Dynamics of Early & Late Universe Cosmology

by

William Thomas Emond

Thesis submitted to
The University of Nottingham
for the degree of
Doctor of Philosophy

August 2018

Abstract

In this thesis we discuss two key problems: the cosmological constant problem (CCP), an issue that primarily manifests itself in late universe cosmology; and the process of thermalisation during the post-inflationary reheating phase of the early universe.

We start by giving a brief review of general relativity, discussing both its successes and failures, in particular, why one might consider modifications of it. We then delve into the aspects of early and late universe cosmology that we aim to address in the research discussed in this thesis. Starting with an overview of the inflationary paradigm, and the need to reheat the universe post-inflation, we give a review of previous research that has been conducted in this area. We then move on to discuss the CCP in detail, in particular, why it is such an issue. After setting the scene for this problem, we proceed to discuss how to approach finding a resolution to it, highlighting certain stumbling blocks that one needs to be mindful of.

Having set the scene, we then present a potential solution to the CCP, involving a scalar-tensor modified theory of gravity, so-called Horndeski theory. Building upon a class of Horndeski theories providing self-tuning solutions to the CCP, we provide a generalisation in which matter interacts with gravity via a disformal coupling to the spacetime metric. We establish the form of the disformally self-tuning Lagrangian on a cosmological Friedmann-Robertson-Walker background, and show that there exist non-trivial self-tuning solutions.

In the latter half of this thesis, we move on to review the literature on the non-perturbative description of the early stages of reheating, so-called preheating. With the motivation to study the less well understood thermalisation process that must necessarily take place in this phase, we then present a toy model preheating theory, in which we account for the effects of thermalisation from its onset. Within the density matrix formalism, we derive a (self-consistent) set of quantum Boltzmann equations, which are able to describe the evolution of an ensemble of self-interacting scalar particles that are subject to an oscillating mass term. In particular, we apply this to the preheating scenario in order to study the evolution of scalar particle number densities throughout this process. We then conclude by discussing our numerical analysis of the Boltzmann equations, drawing attention to some important results and features that manifest using this approach, in particular, how the process differs from the standard analysis through the inclusion of thermalisation.

List of papers

This thesis contains material from the following two papers:

'*Disformally self-tuning gravity*'
W. T. Emond & P. M. Saffin
JHEP 1603 (2016) 161
arXiv:1511.02055

'*Boltzmann equations for preheating*'
William T. Emond, Peter Millington & Paul M. Saffin
JCAP 09 (2018) 041
arXiv:1807.11726

The original material contained in these two papers is discussed in detail in §4 and §6, respectively.

Acknowledgements

There are a number of people that I would like to thank, without whom this thesis would not have been possible. It seems like a lifetime ago that I decided to stop my studies in interior design and pursue a degree in physics. This would not have been possible without the continued support and encouragement from my Mum, Dad, brother, sister, and Grandma and Grandad. To you all, I am truly grateful.

I would particularly like to thank my supervisor Paul. Your enthusiasm for physics, guidance, and support have made the last four years very enjoyable. I have certainly learned a lot from our discussions and collaborations.

A big thank you also to Peter Millington for his support and collaboration with me in the last project of my PhD, and for putting up with my many questions. My knowledge of physics has definitely benefited. Thank you also to the many people whose willingness to discuss physics with me have helped me greatly throughout my PhD career, including (but not limited to): Clare Burrage, Ed Copeland, Silvia Nagy, Florian Niedermann, Tony Padilla and David Stefanyszyn.

Life as a PhD student can be hard at times, but having supportive friends can certainly make it a bit easier more enjoyable. I am grateful to all of my friends for being there for me. In particular, I would like to thank Alan, Anna, Christian, Connor, Ellie, Jim, Johannes, Tom, and Toby. I would also like to thank my office friends: Ben, Eugenio, Felipe, Mehdi and Ben, for keeping me sane during long hours at the office.

Finally, I would like to thank the most important person in my life, Alex. Your love and support of me throughout this endeavour has made the whole thing possible. For that, I am most grateful. Thank you.

Conventions & notation

The following conventions are adopted throughout this thesis (unless stated otherwise):

- We work in units where $\hbar = c = 1$. In this units system, any quantity has dimensions of energy and as such can be expressed in terms of electronvolts, where $1\text{eV} \approx 1.602 \times 10^{-19}\text{J}$.
- The metric signature used is $(-+++)$.
- We use the reduced Planck mass $M_{\text{Pl}} = \frac{1}{\sqrt{8\pi G}} \approx 2.435 \times 10^{18}$ GeV, where G is the gravitational constant.
- Spacetime coordinates are given in 4-vector form, i.e. $x^\mu = (t, \mathbf{x})$. In particular, functions f of spacetime coordinates will be expressed as $f(x) \coloneqq f(t, \mathbf{x})$.
- We adopt the standard practice of refering to tensors $\mathbf{A}$ via their components $A^{\mu_1 \dots \mu_n}{}_{\nu_1 \dots \nu m}$.

The following is a list of notation used throughout this thesis:

- Operators are denoted with a circumflex, i.e. $\hat{\mathcal{O}}$.
- A dot over a function (or an operator) denotes a derivative with respect to coordinate time t, for example $\dot{f} = \frac{\mathrm{d}f}{\mathrm{d}t}$.
- Partial derivatives of (a set of) functions $f_j(q_1, \dots, q_n)$ are either denoted as $\partial_i f_j \coloneqq \frac{\partial f_j}{\partial q^i}$, or f_{j,q^i}. The same notation is used for derivatives of operators $\hat{\mathcal{O}}$ (with the appropriate replacement $f_j \to \hat{\mathcal{O}}$).
- For spacetime derivatives in Minkowski spacetime we adopt the shorthand notation $\partial_\mu \coloneqq \frac{\partial}{\partial x^\mu}$. In more general settings, i.e. curved spacetimes, we denote the covariant derivative as ∇_μ, which for an arbitrary tensor $A^{\mu_1 \dots \mu_n}{}_{\nu_1 \dots \nu_m}$, is given by

$$\begin{aligned}
\nabla_\sigma A^{\mu_1 \dots \mu_n}{}_{\nu_1 \dots \nu_m} = \partial_\sigma A^{\mu_1 \dots \mu_n}{}_{\nu_1 \dots \nu_m} &+ \Gamma^{\mu_1}{}_{\alpha\sigma} A^{\alpha \dots \mu_n}{}_{\nu_1 \dots \nu_m} + \dots + \Gamma^{\mu_n}{}_{\alpha\sigma} A^{\mu_1 \dots \alpha}{}_{\nu_1 \dots \nu_m} \\
&- \Gamma^{\alpha}{}_{\nu_1\sigma} A^{\mu_1 \dots \mu_n}{}_{\alpha \dots \nu_m} - \dots - \Gamma^{\alpha}{}_{\nu_m\sigma} A^{\mu_1 \dots \mu_n}{}_{\nu_1 \dots \alpha} .
\end{aligned}$$

