xi^d. \tag{6.7.26}$$

Since there is no other term involving $\xi$ contracted with the same combination of derivatives of $\Phi$, for (6.7.25) to hold for $\lambda$ independent of $\xi$ it is necessary that this term vanishes in a generic background, i.e. $F'(\Phi) = 0$. However, for such choice of $F(\Phi)$, EdGB theory would reduce to GR.

We can deduce from this argument that any $f_*$ which solves $\det C = 0$ would necessarily depend on $\xi$, which is not allowed. This implies that for any "good" choice of the function $f$, in a generic background (for which $\mathcal{F}_{ab}$ does not have any special properties) the system (6.7.18) has no non-trivial solution. Therefore we can conclude that, in a generic weak field background, for any choice of generalised harmonic gauge, the only elements of $\ker \delta P_{gg}(\xi)$ are "pure gauge" vectors, i.e. $r_{ab} = \xi_{(a}X_{b)}$.

[65]Note that there is only one choice of sign in front of the square root for which $f_*$ satisfies the smallness assumptions (6.2.31).

Going back to our original question, this result implies that if a $t_{ab}$ solving (6.7.15) exists, then this solution will be unique up to the addition of multiples of "pure gauge" vectors; that is (6.7.15) completely fixes the non-gauge part of $t_{ab}$ in terms of $\chi$. If we then substituted such $t_{ab}$ in the second row of (6.7.14), we would obtain a linear homogeneous equation for $\chi$ which, for a generic background, would only admit the solution $\chi = 0$. This would in turn imply that the "non-gauge" part of $t_{ab}$ had to vanish, i.e. $t_{ab} = \xi_{(a}Y_{b)}$. Hence we can conclude that, for any choice of generalised harmonic gauge, in a generic "weak field" background we have that $\dim\ker \delta P(\xi) = 4$ and thus the linearised EdGB equations are not strongly hyperbolic in this setting, since Condition 2 is not satisfied.

## 6.8 Strong hyperbolicity: the general case

We will now investigate whether general Horndeski theories are strongly hyperbolic or they suffer from problems similar to EdGB theory.

### The $\mathcal{G}_5 = \mathcal{G}_5(\Phi)$ case

Consider first the case in which $\mathcal{G}_5 = \mathcal{G}_5(\Phi)$. The corresponding contribution to the Horndeski Lagrangian will read

$$\mathcal{L}_5 = \mathcal{G}_5(\Phi) G_{ab} \nabla^a \nabla^b \Phi. \tag{6.8.1}$$

However, it can be shown that this is equivalent — up to a total derivative term, which does not contribute to the equations of motion — to [31, 106]

$$\mathcal{L}_5 = -\partial_\Phi \mathcal{G}_5 X R - \partial_\Phi \mathcal{G}_5 \delta^{ac}_{bd} \nabla_a \nabla^b \Phi \nabla_c \nabla^d \Phi + 3\partial^2_\Phi \mathcal{G}_5 X \Box \Phi - 2\partial^3_\Phi \mathcal{G}_5 X^2, \tag{6.8.2}$$

and hence the full Horndeski Lagrangian can be rewritten as

$$\mathcal{L} = \mathcal{L}_1 + \tilde{\mathcal{L}}_2 + \tilde{\mathcal{L}}_3 + \tilde{\mathcal{L}}_4 \tag{6.8.3}$$

where

$$\tilde{\mathcal{G}}_2 = \mathcal{G}_2 - 2\partial^3_\Phi \mathcal{G}_5 X^2 \qquad \tilde{\mathcal{G}}_3 = \mathcal{G}_3 + 3\partial^2_\Phi \mathcal{G}_5 X \qquad \tilde{\mathcal{G}}_4 = \mathcal{G}_4 - \partial_\Phi \mathcal{G}_5 X. \tag{6.8.4}$$

As this is effectively equivalent to an Horndeski theory with $\tilde{\mathcal{G}}_5 = 0$, it will be strongly hyperbolic around a generic "weak field" background, in a generalised harmonic gauge if, and only if, $\partial_X \tilde{\mathcal{G}}_4 = 0 = \tilde{\mathcal{G}}_5$, i.e. if, and only if,

$$\partial_X \mathcal{G}_4 = \partial_\Phi \mathcal{G}_5. \tag{6.8.5}$$

However, with this choice the theory simply reduces to one of those discussed in Section 6.6. In order to avoid this degeneracy, we will need $\mathcal{G}_5$ to depend on $X$.

### The $\mathcal{G}_5 = \mathcal{G}_5(\Phi, X)$ case

Let us consider the most general Horndeski theory, i.e. $\mathcal{G}_i \neq 0$ and $\mathcal{G}_5 = \mathcal{G}_5(\Phi, X) \neq 0$. The equations of motion and the principal symbol for this theory are reported in Chapter 3.

The main obstruction to strong hyperbolicity in EdGB theory, as discussed in the previous section, arose from the fact that the dimension of the kernel of $\delta P_{gg}(\xi)$ was not large enough. We will therefore proceed to study the corresponding operator in the general Horndeski theory. From Eq. (3.3.16a) we see that, for null $\xi$, it takes the form

$$
\begin{aligned}
(\delta P_{gg}(\xi)\cdot t)^a{}_b = &\left(F_1(\Phi, X) - \frac{1}{2} f(f+2)\right)\delta^{ac_1c_2}_{bd_1d_2}\xi_{c_1}\xi^{d_1}t_{c_2}{}^{d_2} \\
&+ F_2(\Phi, X)\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\xi_{c_1}\xi^{d_1}t_{c_2}{}^{d_2}\nabla_{c_3}\Phi\nabla^{d_3}\Phi \\
&+ F_3(\Phi, X)\delta^{ac_1c_2c_3}_{bd_1d_2d_3}\xi_{c_1}\xi^{d_1}t_{c_2}{}^{d_2}\nabla_{c_3}\nabla^{d_3}\Phi
\end{aligned} \tag{6.8.6}
$$

where

$$
F_1(\Phi, X) = -\frac{1}{2}(\mathcal{G}_4 - 2X\partial_X\mathcal{G}_4 + X\partial_\Phi\mathcal{G}_5) \tag{6.8.7}
$$

$$
F_2(\Phi, X) = -\frac{1}{2}(\partial_X\mathcal{G}_4 - \partial_\Phi\mathcal{G}_5) \tag{6.8.8}
$$

$$
F_3(\Phi, X) = \frac{1}{2}X\partial_X\mathcal{G}_5. \tag{6.8.9}
$$

We will assume $F_3 \neq 0$, for otherwise the theory would reduce to the case $\mathcal{G}_5 = \mathcal{G}_5(\Phi)$ discussed earlier. Consider now the condition (6.7.14) for a vector $T = (t_{ab}, \chi)^T$ to belong to $\ker \delta P(\xi)$. The first row reads

$$
\delta P_{gg}(\xi)^{abcd}t_{cd} + \delta P_{g\Phi}(\xi)\chi = 0. \tag{6.8.10}
$$

In the EdGB case, we used the fact that $\ker \delta P_{gg}(\xi)$ only contained "pure gauge" elements to conclude that if this equation admitted a solution then it would fix completely the "non-gauge" part of $t_{ab}$ in terms of $\chi$, implying that the dimension of $\ker \delta P(\xi)$ could not be large enough for strong hyperbolicity to hold. We will now show that the same statement holds in the general case. Consider the equation

$$
(\delta P_{gg}(\xi)\cdot r)_{ab} = 0. \tag{6.8.11}
$$

Its tensorial structure is essentially identical to that of the corresponding equation in EdGB theory, Eq. (6.7.17). In fact, defining

$$
\tilde{\mathcal{F}}_{ab} = F_2(\Phi, X)\nabla_a\Phi\nabla_b\Phi + F_3(\Phi, X)\nabla_a\nabla_b\Phi \tag{6.8.12}
$$

$$
\tilde{f} = -1 + \sqrt{(1+f)^2 - 2F_1} \tag{6.8.13}
$$

Eq. (6.8.11) takes the form

$$\delta^{acc_1c_2}_{bdd_1d_2}\xi_{c_1}\xi^{d_1}r_{c_2}{}^{d_2}\tilde{\mathcal{F}}_c{}^d - \frac{1}{2}\tilde{f}(\tilde{f}+2)\delta^{ac_1c_2}_{bd_1d_2}\xi_{c_1}\xi^{d_1}r_{c_2}{}^{d_2} = 0. \tag{6.8.14}$$

Note that even if $F_1 = 0$ or $F_2 = 0$, the form of this equation would be unchanged and hence we do not need to consider these cases separately. We can study this system in the same way as we studied (6.7.18). Recall that this is a system of six equations for the six "non-gauge" components of $r_{ab}$ (the "pure gauge" components do not appear in these equations). An analogous argument allows us to conclude that there is no admissible choice of $f$ for which this system would generically admit non-trivial solutions.[66] Therefore we can deduce that for a generic "weak field" background and for any choice of generalised harmonic gauge, $\ker \delta P_{gg}(\xi)$ will only contain "pure gauge" elements.

Since $\ker \delta P_{gg}(\xi)$ contains only "pure gauge" elements, then if a solution to (6.8.10) exists it will be unique up to addition of multiples of "pure gauge" vectors. In other words, (6.8.10) completely fixes the non-gauge part of a $t_{ab}$ in $\ker \delta P(\xi)$ in terms of $\chi$. If we then substitute such $t_{ab}$ into the second row of (6.7.14), we will obtain a linear homogeneous equation for $\chi$, which for generic background has only trivial solutions $\chi = 0$. This implies that all elements in $\ker \delta P(\xi)$ have vanishing "non-gauge" part, i.e. $\ker \delta P(\xi)$ only contains "pure gauge" elements.

Finally, since $\dim \ker \delta P(\xi) = 4 < 8$, we conclude that the necessary condition for strong hyperbolicity — i.e., Condition 2 — is not satisfied and hence the most general Horndeski theory fails to be strongly hyperbolic in a generic "weak field" background for any choice of generalised harmonic gauge.

[66] In this setting, Eq. (6.7.25) would have a solution independent of $\xi$ if, and only if, $F_3 = 0$. However, since we are assuming $F_3 \neq 0$, this does not happen. To see that this condition is necessary, we could repeat the argument following Eq. (6.7.25) where now $F_3$ plays the role of $F'$ in Eq. (6.7.26).

## 6.9 Summary of results

In summary, we have proven that all Horndeski theories are *weakly hyperbolic* around a weak field background, for any choice of generalised harmonic gauge.

We further showed that any Horndeski theory for which $\mathcal{G}_5 \neq 0$ or $\partial_X \mathcal{G}_4 \neq 0$ fails to be strongly hyperbolic around a weak field background, for any choice of generalised harmonic gauge. On the other hand, we proved that there exists a *unique* choice of generalised harmonic gauge, Eq. (6.6.52), which makes the class of Horndeski theories with $\partial_X \mathcal{G}_4 = \mathcal{G}_5 = 0$, i.e.,

$$\mathcal{L} = (1 + \mathcal{G}_4(\Phi))R + X - V(\Phi) + \mathcal{G}_2(\Phi, X) + \mathcal{G}_3(\Phi, X)\Box\Phi. \tag{6.9.1}$$

strongly hyperbolic around a weak field background.