- Higher-order spacetime derivatives are determined through repeated operation of the covariant derivative. A particularly important example is the d'Alembert operator $\Box := \nabla^\mu \nabla_\mu$. In this thesis, we also use other combinations: $(\nabla_\mu \nabla_\nu f)^2 = \nabla^\mu \nabla^\nu f \nabla_\nu \nabla_\mu f$ and $(\nabla_\mu \nabla_\nu f)^3 = \nabla^\mu \nabla_\nu f \nabla^\nu \nabla_\lambda f \nabla^\lambda \nabla_\mu f$ where f is a proxy spacetime-dependent function.

- In §6 we shall also adopt the following compact notation for spatial integrals:

$$\int_{\mathbf{x}_1,\, \dots,\, \mathbf{x}_n} := \int_{-\infty}^{+\infty} \mathrm{d}x^1 \, \cdots \, \int_{-\infty}^{+\infty} \mathrm{d}x^n$$

and similarly for momentum integrals:

$$\int_{\mathbf{p}_1,\, \dots,\, \mathbf{p}_n} := \int_{-\infty}^{+\infty} \frac{\mathrm{d}^3\mathbf{p}_1}{(2\pi)^3\sqrt{2\omega_{\mathbf{p}_1}(t)}} \, \cdots \, \int_{-\infty}^{+\infty} \frac{\mathrm{d}^3\mathbf{p}_n}{(2\pi)^3\sqrt{2\omega_{\mathbf{p}_n}(t)}}$$

where $\omega_{\mathbf{p}_i}(t) = \sqrt{|\mathbf{p}_i|^2 + m_{\text{eff}}^2(t)}$ (in which the time-dependent effective mass $m_{\text{eff}}^2(t)$ is defined as in §6.1.1).

- The various curvature terms (tensors and scalars), prevalent in the settings of general relativity and the model(s) of modified gravity that we consider in this thesis, are

$$\begin{aligned}
R^\alpha{}_{\mu\beta\nu} &= \partial_\beta \Gamma^\alpha_{\nu\mu} - \partial_\nu \Gamma^\alpha_{\beta\mu} + \Gamma^\alpha_{\beta\lambda}\Gamma^\lambda_{\nu\mu} - \Gamma^\alpha_{\nu\lambda}\Gamma^\lambda_{\beta\mu}\,, \\
R_{\mu\nu} &= R^\alpha{}_{\mu\alpha\nu} = g^{\alpha\beta} R_{\alpha\mu\beta\nu}\,, \\
R &= g^{\mu\nu} R_{\mu\nu}\,, \\
\hat{\mathcal{G}} &= R^2 - 4R^{\mu\nu}R_{\mu\nu} + R^{\mu\nu\rho\lambda}R_{\mu\nu\rho\lambda}\,, \\
P^{\mu\nu\alpha\beta} &= \frac{1}{4}\varepsilon^{\mu\nu\lambda\sigma} R_{\lambda\sigma\gamma\delta}\varepsilon^{\gamma\delta\alpha\beta}\,.
\end{aligned}$$

where $R^\alpha{}_{\mu\beta\nu}$ and $R_{\mu\nu}$ are the Riemann and Ricci curvature tensors, respectively, $R = g^{\mu\nu}R_{\mu\nu}$ is the scalar curvature, $\hat{\mathcal{G}}$ is the Gauss-Bonnet combination, and $P^{\mu\nu\alpha\beta}$ is the double-dual of the Riemann tensor. The $\Gamma^\alpha_{\nu\mu}$ are the (coefficients of the) affine connection, $\Gamma^\alpha_{\nu\mu} = \frac{1}{2}g^{\alpha\beta}\big(\partial_\nu g_{\beta\mu} + \partial_\mu g_{\nu\beta} - \partial_\beta g_{\nu\mu}\big)$, $\varepsilon^{\mu\nu\lambda\sigma}$ is the Levi-Civita tensor (density), and $g_{\mu\nu}$ the metric tensor.

Abbreviations

Here we list (in alphabetical order) the abbreviations of commonly used words throughout this thesis:

BBN	Big-bang nucleosynthesis.
CCP	Cosmological constant problem.
CMB	Cosmic microwave background.
CM	Centre-of-mass.
EEP	Einstein equivalence principle.
EFE	Einstein field equation.
EFT	Effective field theory.
EH	Einstein-Hilbert.
EOM	Equation(s) of motion.
FRW	Friedmann-Robertson-Walker.
GR	General relativity.
HBB	Hot big-bang.
IR	Infrared.
QED	Quantum electrodynamics.
QFT	Quantum field theory.
PDE	Partial differential equation.
SM	Standard model (of particle physics).
UV	Ultraviolet.
VEV	Vacuum expectation value.

Contents

Chapter 1

Introduction

1.1 Overview

In our quest to understand the universe, and how it came into existence, cosmology has proven to be a powerful tool, enabling us to develop descriptions of its very early moments, to the formation of structure, and importantly, track its evolution to the present day (and beyond). At the heart of cosmology is gravity. Indeed, despite it being significantly weaker than the other three fundamental forces: electromagnetism, weak and strong nuclear forces[1], on the scales that are dealt with in cosmology, it is often the most relevant interaction. That being said, a thorough understanding of non-gravitational physics is vital in order to describe the complex interactions that occur, particularly in the early universe, giving rise to the rich structure and abundances of elements that we observe in the present day universe. Thus, if we are to attain a complete description of the universe, a thorough and consistent understanding of both gravity and the three non-gravitational forces is needed to fully understand its history, how it came into being and how it has evolved into the complex cosmos that we observe today. As such, in our efforts to attain these goals, it is of great importance to have robust theories describing the gravitational and non-gravitational interactions, and furthermore, a complementary cosmological model to specify the evolution and composition of the universe.