### Causality

Causal properties of theories of the form (6.9.1) have been discussed in Ref. [27].[67] It is interesting to discuss causality using our results above. We showed above that, in an appropriate generalised harmonic gauge, a null co-vector $\xi_a$ is characteristic if, and only if, either $g^{ab}\xi_a\xi_b = 0$ or $f^{ab}\xi_a\xi_b = 0$, where $f^{ab}$ is defined by (6.6.55). Furthermore, if $\xi_a$ satisfies the former condition then $P(\xi)$ generically has a 10-dimensional kernel consisting of vectors of the form $(t_{ab}, 0)$ for general $t_{ab}$, whereas if $\xi_a$ satisfies the latter condition then $P(\xi)$ generically has a 1-dimensional kernel consisting of vectors of the form $(t_{ab}, 1)$ with $t_{ab}$ given by (6.6.54). Hence, roughly speaking, causality for the 10 tensor degrees of freedom is determined by $g_{ab}$ whereas causality for the 1 scalar degree of freedom is determined by $f_{ab}$, the inverse of $f^{ab}$. This agrees with Ref. [27]. Of course these degrees of freedom are coupled together so causality for the theory as a whole is determined by both metrics $g_{ab}$ and $f_{ab}$. More precisely, the characteristic surfaces of the theory are surfaces which are null w.r.t. either $g_{ab}$ or $f_{ab}$.

### Non-linear considerations

The above discussion shows that there exists a preferred generalised harmonic gauge (6.6.52) for which a theory of the form (6.9.1) is strongly hyperbolic when linearised around a generic weak field background. We can now ask: does this generalised harmonic gauge condition for the linearised theory arise by linearising a generalised harmonic gauge condition for the non-linear theory?

[67] Ref. [27] assumed $\mathcal{G}_4 = 0$ but for a theory of the form (6.9.1) we can always set $\mathcal{G}_4 = 0$ using a field redefinition, specifically a conformal transformation.

It is important to address this question for, in the end, what we are interested in is the well-posedness of the full non-linear problem. As discussed earlier, for the non-linear problem to be well-posed it is necessary that the all the problems obtained by linearising it (i.e. linearising both the equations of motion and the gauge condition) be well-posed. We saw in the previous sections that the requirement of strong hyperbolicity forced a unique choice of generalised harmonic gauge at the linear level. If such a choice of gauge cannot be realised by linearising a corresponding non-linear gauge condition, we can conclude that there is no choice of generalised harmonic gauge which makes the non-linear problem well-posed.

Consider a non-linear generalised harmonic gauge condition of the form

$$\frac{1}{\sqrt{-g}}\partial_\nu\left(\sqrt{-g}g^{\mu\nu}\right) = J^\mu(g,\Phi,\partial\Phi). \tag{6.9.2}$$

Note that we would not want $J^\mu$ to depend on second or higher derivatives of $\Phi$ because this would give a gauge-fixed equation of motion involving third derivatives of $\Phi$.

Linearising around a general background solution gives

$$\nabla_\nu h^{\mu\nu} - \frac{1}{2}\nabla^\mu h^\nu{}_\nu + \frac{\partial J^\mu}{\partial(\partial_\nu\Phi)}\partial_\nu\psi = \dots \tag{6.9.3}$$

where the ellipsis denotes terms that do not involve derivatives of $h_{\mu\nu}$ or $\psi$ and therefore do not influence hyperbolicity. Comparing with (6.2.7) we see that the linearised gauge condition has

$$\frac{\mathcal{H}^{\mu\nu}}{1+f} = -\frac{\partial J^\mu}{\partial(\partial_\nu\Phi)}. \tag{6.9.4}$$

It follows that the functions appearing in the linearised gauge condition must satisfy the integrability condition

$$\frac{\partial}{\partial(\partial_\rho\Phi)}\left(\frac{\mathcal{H}^{\mu\nu}}{1+f}\right) = \frac{\partial}{\partial(\partial_\nu\Phi)}\left(\frac{\mathcal{H}^{\mu\rho}}{1+f}\right). \tag{6.9.5}$$

Plugging in the functions (6.6.52), this equation reduces to

$$\partial_X\mathcal{G}_3\left(g^{\mu\nu}\partial^\rho\Phi - g^{\mu\rho}\partial^\nu\Phi\right) = 0. \tag{6.9.6}$$

By contracting this equation it is easy to see that the only way this can hold in a generic background is if $\partial_X\mathcal{G}_3 = 0$. But if $\mathcal{G}_3$ is independent of $X$, then $\mathcal{L}_3$ is equivalent (up to a total derivative) to $\mathcal{L}_2$.[68] In other words, requiring the above

[68]Explicitly, neglecting terms which do not contribute to the equations of motion, we have $\mathcal{L}_3 = \Box\Phi\mathcal{G}_3(\Phi) = 2X\partial_\Phi\mathcal{G}_3(\Phi)$. We see that this is equivalent to $\mathcal{L}_2$ with $\mathcal{G}_2(\Phi,X) = 2X\partial_\Phi\mathcal{G}_3(\Phi)$.

equation to hold implies $\mathcal{G}_3 = 0$. If $\mathcal{G}_3 = 0$ then we can find a source function $J^\mu$ consistent with equation (6.9.4):

$$J^\mu = -\frac{\partial_\Phi \mathcal{G}_4}{1 + \mathcal{G}_4} \partial^\mu \Phi. \tag{6.9.7}$$

In summary, we have imposed the requirement that the preferred generalised harmonic gauge condition for the linearised theory arises by linearising a generalised harmonic gauge condition for the non-linear theory. The result is that this requirement excludes theories with non-trivial $\mathcal{G}_3$. So demanding that there exists a generalised harmonic gauge for which the non-linear theory is strongly hyperbolic in a generic weak-field background restricts the theory to one of the form

$$\mathcal{L}_* = (1 + \mathcal{G}_4(\Phi))R + X - V(\Phi) + \mathcal{G}_2(\Phi, X). \tag{6.9.8}$$

Since $\mathcal{G}_4$ can be eliminated by a field redefinition (footnote 67), this theory is equivalent to Einstein gravity coupled to a "k-essence" theory. With the gauge choice (6.9.7), this theory is not just strongly hyperbolic, it is *symmetric* hyperbolic (see Footnote 63 and Definition 4, page 16).