Since the non-gravitational physics of our universe is inherently quantum, it seems reasonable to expect that the gravitational sector might also be quantum in nature. This being the case, one ultimately seeks to determine a complete quantum theory of both gravity, and the non-gravitational forces. At present, the standard model of particle physics provides us with an empirically well tested quantum theory for the non-gravitational forces, however, a complete quantum description of gravity is yet to be attained. Nonetheless, in many situations, it is possible to combine our classi-

[1] There is a hierarchy of $\sim \mathcal{O}(10^{42})$ between electromagnetism and gravity. Similarly, there are hierarchies of $\sim \mathcal{O}(10^{24})$ and $\sim \mathcal{O}(10^{38})$ between the weak and strong nuclear forces and gravity, respectively.

cal understanding of gravity with our quantum description of electromagnetism, and the weak and strong nuclear forces, to successfully further our knowledge of how the universe "ticks".

In this chapter, we briefly discuss the current best (classical) description of gravity, general relativity, and how it can be interpreted in the context of effective field theory as a low energy quantum description of a massless spin-2 particle. We shall also touch upon its theoretical and empirical successes and failures. We then move on, in §2 and §3, to summarise some important problems existing in cosmology, our current understanding of both the early and late universe, and how these are related to the interplay between contemporary descriptions of gravity and non-gravitational physics. §4 is then dedicated to a detailed discussion, and derivation of a solution to the CCP (introduced in §3). In §5 we move on to address the current state of affairs in the theory of post-inflationary reheating in the universe, the non-perturbative aspects of it, and in particular, the less well understood details of how the particles produced in this phase thermalise. §6 is then focused on a detailed derivation and analysis of a toy model of reheating that aims to capture the effects of thermalisation from its onset, serving to highlight the need to include these contributions throughout the full reheating process if one is to correctly estimate the reheating temperature of the universe. Finally, in §7, we shall make some concluding remarks on the topics discussed throughout this thesis, and in particular, the outcomes and future research directions of the projects that we have reviewed.

1.2 GR: A robust theory of gravity

At present, the prevailing theory of gravity is general relativity; first formulated by Einstein over 100 years ago to reconcile inconsistencies between Newtonian gravity and special relativity [1].[2] Einstein posited that gravitation is nothing but the manifestation of the local curvature of spacetime sourced by the presence of matter, a statement that was based on the founding principle of GR: the *principle of equivalence*. In fact, to be precise, it is the *Einstein equivalence principle*:

> **EEP**: The outcome of any local non-gravitational experiment in a freely falling laboratory is independent of the velocity of the laboratory and its position in spacetime.

[2] In the following section we make use of Refs. [2] and [3], and we refer the reader to them for further details.

That is, all forms of matter and energy (by which we mean all massive and massless particles, and all forms of non-gravitational energy) will undergo the same acceleration, independently of their properties, when subject only to gravity. In particular, the EEP implies that locally[3] it is impossible, even in principle, by means of observing any non-gravitational physics, to distinguish between the effects of an arbitrary gravitational field and those of a uniformly accelerating reference frame in Minkowksi spacetime. Note, however, that the effects of a true gravitational field can never be completely transformed away by a suitable choice of coordinates. Indeed, the ability to negate the effects of gravity locally relies on the fact that, within infinitesimal regions of spacetime, the gravitational field is (approximately) homogeneous. As we extend such spacetime regions to finite sizes, there will generically be inhomogeneities in the gravitational field. These lead to tidal forces which can be detected.

The fact that the local effects of a gravitational field are independent of the nature of any given particle suggests that the effect of gravity is not related to the properties of matter, but is an intrinsic feature of spacetime itself. Indeed, such observations suggest that one should view gravity as the manifestation of spacetime curvature, the geometry of which should be described by a metric tensor $g_{\mu\nu}(x)$. Furthermore, the local curvature of spacetime should be attributed to the presence of matter (i.e. all forms of mass and energy source local spacetime curvature). Consequently, in the absence of external forces, the trajectories of all species of particles lie along the geodesics of $g_{\mu\nu}(x)$. The tidal forces that manifest over finite regions within a gravitational field then result in relative accelerations between neighbouring geodesics (so-called geodesic deviation).

Moreover, the EEP implies that within a sufficiently small neighbourhood of each spacetime point, the (non-gravitational) laws of physics reduce to those of special relativity. With this is mind one can uniquely recast their corresponding mathematical descriptions into tensorial form[4] such that they are generally covariant, i.e. they are valid in *all* frames of reference in the presence of an arbitrary gravitational field. The EEP is often further extended to the so-called *strong equivalence principle* to include local gravitational experiments, taking into account bodies whose gravitational self-energy contribute significantly to their overall mass. Thus, one posits that

[3]By *local* we mean that we are in a small enough volume of spacetime such that the tidal effects of gravitation can be neglected.

[4]The is the so-called principle of general covariance - the requirement that the form of physical laws should be independent of any reference frame (i.e. they should be coordinate independent).

it is impossible to locally distinguish between gravitation and uniform acceleration through observing any kind of local experiment, both non-gravitational and gravitational.

The supposition that gravity is a geometric phenomenon suggests that a theory describing it should be formulated within the framework of differential geometry. Indeed, if spacetime is represented by a pseudo-Riemannian manifold $\mathcal{M}$ equipped with a metric $g_{\mu\nu}$, then within infinitesimally small regions its geometry is that of Minkowski spacetime, however for finite regions its curvature is apparent. This conforms with the EEP: that in the presence of a gravitational field, locally (i.e. infinitesimal regions) the geometry is Minkowksi, however, over finite regions tidal forces arise as a consequence of the non-vanishing gravitational field. GR formalises the notion that gravity is the manifestation of spacetime geometry through the construction of a gravitational action from curvature invariants. This is the Einstein-Hilbert action, which is given by

$$S_{\text{EH}}[g_{\mu\nu}] \;=\; \int_{\mathcal{M}} \mathrm{d}^4x\sqrt{-g}\left[\frac{M_{\text{Pl}}^2}{2}R - \Lambda_{\text{bare}}\right], \tag{1.2.1}$$

where $g = \det(g_{\mu\nu})$ is the determinant of the metric tensor, R is the Ricci scalar, and Λ_{bare} is a bare cosmological constant. Note that the theory has two inherent dimensionful scales built into it by the presence of the (reduced) Planck mass M_{Pl} and a cosmological constant term. We include matter in the theory by a universal minimal coupling (this universality is dictated by the EEP) to the metric tensor[5], such that the total action is a linear combination of gravitational action and the action describing the field theory sector,

$$S[g_{\mu\nu},\Psi] \;=\; S_{\text{EH}}[g_{\mu\nu}] + S_{\text{M}}[g_{\mu\nu},\Psi]\,, \tag{1.2.2}$$

where Ψ generically denotes the matter fields. The equations of motion for this theory are found by varying the action with respect to $g_{\mu\nu}$ (where $g_{\mu\nu}$ satisfies Dirichlet boundary conditions, i.e. it is fixed on the spacetime boundary[6]), and due to the

[5]By *minimal coupling* we mean that one introduces a covariant integration measure $\mathrm{d}^4x \to \mathrm{d}^4x\sqrt{-g}$ to the matter action and promotes partial derivatives to covariant derivatives $\partial_\mu \to \nabla_\mu$.

[6]Technically one should include a Gibbons-Hawking term [4] in eq. (1.2.2) to account for spacetime manifolds with boundaries (see e.g. ref. [5] for a detailed review). Here we shall implicitly assume its presence in the derivation of the EOM.

symmetry of the metric tensor, correspond to 10 partial differential equations

$$\frac{\delta S}{\delta g^{\mu\nu}} = M_{\rm Pl}^2 G_{\mu\nu} + g_{\mu\nu}\Lambda_{\rm bare} - T_{\mu\nu} = 0$$

$$\Rightarrow \quad G_{\mu\nu} = \frac{1}{M_{\rm Pl}^2}\left[T_{\mu\nu} - g_{\mu\nu}\Lambda_{\rm bare}\right], \tag{1.2.3}$$

where $G_{\mu\nu}$ is the Einstein tensor, and $T_{\mu\nu}$ is the energy-momentum tensor of the field theory sector and is defined as

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S_{\rm M}}{\delta g^{\mu\nu}}. \tag{1.2.4}$$

These are generally referred to as the Einstein field equations. We see then that the metric is a dynamical quantity, i.e. it is a solution to the EFEs [eq. (1.2.3)] whose form is dictated by the local matter configuration (described by $T_{\mu\nu}$). We see then, that the geometry and the behaviour of the matter fields are dynamically determined - the theory is free of prior geometry, i.e *background independent* (this is true up to the presence of the predetermined cosmological constant). Consequently, it is found that given the same initial data, one can have more than one solution to the EFEs, eq. (1.2.3) (analogous to the gauge redundancy of quantum electrodynamics in which two potentials A^μ and A'^μ, related by a gauge transformation $A^\mu \to A^\mu + \partial_\mu \alpha(x)$, can be obtained from Maxwell's equations given the same initial data). Indeed, this is the case if two given metrics $\tilde{g}_{\mu\nu}$ and $g_{\mu\nu}$ are related by a diffeomorphism $\phi : \mathcal{M} \to \mathcal{M}$, i.e. $\tilde{g}_{\mu\nu} = (\phi^* g)_{\mu\nu}$. For the theory to be deterministic it must be that these two solutions are physically equivalent.[7] Thus, a given gravitational field cannot be represented by a single mathematical field, but instead an equivalence class of diffeomorphic fields. This necessitates that the full action [eq. (1.2.2)] is *diffeomorphism invariant.*[8] In fact, when considered in isolation, the EH action [eq. (1.2.1)] is diffeomorphism invariant, and this implies that $G_{\mu\nu}$ satisfies the (contracted) Bianchi identity $\nabla_\mu G^{\mu\nu} = 0$.[9] As a consequence, this requires that the matter action should simultaneously be diffeomorphism invariant such that the full action [eq. (1.2.2)] is. This condition enforces the on-shell covariant conservation of $T_{\mu\nu}$, i.e. $\nabla_\mu T^{\mu\nu} = 0$. The fact that GR is diffeomorphism invariant is captured

[7]Since we have no way of distinguishing between them, we would otherwise lose predictability as we would not be able to uniquely determine the gravitational field given any initial data.

[8]See e.g. ref. [6] for a detailed analysis of Einstein's so-called "hole" argument.

[9]Note that this is an off-shell statement, i.e. it is true for *any* metric, regardless of whether it is a solution to the EFEs or not.

in the field equations [eq. (1.2.3)]. Indeed, the Bianchi identity reveals that 4 of the 10 PDEs are in fact constraint equations (they cannot propagate the initial data). Moreover, the freedom to choose our spacetime coordinates introduces a fourfold degeneracy (corresponding to the gauge freedom of GR). Taking these details into account, the theory actually propagates only two physical degrees of freedom. These are the polarisations of the graviton.

At this point we observe that Λ_{bare} has been included in the EH action [eq. (1.2.1)] simply because it is not prevented by diffeomorphism invariance and hence there is no a priori reason to exclude it. However, as we shall discuss in detail in §2, there are contributions from the vacuum energies ρ_{vac} of the matter fields that manifest themselves as a cosmological constant. Accordingly, since gravity has no way of distinguishing between the two contributions, it is actually the net cosmological constant $\Lambda = \Lambda_{\text{bare}} + \rho_{\text{vac}}$ that sources curvature in Einstein's field equations [eq. (1.2.3)].

1.3 Experimental & theoretical successes of GR

Phenomenologically, GR has had many successes, indeed, some examples of experimental confirmation of its predictions include the perihelion precession of Mercury, the bending of light in a gravitational field (due to the local warping of spacetime by massive bodies) and gravitational redshifting (a consequence of the equivalence principle in the case where light is propagating out of a gravitational potential well). So far, GR has passed every experimental test, on Earth and in the solar system, highlighting its accomplishment as a theory of gravity, see for example refs. [7–10]. The pinnacle of this success has to be the experimental discovery of gravitational waves, the final prediction of GR to be experimentally verified. First observed from a binary black hole merger in 2016 [11], followed by a flurry of further detections [12–15], and most recently from a binary neutron star merger [16]. These detections have given further credence to predictive power of GR, and furthermore, the most recent detection has enabled even more detailed analyses due to data being collected from both gravitational and electromagnetic (EM) waves generated by the event. This data has further supported the validity of GR, with stringent bounds being placed on the difference between the speed of gravitational and EM waves ($-3\times10^{-15} \leq \frac{c_g - c}{c} \leq 5\times10^{-16}$, where c_g is the gravitational wave speed and c is the speed of light), as well as violations of Lorentz invariance and the equivalence principle (by constraining the Shapiro delay between gravitational and EM radiation[10]) [16].