# Chapter 7

# Conclusions

In this dissertation we studied certain mathematical properties of two important classes of Modified Gravity theories: Lovelock and Horndeski theories.

In the first part of the thesis, Chapter 4, we studied aspects of the causal structure of Lovelock theories in static, spherically symmetric backgrounds.

The core of the thesis, Chapters 5 and 6, has been devoted to the study of the initial value problem for these theories.

We will now summarise and discuss the main results.

### Causality

In Chapter 4 we studied the propagation of high frequency gravitons on static, spherically symmetric black hole spacetimes in Einstein–Gauss–Bonnet gravity. In this theory gravitons do not necessarily propagate at the speed of light. In particular, they can propagate faster than light. It was recently argued that, thanks to this property, gravitons could experience a negative Shapiro time delay and that this fact could be used to construct a "time machine", implying that these theories could violate causality. These arguments, however, need to be treated with care: the negative Shapiro time delay was observed in a singular shock-wave geometry, while the argument for the construction of the "time machine" is not obviously realisable dynamically.

Our main result was to confirm that gravitons with certain polarisations can indeed experience a negative time delay when scattering off a regular geometry, namely a small, spherically symmetric black hole spacetime. Interestingly, we observed that gravitons of such polarisation feel a repulsive gravitational force at distances comparable to the length scale set by the coupling constants. In particular, we found that for certain values of the impact parameter a graviton can suffer no net deflection, implying that it is possible for a graviton to experience a negative time delay when scattering from and to infinity! We then examined the "time machine" construction and we showed that it is not possible to realise it as the evolution of "good" initial data. The construction relies on the possibility of boosting a black hole to a speed arbitrarily close to the speed of light. We showed that one cannot boost an object arbitrarily fast as this would cause the leaves of "constant time" to become non-spacelike, hence causing time-evolution to break down. This implies that the construction cannot be physically realised and hence that closed causal curves cannot form in this way, suggesting that these theories may, in fact, not be causally pathological.

## Initial value problem

The main results of this thesis concern the well-posedness of the initial value problem for Lovelock and Horndeski theories.

We have shown that, in harmonic gauge, the linearised equation of motion of a Lovelock theory is always weakly hyperbolic in a weakly curved background. However, it is not strongly hyperbolic in a generic weak-field background.

We have shown that, in a generalised harmonic gauge, the linearised equation of motion of a Horndeski theory is always weakly hyperbolic in a weak-field background. For some Horndeski theories, a generalised harmonic gauge can be found for which the linearised equation of motion is also strongly hyperbolic in a weak field background. In particular this is true for theories of the form (6.9.1). However, for more general Horndeski theories, we have shown that there is no generalised harmonic gauge for which the equation of motion is strongly hyperbolic in a generic weak-field background. Furthermore, even for theories of the form (6.9.1), imposing the requirement that the gauge condition for the linearised theory is the linearisation of a generalised harmonic gauge condition for the non-linear theory restricts the theory further, to one of the form (6.9.8):[69]

$$\mathcal{L}_* = (1 + \mathcal{G}_4(\Phi))R + X - V(\Phi) + \mathcal{G}_2(\Phi, X).$$

Without strong hyperbolicity, the best one can hope for is that the linearised equation of motion is locally well-posed with a "loss of derivatives". This means that the $k$-th Sobolev norm $H^k$ of the fields at time $t$ cannot be bounded in terms of its initial value but only in terms of the initial value of some higher Sobolev norm $H^{k+l}$ with $l > 0$. Whether or not even this can be done depends on the nature of the terms with fewer than two derivatives in the equation of motion [53]. But even if this can

[69]Note that this coincides with the class of Horndeski theories which has not been ruled out by the neutron star merger GW170817 [37, 38, 39, 40].

be achieved, the loss of derivatives is likely to be fatal for any attempt to prove that the *non-linear* equation is locally well-posed in some Sobolev space, as is the case for the Einstein equation.[70] This is because establishing well-posedness for a non-linear equation usually involves a "bootstrap" argument in which one assumes some bound on the $H^k$ norm and then uses the energy estimate to improve this bound, thereby closing the bootstrap. This is not possible if the energy estimate exhibits a loss of derivatives.

Note that our result is a statement about the *full* equations of motion. If one restricts the equations of motion by imposing some symmetry on the solution (e.g. spherical symmetry) then it is possible that the resulting equations might be strongly hyperbolic. This is because the resulting class of background spacetimes would be non-generic and, as we have seen, for non-generic backgrounds it is possible for the equation of motion to be strongly hyperbolic even if it is not strongly hyperbolic for a generic background.