Empirical successes aside, GR is also theoretically extremely robust. Indeed, from minimal assumptions it can be derived as the unique geometrical theory of gravity, starting from either a geometrical or a field theoretic approach. In the framework of differential geometry, it was first proven by Lovelock that if one wishes to construct a theory of gravity from an action principle, in four dimensional (pseudo-) Riemannian space, from the spacetime metric $g_{\mu\nu}(x)$ and its derivatives alone, then the unique EOM following from this action are Einstein's (vacuum) equations with the addition of a bare cosmological constant [eq. (1.2.3)] [18, 19]. We note here that while Lovelock's theorem identifies Einstein's equations as the unique second-order field equations for a metric theory of gravity in 4D (pseudo-) Riemannian space, it does not preclude the presence of higher-order curvature terms in the gravitational action, i.e. the EH action [eq. (1.2.1)] is not the unique action constructed from $g_{\mu\nu}(x)$ that leads to these equations. Indeed, in four dimensions (or less), the most general action that one can construct from $g_{\mu\nu}(x)$ is given by[11]

$$S[g_{\mu\nu}] \;=\; \int \mathrm{d}^4x\sqrt{-g}\left[\alpha R - 2\Lambda + \beta\varepsilon^{\mu\nu\rho\sigma}R^{\alpha\beta}_{\ \ \mu\nu}R_{\alpha\beta\rho\sigma} + \gamma\Big(R^2 - 4R^{\mu\nu}R_{\nu\mu} + R^{\mu\nu}_{\ \ \rho\sigma}R^{\rho\sigma}_{\ \ \mu\nu}\Big)\right], \tag{1.3.1}$$

where α, Λ, β and γ are constants, $R^{\mu}_{\ \alpha\nu\beta}$ is the Riemann curvature tensor, $R_{\mu\nu}$ the Ricci curvature tensor, and $\varepsilon^{\mu\nu\rho\sigma}$ the Levi-Civita tensor density. It can be shown the third and fourth terms in this expression do not contribute to the EOM, and so we recover Einstein's equations.[12]

In the context of field theory, one can arrive uniquely at GR as the low-energy limit of a local Lorentz invariant theory for interacting massless spin-2 particles [21–26]. In brief, one starts by considering a theory for a local massless spin-2 field operator $h_{\mu\nu}(x)$ whose quanta we shall refer to as gravitons. We then construct a Lagrangian for this theory by writing down all possible terms, quadratic in derivatives of $h_{\mu\nu}(x)$, that are compatible with Lorentz invariance and locality, of which the general form

[10]The Shapiro delay is a gravitational phenomenon, predicted in GR, in which the time taken for a photon to propagate past a massive body is longer than it would be in its absence, due to gravitational spacetime dilation increasing the path length [17].

[11]Note, however, that the $R^{\alpha\beta}_{\ \ \mu\nu}R_{\alpha\beta\rho\sigma}$ term is forbidden if one assumes parity, due to the presence of the Levi-Civita tensor $\varepsilon^{\mu\nu\rho\sigma}$.

[12]For further discussion of Lovelock's theorem c.f. this review [20]

is

$$\mathcal{L} \;=\; \alpha\partial_\lambda h^{\mu\nu}\partial^\lambda h_{\mu\nu} + \beta\partial_\mu h^{\mu\nu}\partial^\lambda h_{\nu\lambda} + \gamma\partial_\nu h\partial_\mu h^{\mu\nu} + \sigma\partial_\mu h\partial^\mu h\,, \tag{1.3.2}$$

where α, β, γ and σ are constants and $h := \eta^{\mu\nu}h_{\mu\nu}$ is the trace of $h_{\mu\nu}$. By insisting upon both Lorentz and gauge invariance of the theory, we require that eq. (1.3.2) must be invariant under Lorentz transformations $x^\mu \to \Lambda^\mu{}_\nu x^\nu$, and simultaneously invariant (up to boundary terms) under the following transformation of $h_{\mu\nu}$

$$h_{\mu\nu}(x) \to h_{\mu\nu}(x) + \partial_\mu\xi_\nu(x) + \partial_\nu\xi_\mu(x)\,, \tag{1.3.3}$$

where $\xi^\mu(x)$ is a generic vector field. In fact, eq. (1.3.3) corresponds exactly to a gauge transformation of a massless spin-2 field. We can exploit this gauge freedom to determine the constants in eq. (1.3.2); indeed, by replacing of occurrences of $h_{\mu\nu}$ in eq. (1.3.2) with the transformation given by eq. (1.3.3), and integrating by parts (neglecting boundary terms), we obtain constraints on α, β, γ and σ such that

$$\alpha \;=\; -\frac{1}{2}\beta\,, \quad \gamma \;=\; -\beta\,, \quad \sigma \;=\; \frac{1}{2}\beta\,. \tag{1.3.4}$$

Therefore, by choosing $\beta = -\frac{1}{2}$ we find that eq. (1.3.2) takes the form,

$$\mathcal{L} \;=\; \frac{1}{4}\partial_\lambda h^{\mu\nu}\partial^\lambda h_{\mu\nu} - \frac{1}{2}\partial_\mu h^{\mu\nu}\partial^\lambda h_{\nu\lambda} + \frac{1}{2}\partial_\nu h\partial_\mu h^{\mu\nu} - \frac{1}{4}\partial_\mu h\partial^\mu h\,. \tag{1.3.5}$$