Our results demonstrate that we do not have local well-posedness for the harmonic gauge Lovelock equation of motion for *general* initial data. So the situation is worse than for the Einstein equation, for which the harmonic gauge equation of motion is locally well-posed for any initial data [42]. But in practice we are not interested in general initial data, but only in initial data satisfying the harmonic gauge condition. Since the failure of strong hyperbolicity appears to be associated to modes which violate the harmonic gauge condition, perhaps we could restrict to initial data satisfying this condition *exactly* and thereby obtain a well-posed problem. One could not do this numerically on a computer because the gauge condition could never be imposed exactly – there would always be numerical error. But perhaps this could be done in principle. One way to proceed would be to consider sequences of analytic initial data, satisfying the gauge condition, which approach some specified smooth initial data. For analytic data one can solve the equation of motion locally [77]. If one could prove that the resulting analytic solution satisfies an energy estimate without a loss of derivatives (because it satisfies the gauge condition), then perhaps it would be possible to establish local well-posedness. Having said this, we note that one could make exactly the same remarks about the Einstein equation written in a "bad" (non-strongly hyperbolic) gauge so it is far from clear that this method has any chance of succeeding.

[70] It is conceivable that one might have local well-posedness in some much more restricted function space, such as a Gevrey space. This class of functions consists of $C^\infty$ functions whose successive derivatives satisfy inequalities weaker than those required for the convergence of a Taylor series. The important property of such functions is that they are "almost" analytic but they are not determined by their values in the neighbourhood of a point [49].

If the equation of motion is not strongly hyperbolic in (generalised) harmonic gauge then could there be some other gauge in which it is strongly hyperbolic? For example, maybe one could modify the (generalised) harmonic gauge condition to include additional terms involving first derivatives of $h_{ab}$, contracted in some way with the background curvature tensor (or scalar field). But this raises the question of whether it is always possible to impose the new gauge condition via a gauge transformation (cf. Eqs. (6.2.8)–(6.2.10)). This would involve solving an equation for the gauge parameters. We would then have to analyse whether this new equation has a well-posed initial value problem, and whether the resulting gauge condition is propagated by the gauge-fixed equation of motion. This may amount to analysing equations that suffer from the same kind of problems as the equations we have discussed in this Thesis.

In this Thesis, we have been working with equations of motion for the metric. An alternative approach would be to derive an equation of motion for curvature. The Bianchi identity can be used to write $\nabla^e \nabla_e R_{abcd}$ in terms of second derivatives of the Ricci tensor, and terms with fewer than two derivatives of curvature. For the Einstein equation, one can eliminate the Ricci tensor to obtain a non-linear wave equation for the Weyl tensor. This equation is strongly hyperbolic and admits a well-posed initial value problem. For a Lovelock theory one cannot solve explicitly for the Ricci tensor but one could still replace the Ricci tensor terms using the expression obtained from the equation of motion of the theory. This gives an equation of motion for the Riemann tensor. In contrast with what happens for the Einstein equation, the resulting equation is subject to a constraint, which is simply the Lovelock equation of motion. If this constraint is satisfied by the initial data then it will be satisfied by any solution of the equation of motion for the Riemann tensor. The situation looks analogous to the case of the harmonic gauge equation of motion for the metric, but with more indices. It seems very likely that this equation of motion for the Riemann tensor will fail to be strongly hyperbolic in a generic background.

Another approach would be to investigate equations of motion based on a space-time decomposition of the metric, as in the ADM formalism. It is known that the ADM formulation of the Einstein equation gives equations that are not strongly hyperbolic [47]. However, suitable modification of the ADM method gives equations that *are* strongly hyperbolic [47, 61]. Perhaps something similar would work for Lovelock or Horndeski theories. However, it appears that there is no obvious way of extending the approaches used for the Einstein equation to Lovelock theories [107].

Moreover, note that the way we obtained the generalised harmonic gauge equations is by no means unique. One could, for example, use different $\mathcal{H}_{ab}$ in the metric and scalar field gauge fixing terms. However, fixing the gauge in this way would break the symmetry of the principal symbol: $P_{g\Phi} \neq P_{\Phi g}$. Our proof of weak hyperbolicity relies crucially on this symmetry and therefore it seems unlikely that such modification of the gauge fixing term could be useful.

Finally, there is also the possibility that these theories may only be well-posed for initial data belonging to some highly restrictive class. In fact, the situation could be even worse: these theories might *not* admit a well-posed initial value problem at all! This scenario would likely be fatal for these theories since, as discussed several times in this thesis, any sensible classical theory must admit a well-posed initial value problem. In other words, this would lead to the satisfying conclusion that these theories are *unviable* as physical alternatives to General Relativity.

# Bibliography

[1] A. Einstein, *Die Feldgleichungen der Gravitation*, *Sitzungsberichte der Königlich Preußischen Akad. der Wissenschaften* (1915) 844–847.

[2] A. Einstein, *Die Grundlage der allgemeinen Relativitätstheorie*, *Ann. Phys.* **354** (1916) 769–822.

[3] F. W. Dyson, A. S. Eddington and C. Davidson, *A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919*, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **220** (1920) 291–333.

[4] The LIGO Collaboration and the VIRGO Collaboration, *Observation of Gravitational Waves from a Binary Black Hole Merger*, *Phys. Rev. Lett.* **116** (2016) 061102, [1602.03837].

[5] R. Penrose, *Gravitational collapse and space-time singularities*, *Phys. Rev. Lett.* **14** (1965) 57–59, [1001.4815].

[6] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time*. Cambridge University Press, 1973.

[7] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich et al., *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, *Astron. J.* **116** (1998) 1009–1038, [astro-ph/9805201].

[8] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro et al., *Measurements of* $\Omega$ *and* $\Lambda$ *from 42 High-Redshift Supernovae*, *Astrophys. J.* **517** (1999) 565–586, [astro-ph/9812133].

[9] A. Einstein, *Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie*, *Sitzungsberichte der Königlich Preußischen Akad. der Wissenschaften* (1917) 142–152.