One can then determine the field equations for this spin-2 field in the usual manner, from the Euler-Lagrange equations,

$$\Box h_{\mu\nu} - \eta_{\mu\nu}\Box h - 2\partial^\lambda\partial_{(\mu}h_{\nu)\lambda} + \eta_{\mu\nu}\partial_\lambda\partial_\sigma h^{\lambda\sigma} + \partial_\mu\partial_\nu h \;=\; 0\,, \tag{1.3.6}$$

where $\Box$ is the d'Alembert operator. Remarkably, these are exactly the linearised Einstein field equations in a Minkowski vacuum. To see this, consider a small perturbation of the metric $g_{\mu\nu}$ around a Minkowski background $\eta_{\mu\nu}$,

$$g_{\mu\nu} \;=\; \eta_{\mu\nu} + h_{\mu\nu}\,, \tag{1.3.7}$$

where $h_{\mu\nu}$ is a symmetric rank $(0,2)$ tensor that we require to be *small*, in the sense that, in any given coordinate system, its components are infinitesimally small, $|h_{\mu\nu}| \ll 1$. One can find the inverse metric from eq. (1.3.7) and the fact that $\delta g^{-1} =$

$g^{-1}\delta g g^{-1}$ (where $\delta g = h$). Indeed, one finds that,

$$g^{\mu\nu} \;=\; \eta^{\mu\nu} - h^{\mu\nu} + h^{\mu\lambda}h^{\nu}_{\lambda} + \mathcal{O}(h^3)\,. \tag{1.3.8}$$

Furthermore, the determinant of the metric $g := \sqrt{|g_{\mu\nu}|}$ can be expanded in powers of $h_{\mu\nu}$ to give

$$g \;=\; -\big(1 + h + \mathcal{O}(h^2)\big) \tag{1.3.9}$$

leading to

$$\sqrt{-g} \;=\; 1 + \frac{1}{2}h + \mathcal{O}(h^2)\,. \tag{1.3.10}$$

Since the perturbation is small, we can consider only the leading order contributions to the Einstein-Hilbert action, i.e. up to quadratic-order in the Lagrangian, such that we obtain the linear approximation of Einstein's field equations. Equipped with this information one can show that the quadratic-order Einstein-Hilbert action $\delta_{(2)}S_{EH}$ has the following form,

$$\delta_{(2)}S_{EH} \;=\; -\frac{M_{Pl}^2}{2}\int \mathrm{d}^4x\Big[\frac{1}{2}\partial_\lambda h^{\mu\nu}\partial^\lambda h_{\mu\nu} - \partial_\mu h^{\mu\nu}\partial^\lambda h_{\nu\lambda} + \partial_\nu h\partial_\mu h^{\mu\nu} - \frac{1}{2}\partial_\mu h\partial^\mu h\Big]\,, \tag{1.3.11}$$

we see then that the Lagrangian in eq. (1.3.11) has the exact same form as eq. (1.3.5), up to a constant of proportionality, and thus clearly leads to the same equations of motion, the linearised Einstein equations in vacuo. Furthermore, eq. (1.3.11) is invariant under linearised diffeomorphisms, $x^\mu \rightarrow x^\mu - \xi^\mu$ (where ξ^μ is the generator of the diffeomorphism) in which the perturbation to the metric $h_{\mu\nu}$ transforms as in eq. (1.3.3). In this sense we can view $h_{\mu\nu}$ either from a geometric standpoint, as an infinitesimal perturbation to the Minkowksi metric, or from a field theoretic stance, in which $h_{\mu\nu}$ is a massless spin-2 quantum field, propagating on a Minkowski background, whose quanta are identified as gravitons. Moreover, if we were to further couple the graviton field to a set of matter fields, then the only way to achieve this in a Lorentz invariant manner is if all the coupling constants are equal, i.e. Lorentz invariance implies the equivalence principle [27, 28]. Although it would certainly be an overstatement to say that "*all (theoretical) roads lead to GR*", it can be seen that GR can be derived almost uniquely from two completely different starting points, a result that, at the very least shows that GR is theoretically well founded.

1.4 Where GR fails

Despite its many successes, GR is not without its failings and ultimately, it cannot be the full UV completed theory of gravity. As we have shown, GR can be interpreted as low energy description of a QFT for a massless spin-2 field and in fact, it has been shown to be renormalisable at the one-loop level [29]. Unfortunately, this is a fluke, and in continuing this perturbative approach to quantum gravity, one finds that it becomes non-renormalisable at two-loops [30, 31]. The problem is exacerbated further when matter is introduced, in fact, GR coupled to matter is not even renormalisable at one-loop [29]. Therefore, GR is at best an effective description of the ultimate quantum theory of gravity.[13] As an effective field theory, the cut-off scale of GR is at the Planck mass, $M_{\rm Pl}$. At energies well below this scale, $E \ll M_{\rm Pl}$, GR is a well-defined QFT and is extremely predictive. However, when $E \sim M_{\rm Pl}$ the EFT breaks down since the theory becomes strongly coupled, and hence non-perturbative - higher-order quantum effects of gravity can no longer be neglected. This breakdown of perturbation theory signals the need for a full UV completion of GR. As alluded to earlier, this issue is related to the non-renormalisablity of GR, i.e. it is not possible to cancel the UV divergences (that necessarily appear as we consider higher-order loop effects) with a finite number of counter-terms, a consequence of the fact that GR has a coupling constant of negative mass dimension, $[1/M_{\rm Pl}] = -1$.

Indeed, this can readily be seen by power-counting. From eq. (1.3.5) we see that each term in the Lagrangian for the graviton contains only kinetic terms (due to it being massless), and as such, in momentum space the graviton propagator behaves as $\sim \frac{1}{k^2}$, whilst each vertex contributes a factor of k^2, furthermore, each loop integral supplies a factor of k^4. Accordingly, for V vertices, P internal lines and L loops, the superficial degree of divergence, D, of a Feynman diagram in GR is given by $D = 4L + 2V - 2P$. From the topological relation, $L = 1 - V + P$, it follows that $D = 2L+2$, where L is the number of loops in a given diagram [33, 34]. We therefore see that the degree of divergence increases as the number of loops increases, rendering GR non-renormalisable. Various attempts have been made to find a UV completion of GR, the most prominent being loop quantum gravity [35] and string theory [36, 37].[14]

[13] We refer the reader to these notes [32] for an detailed introduction to GR as an EFT.

[14] One might argue that string theory is currently the more promising of the two, given that it contains GR as a low energy limit, and is furthermore a quantum theory of not just gravity, but all the fundamental forces (see, e.g., [38] for further details). Loop quantum gravity, on the other hand, is purely a quantum theory of gravity, and is unable (at present) to recover GR in the semi-classical limit (see, e.g. [39] for a review).

Quantum issues aside, GR faces classical problems too. Astrophysical data has shown that rotation curves of galaxies are not in agreement with what one predicts from GR given the distribution of baryonic matter within them. Indeed, one expects the dominant contribution to the gravitational field to be from the baryonic matter distribution, which is predominantly at the core of the galaxy, $M(r) \sim M = \text{const.}$ Using the Newtonian limit of GR and assuming circular orbits for simplicity, this leads to the following velocity profile of the outer stars,

$$\frac{GM}{r^2} = \frac{v^2}{r} \qquad \Longrightarrow \qquad v(r) \propto \frac{1}{\sqrt{r}}\,, \tag{1.4.1}$$

however, what is actually observed is that the velocity profile flattens out as one moves further from the galatic core and is roughly constant, $v \approx$ const. at large radii [40, 41]. One can account for this flattening by assuming that the matter distribution within the galaxy grows linearly with distance, i.e. $M(r) \propto r$, however, as we have stated, the observable matter is largely found near the galatic core. As such, GR coupled to ordinary baryonic matter cannot account for this observation. To resolve this issue, it has been postulated that an exotic, non-baryonic form of matter, generically referred to as *dark matter* exists, interacting weakly with the SM fields, thus explaining why it is not directly observable in non-gravitational physics. In the present day, there is a significant amount of astrophysical data that points towards the necessity of dark matter; sophisticated analyses using x-ray emissions, gravitational lensing and mass modelling from galaxy clusters provide overwhelming evidence for dark matter [42, 43].

A further issue that arises is the fact that observational evidence indicates that the universe is currently undergoing a late-time period of acceleration. Naively, one would expect the universe to be decelerating due to the presence of baryonic and dark matter, however, the experimental data suggests otherwise. Indeed, this was first noticed from studying Type 1A supernovae which were observed to be dimmer than would be expected assuming that the universe is presently dominated by non-relativistic baryonic matter [44, 45]. Moreover, earlier observations suggested that the net density parameter Ω_{M} (a measure of the energy density of the universe) for baryonic and dark matter was $\Omega_{\mathrm{M}} \sim 0.1 - 0.3$ [46, 47]. Now, assuming a period of cosmological inflation [48], it was expected that $\Omega_{\mathrm{M}} \sim 1$, and so there was a disecrepancy of ~ 0.7. One can explain both of these observations by postulating that, alongside the energy provided by baryonic and dark matter, there exists another form of energy, generically referred to as *dark energy*.

The simplest candidate for dark energy is a cosmological constant Λ quantifying the intrinsic vacuum energy of spacetime. As it is a constant energy density associated with the vacuum it is unaffected by the cosmic expansion and is therefore able to drive a late-time accelerated expansion. Since the initial discovery of the accelerated expansion of the universe and the measurements of Ω, further compelling evidence has been provided by experimental data. For example, the WMAP and Planck experiments of the Cosmic Microwave Background have made precise measurements suggesting that baryonic matter and dark matter comprise $\sim 31\%$ (with the majority, $\sim 27\%$, coming from dark matter) of the total energy density of the universe, which infers that dark energy must account for the remaining $\sim 68\%$. They have furthermore placed bounds on the density parameter and the equation of state for dark energy, indicating that $\Omega = \Omega_{\mathrm{M}} + \Omega_{\mathrm{DE}} \approx 1$, and $w \approx -1$, implying that dark energy behaves like a cosmological constant [49–51].

Now, GR is perfectly capable of providing an explanation for this. The presence of a cosmological constant term, $\Lambda = \Lambda_{\mathrm{bare}} + \rho_{\mathrm{vac}}$, in the full action for GR [eq. (1.2.2)] fulfills the essential requirements for a dark energy candidate - it can account for the remaining energy density required by observations, and furthermore, it is a constant energy density. Indeed, the fact that Λ does not redshift, as all other forms of energy density do, means that it will eventually completely dominate the total energy density, causing the universe to evolve into a period of exponential expansion, a so-called de Sitter phase [52]. The cosmological constant then determines the magnitude of the spacetime curvature during this phase, and current observational data constrains its value to $\Lambda_{\mathrm{obs}} \sim (\mathrm{meV})^4$ [51].

The problem arises when one takes into account that the matter sector is described in the framework of quantum field theory, which implies that each particle species has an associated vacuum energy density. As a consequence of the equivalence principle, vacuum energy should couple to gravity identically to all other forms of matter and energy, i.e. it should gravitate. The issue here is that a priori the vacuum energy densities from each of the massive particle species should contribute to the cosmological constant, and in the context of the SM, these energy densities are extremely sensitive to high energy physics. As such the natural value of the cosmological constant is many orders of magnitude larger than its observed value, even if we only include the contribution from the lightest SM particle, the electron[15]. Indeed, it is found that the cosmological horizon would be less than the distance between the Earth and the Moon, $d_H \lesssim 10^6$km [53], whereas current observations indicate that $d_H \sim H_0^{-1} \sim 10^{23}$km [51] (where H_0 is the current value of the Hubble parameter).

One can naively reconcile this issue by introducing the bare cosmological constant Λ_{bare}, itself divergent, such that it can absorb the divergences present in the vacuum energy density ρ_{vac}. In doing so, what actually gravitates is then the finite combination of these two quantities, i.e. $\Lambda = \Lambda_{\text{bare}} + \rho_{\text{vac}}$. The problem is that this presents us with a serious amount of fine-tuning - there must be cancelling to an accuracy of (at least) 1 part in 10^{60} between the finite parts of Λ_{bare} and ρ_{vac} in order to match the observational bound of $\Lambda \lesssim (\text{meV})^4$. Things get much worse, however, once one considers radiative corrections to ρ_{vac}. It is found that, as one includes higher-order loop corrections, they are not significantly suppressed relative to the lower order contributions, and in fact, one is forced to *re-tune* Λ_{bare} at each level in perturbation theory to a similar degree of accuracy as the lower order contributions. As such, the vacuum energy density is *radiatively unstable*, implying that it is very sensitive to the details of UV physics of which we are ignorant. This is the crux of what is known as the *cosmological constant problem*, which we shall review in detail in §3.