[10] S. Weinberg, *The cosmological constant problem*, *Rev. Mod. Phys.* **61** (1989) 1–23, [astro-ph/0005265].

[11] D. Lovelock, *The Four-Dimensionality of Space and the Einstein Tensor*, *J. Math. Phys.* **13** (1972) 874–876.

[12] D. Lovelock, *The Einstein Tensor and its Generalizations*, *J. Math. Phys.* **12** (1971) 498–501.

[13] T. Kaluza, *Zum Unitätsproblem der Physik*, *Sitzungsberichte der Königlich Preußischen Akad. der Wissenschaften* (1921) 966–972.

[14] O. Klein, *Quantentheorie und fünfdimensionale Relativitätstheorie*, *Zeitschrift für Phys.* **37** (1926) 895–906.

[15] O. Klein, *The Atomicity of Electricity as a Quantum Theory Law*, *Nature* **118** (1926) 516–516.

[16] M. B. Green and J. H. Schwarz, *Anomaly cancellations in supersymmetric $D = 10$ gauge theory and superstring theory*, *Phys. Lett. B* **149** (1984) 117–122.

[17] B. Zwiebach, *Curvature squared terms and string theories*, *Phys. Lett. B* **156** (1985) 315–317.

[18] R. Metsaev and A. Tseytlin, *Order $\alpha'$ (two-loop) equivalence of the string equations of motion and the $\sigma$-model Weyl invariance conditions*, *Nucl. Phys. B* **293** (1987) 385–419.

[19] J. Maldacena, *The large N Limit of superconformal field theories and supergravity*, *Adv. Theor. Math. Phys.* **2** (1998) 231–252, [hep-th/9711200].

[20] E. Witten, *Anti De Sitter Space And Holography*, *Adv. Theor. Math. Phys.* **2** (1998) 253–290, [hep-th/9802150].

[21] G. W. Horndeski, *Second-order scalar-tensor field equations in a four-dimensional space*, *Int. J. Theor. Phys.* **10** (1974) 363–384.

[22] P. Jordan, *Schwerkraft und Weltall.* Vieweg, 1955.

[23] C. Brans and R. H. Dicke, *Mach's Principle and a Relativistic Theory of Gravitation*, *Phys. Rev.* **124** (1961) 925–935.

[24] P. A. M. Dirac, *A New Basis for Cosmology*, *Proc. R. Soc. A Math. Phys. Eng. Sci.* **165** (1938) 199–208.

[25] B. Ratra and P. J. E. Peebles, *Cosmological consequences of a rolling homogeneous scalar field*, *Phys. Rev. D* **37** (1988) 3406–3427.

[26] R. R. Caldwell, R. Dave and P. J. Steinhardt, *Cosmological imprint of an energy component with general equation of state*, *Phys. Rev. Lett.* **80** (1998) 1582–1585, [astro-ph/9708069].

[27] C. Deffayet, O. Pujolàs, I. Sawicki and A. Vikman, *Imperfect dark energy from kinetic gravity braiding*, *J. Cosmol. Astropart. Phys.* **2010** (2010) 026–026, [1008.0048].

[28] A. H. Guth, *Inflationary universe: A possible solution to the horizon and flatness problems*, *Phys. Rev. D* **23** (1981) 347–356.

[29] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw et al., *Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: cosmological interpretation*, *Astrophys. J. Suppl. Ser.* **192** (2011) 18, [1001.4538].

[30] The PLANCK Collaboration, *Planck 2013 results. XVI. Cosmological parameters*, *Astron. Astrophys.* **571** (2014) A16, [1303.5076].

[31] T. Kobayashi, M. Yamaguchi and J. Yokoyama, *Generalized G-Inflation: –Inflation with the Most General Second-Order Field Equations–*, *Prog. Theor. Phys.* **126** (2011) 511–529.

[32] T. Qiu, J. Evslin, Y. F. Cai, M. Li and X. Zhang, *Bouncing Galileon cosmologies*, *J. Cosmol. Astropart. Phys.* **2011** (2011) 036–036, [1108.0593].

[33] D. A. Easson, I. Sawicki and A. Vikman, *G-bounce*, *J. Cosmol. Astropart. Phys.* **2011** (2011) 021–021.

[34] The LIGO Collaboration and the VIRGO Collaboration, *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral*, *Phys. Rev. Lett.* **119** (2017) 161101.

[35] A. Goldstein, P. Veres, E. Burns, M. S. Briggs, R. Hamburg, D. Kocevski et al., *An Ordinary Short Gamma-Ray Burst with Extraordinary Implications: Fermi-GBM Detection of GRB 170817A*, *Astrophys. J.* **848** (2017) L14.

[36] The LIGO Collaboration and the VIRGO Collaboration, *Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A*, *Astrophys. J.* **848** (2017) L13.

[37] P. Creminelli and F. Vernizzi, *Dark Energy after GW170817 and GRB170817A*, *Phys. Rev. Lett.* **119** (2017) 251302.

[38] J. Sakstein and B. Jain, *Implications of the Neutron Star Merger GW170817 for Cosmological Scalar-Tensor Theories*, *Phys. Rev. Lett.* **119** (2017) 251303.

[39] J. M. Ezquiaga and M. Zumalacárregui, *Dark Energy After GW170817: Dead Ends and the Road Ahead*, *Phys. Rev. Lett.* **119** (dec, 2017) 251304.

[40] T. Baker, E. Bellini, P. G. Ferreira, M. Lagos, J. Noller and I. Sawicki, *Strong Constraints on Cosmological Gravity from GW170817 and GRB 170817A*, *Phys. Rev. Lett.* **119** (2017) 251301.

[41] X. O. Camanho, J. D. Edelstein, J. Maldacena and A. Zhiboedov, *Causality Constraints on Corrections to the Graviton Three-Point Coupling*, 1407.5597.

[42] Y. Choquet-Bruhat, *Théorème d'existence pour certains systèmes d'équations aux dérivées partielles non linéaires*, *Acta Math.* (1952) .

[43] G. Papallo and H. S. Reall, *Graviton time delay and a speed limit for small black holes in Einstein-Gauss-Bonnet theory*, *J. High Energy Phys.* **2015** (2015) 109, [1508.05303].

[44] G. Papallo and H. S. Reall, *On the local well-posedness of Lovelock and Horndeski theories*, *Phys. Rev. D* **96** (2017) 044019, [1705.04370].

[45] G. Papallo, *On the hyperbolicity of the most general Horndeski theory*, *Phys. Rev. D* **96** (2017) 124036, [1710.10155].

[46] O. Sarbach and M. Tiglio, *Continuum and Discrete Initial-Boundary Value Problems and Einstein's Field Equations*, *Living Rev. Relativ.* **15** (2012) 191, [1203.6443].

[47] G. Nagy, O. E. Ortiz and O. A. Reula, *Strongly hyperbolic second order Einstein's evolution equations*, *Phys. Rev. D* **70** (2004) 044012, [gr-qc/0402123].

[48] D. Christodoulou, *Mathematical problems of general relativity I.* European Mathematical Society, 2008.

[49] Y. Choquet-Bruhat, *General relativity and the Einstein equations.* Oxford University Press, 2008.

[50] R. Courant and D. Hilbert, *Methods of Mathematical Physics vol. 2.* Interscience, 1962.

[51] L. C. Evans, *Partial Differential Equations.* American Mathematical Society, 1998.

[52] J. Hadamard, *Sur les problèmes aux dérivées partielles et leur signification physique*, *Princet. Univ. Bull.* **13** (1902) 49–52.

[53] H. O. Kreiss and J. Lorenz, *Initial-boundary value problems and the Navier-Stokes equations*, vol. 136. Society for Industrial and Applied Mathematics, 1989.

[54] F. John, *Blow-up of solutions of nonlinear wave equations in three space dimensions*, *Manuscripta Math.* **28** (1979) 235–268.

[55] D. Christodoulou and S. Klainerman, *The Global Nonlinear Stability of the Minkowski Space.* Princeton University Press, 1993.

[56] A. D. Rendall, *Reduction of the Characteristic Initial Value Problem to the Cauchy Problem and its Applications to the Einstein Equations*, *Proc. R. Soc. A Math. Phys. Eng. Sci.* **427** (1990) 221–239.

[57] J. Luk, *On the Local Existence for the Characteristic Initial Value Problem in General Relativity*, *Int. Math. Res. Not.* (2011) 1–43, [1107.0898].

[58] Y. Choquet-Bruhat and R. Geroch, *Global aspects of the Cauchy problem in general relativity*, *Commun. Math. Phys.* (1969) .

[59] J. Sbierski, *On the Existence of a Maximal Cauchy Development for the Einstein Equations: a Dezornification*, *Ann. Henri Poincaré* **17** (2015) 301–329, [1309.7591].

[60] R. Arnowitt, S. Deser and C. W. Misner, *Dynamical Structure and Definition of Energy in General Relativity*, *Phys. Rev.* **116** (1959) 1322–1330.

[61] I. Rodnianski and J. Speck, *Stable Big Bang Formation in Near-FLRW Solutions to the Einstein-Scalar Field and Einstein-Stiff Fluid Systems*, 1407.6298.

[62] F. Pretorius, *Evolution of Binary Black-Hole Spacetimes*, *Phys. Rev. Lett.* **95** (2005) 121101.

[63] M. Campanelli, C. O. Lousto, P. Marronetti and Y. Zlochower, *Accurate Evolutions of Orbiting Black-Hole Binaries without Excision*, *Phys. Rev. Lett.* **96** (2006) 111101.

[64] J. G. Baker, J. Centrella, D. I. Choi, M. Koppitz and J. van Meter, *Gravitational-Wave Extraction from an Inspiraling Configuration of Merging Black Holes*, *Phys. Rev. Lett.* **96** (2006) 111102.

[65] T. Nakamura, K. Oohara and Y. Kojima, *General Relativistic Collapse to Black Holes and Gravitational Waves from Black Holes*, *Prog. Theor. Phys. Suppl.* **90** (1987) 1–218.

[66] M. Shibata and T. Nakamura, *Evolution of three-dimensional gravitational waves: Harmonic slicing case*, *Phys. Rev. D* **52** (1995) 5428–5444.

[67] T. W. Baumgarte and S. L. Shapiro, *Numerical integration of Einstein's field equations*, *Phys. Rev. D* **59** (1998) 024007.

[68] C. Gundlach and J. M. Martín-García, *Well-posedness of formulations of the Einstein equations with dynamical lapse and shift conditions*, *Phys. Rev. D* **74** (2006) 024016.

[69] H. Friedrich, *Hyperbolic reductions for Einstein's equations*, *Class. Quantum Gravity* **13** (1996) 1451–1469.

[70] H. S. Reall, N. Tanahashi and B. Way, *Causality and hyperbolicity of Lovelock theories*, *Class. Quantum Gravity* **31** (2014) 205005, [1406.3379].

[71] M. V. Ostrogradsky, *Mémoires sur les équations différentielles, relatives au problème des isopérimètres*, *Mem. Acad. St. Petersbg.* **6** (1850) 385–517.

[72] R. P. Woodard, *Avoiding Dark Energy with 1/R Modifications of Gravity*, astro-ph/0601672.

[73] H. Vermeil, *Notiz über das mittlere Krümmungsmaß einer n-fach ausgedehnten Riemann'schen Mannigfaltigkeit*, *Nachrichten von der Gesellschaft der Wissenschaften* (1917) .

[74] H. Weyl, *Über die physikalischen Grundlagen der erweiterten Relativitätstheorie*, *Phys. Zeitschrift* (1921) .

[75] S. S. Chern, *On the Curvatura Integra in a Riemannian Manifold*, *Ann. Math.* **46** (1945) 674.

[76] C. Aragone, *Stringy characteristics of effective gravity*, *Relativ. Gravitation, SILARG VI* (1987) 60–69.

[77] Y. Choquet-Bruhat, *The Cauchy problem for stringy gravity*, *J. Math. Phys.* **29** (1988) 1891.

[78] I. Shapiro, *Fourth test of general relativity*, *Phys. Rev. Lett.* (1964) .

[79] P. C. Aichelburg and R. U. Sexl, *On the gravitational field of a massless particle*, *Gen. Relativ. Gravit.* **2** (1971) 303–312.

[80] T. Dray and G. 't Hooft, *The gravitational shock wave of a massless particle*, *Nucl. Phys. B* **253** (1985) 173–188.

[81] S. Gao and R. M. Wald, *Theorems on gravitational time delay and related issues*, *Class. Quantum Gravity* **17** (2000) 4999–5008, [gr-qc/0007021].

[82] B. C. Palmer and D. Marolf, *Fast travel through spherically symmetric spacetimes*, *Phys. Rev. D* **67** (2003) 044012, [gr-qc/0211045].

[83] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis and R. Rattazzi, *Causality, analyticity and an IR obstruction to UV completion*, *J. High Energy Phys.* **2006** (2006) 014, [hep-th/0602178].

[84] R. Geroch, *Faster Than Light?*, 1005.1614.

[85] J. P. Bruneton, *Causality and superluminal behavior in classical field theories: Applications to k -essence theories and modified-Newtonian-dynamics-like theories of gravity*, *Phys. Rev. D* **75** (2007) 085013, [gr-qc/0607055].

[86] D. G. Boulware and S. Deser, *String-Generated Gravity Models*, *Phys. Rev. Lett.* **55** (1985) 2656–2660.

[87] K. Izumi, *Causal structures in Gauss-Bonnet gravity*, *Phys. Rev. D* **90** (2014) 044037, [1406.0677].

[88] G. Dotti and R. J. Gleiser, *Gravitational instability of Einstein-Gauss-Bonnet black holes under tensor mode perturbations*, *Class. Quantum Gravity* **22** (2005) L1–L6, [gr-qc/0409005].

[89] R. J. Gleiser and G. Dotti, *Linear stability of Einstein-Gauss-Bonnet static spacetimes: Vector and scalar perturbations*, *Phys. Rev. D* **72** (2005) 124002, [gr-qc/0510069].

[90] D. Christodoulou, *The problem of a self-gravitating scalar field*, *Commun. Math. Phys.* **105** (1986) 337–361.

[91] S. Klainerman, *The null condition and global existence to nonlinear wave equations*, *Lect. Appl. Math* **23** (1986) 293–326.

[92] S. Miao and P. Yu, *On the formation of shocks for quasilinear wave equations*, 1412.3058.

[93] J. Luk and I. Rodnianski, *Local Propagation of Impulsive GravitationalWaves*, *Commun. Pure Appl. Math.* **68** (2015) 511–624, [1301.1072].

[94] S. Willison, *Local well-posedness in Lovelock gravity*, *Class. Quantum Gravity* **32** (2015) 022001, [1409.6656].

[95] J. Lee and R. M. Wald, *Local symmetries and constraints*, *J. Math. Phys.* **31** (1990) 725–743.

[96] T. Kato, *Perturbation Theory for Linear Operators*. Springer Berlin Heidelberg, 1976.

[97] I. C. Gohberg, P. Lancaster and L. Rodman, *Matrices and Indefinite Scalar Products*. Birkhäuser, 1983.

[98] J. M. Martín-García, *xAct: efficient tensor computer algebra for the wolfram language*, 2015.

[99] J. T. Wheeler, *Symmetric solutions to the maximally Gauss-Bonnet extended Einstein equations*, *Nucl. Phys. B* **273** (1986) 732–748.

[100] N. Deruelle and L. Fariña-Busto, *Lovelock gravitational field equations in cosmology*, *Phys. Rev. D* **41** (1990) 3696–3708.

[101] A. De Felice and S. Tsujikawa, *Cosmology of a Covariant Galileon Field*, *Phys. Rev. Lett.* **105** (2010) 111301.

[102] A. De Felice and S. Tsujikawa, *Generalized Galileon cosmology*, *Phys. Rev. D* **84** (2011) 124029, [1008.4236].

[103] X. Gao, *Conserved cosmological perturbation in Galileon models*, *J. Cosmol. Astropart. Phys.* **2011** (2011) 27, [1106.0292].

[104] A. D. Rendall, *Dynamics of k-essence*, *Class. Quantum Gravity* **23** (2005) 1557–1569, [gr-qc/0511158].

[105] P. Kanti, N. E. Mavromatos, J. Rizos, K. Tamvakis and E. Winstanley, *Dilatonic Black Holes in Higher Curvature String Gravity*, *Phys. Rev. D* **54** (1995) 5049–5058, [hep-th/9511071].

[106] D. Bettoni and S. Liberati, *Disformal invariance of second order scalar-tensor theories: Framing the Horndeski action*, *Phys. Rev. D* **88** (2013) 084020, [1306.6724].

[107] J. Keir (Unpublished).