We see then, that a number of phenomenological issues provide motivation to consider modifications to GR. In fact, the problem of dark matter prompted the construction of modified models such as MOND [54], and its relativistic generalisation, TeVeS [55], which attempt to explain dark matter by modifying gravity rather than introducing new forms of exotic matter. The issue of dark energy has also generated a plethora of modified theories of gravity, many of which do not attempt to account for the effects of dark energy, but to "simply" solve the CCP, for example the "fat" graviton [56] and SLED [57]. Since the CCP presents such a big problem, by itself it is very much a warranted motivation for modifying GR. One particular approach that has proven popular, is to accept that the cosmological constant is radiatively unstable, and large, but to address this by preventing it from sourcing curvature. This idea has been explored recently in the *sequester* model [58], in which a global modification is made such that the gravitational equations of GR are only modified in the infinite wavelength limit. This has the advantage of not requiring any

[15] Neglecting the neutrino masses.

screening mechanism to pass local solar system tests, since the theory reduces to GR locally (in fact, it was further shown that the theory can be constructed such that it is manifestly local, with the global modifications corresponding to conserved quantities [59]). Another recent example that adopts this de-gravitation approach is the Fab-Four [60] and its disformal generalisation [61]; here one introduces a homogeneous scalar field that self-adjusts (or self-tunes) dynamically to absorb the vacuum energy in such a way that it does not source curvature. We will discuss these latter two approaches in detail in §4. For further details on modifed theories of gravity, we refer the reader to the following reviews [62, 63].

Chapter 2

Problems in early universe cosmology

2.1 Inflating the early universe

2.1.1 The hot big-bang & the need for inflation

The paradigm shift instigated by GR lead to new ideas about the evolution of the universe. In particular, it was first noted by Georges Lemaître that a solution to Einstein's equations, describing an expanding universe, could be traced back in time to a single point [64]. The subsequent experimental discovery (by Edwin Hubble, [65]) that the universe is expanding prompted the development of a phenomologically successful cosmological model for the observable universe, the so-called hot big-bang model (we refer the reader to ref. [66] for a review of the HBB model). Here we shall recapitulate the successes of the HBB model, as well as give a brief overview of its cosmology. In the following discussion we shall make use of the Refs. [66], [67] and [68], and refer the reader to them for further details.

The standard HBB model has been successful at providing a theoretical explanantion for a number of important experimental observations [69], most notably:

1. it predicts the existence and spectrum of the cosmic microwave background;

2. it provides a mechanism whereby the primordial abundances of light elements in the universe are produced (nucleosynthesis);

3. it also predicts the expansion of the universe, and in addition, provides a framework in which one can describe the gravitational collapse of matter to form galaxies and other large-scale structures observed in the present-day universe.

The HBB model is based on the *cosmological principle* which states that, on scales $\gtrsim \mathcal{O}(100\text{Mpc})$, the matter distribution in the universe should be *homogeneous* and *isotropic.*[1] Qualitatively, this translates to the notion that there are no preferred locations, and no preferred directions in the universe. This provides us with a criterion for the metric describing such a spacetime (at least on the scales at which the principle is valid). Indeed, the coarse-grained spacetime is described by the Friedmann-Robertson-Walker metric,

$$\mathrm{d}s^2 \;=\; -\mathrm{d}t^2 + a^2(t)\gamma_{ij}(\mathbf{x})\mathrm{d}x^i\mathrm{d}x^j\,, \tag{2.1.1}$$

where we work in a spherical-polar comoving coordinate system[2], such that $x^\mu = (t, r, \theta, \phi)$. Spacetime is thus foliated into a set of homogeneous and isotropic spatial (space-like) hypersurfaces Σ_t. The foliation is in terms of the time coordinate t, the so-called cosmic time, which is the time (since the big-bang) as measured by an observer for which the local matter distribution is homogeneous and isotropic at each time-slice t. The metric on each spacelike hypersurface Σ_t is given by,

$$\gamma_{ij}(\mathbf{x}) \;=\; \delta^1_{\,i}\delta^1_{\,j}\,\frac{1}{1-kr^2} + \delta^2_{\,i}\delta^2_{\,j}\,r^2 + \delta^3_{\,i}\delta^3_{\,j}\,r^2\sin^2(\theta)\,, \tag{2.1.2}$$

where k is the *curvature constant*, parameterising the constant spatial curvature of each hyper-surface, which is either flat ($k = 0$), positively ($k = +1$), or negatively ($k = -1$) curved. The function $a(t)$ in eq. (2.1.1) is a dimensionful *scale factor*, parameterising the expansion of the universe as it evolves in time. Qualitatively, it characterises the relative size of spacelike hypersurfaces Σ_t at different times.

The cosmological principle implies that the background energy-momentum tensor is that of a perfect fluid, i.e. $T_{\mu\nu} = (\rho + p)\,u^\mu u^\nu + p\,g_{\mu\nu}$, where $\rho(t)$ and $p(t)$ are the energy density and pressure of the fluid, respectively. Given this, we obtain from Einstein's equations (1.2.3), the so-called *Friedmann equations* describing the evolution

[1]This is evidenced by astronomical and cosmological data which suggest that the distribution of galaxy clusters is highly homogeneous on scales of $\gtrsim \mathcal{O}(100\mathrm{Mpc})$ [70]. The isotropy of the CMB to high levels of precision [51] futhermore backs up the assumption that the universe is isotropic on these scales.

[2]By *co-moving*, we mean that a particle intially at rest within this coordinate system will remain at rest, i.e. it moves along with the Hubble flow (motion solely due to the expansion of the universe).

of an FRW universe,

$$H^2 \;=\; \frac{1}{3M_{\rm Pl}^2}\rho - \frac{k}{a^2}\,, \tag{2.1.3a}$$

$$\dot{H} + H^2 \;=\; -\frac{1}{6M_{\rm Pl}^2}\left(\rho + 3p\right)\,, \tag{2.1.3b}$$

where

$$H(t) \;=\; \frac{\dot{a}(t)}{a(t)} \tag{2.1.4}$$

is the Hubble parameter. Note that, for brevity, we have subsumed the contibution from the cosmological constant ρ_Λ into the total energy density $\rho(t)$ on the right-hand side of eq. (2.1.3a).

Observe from eq. (2.1.3a), that if the spatial geometry is flat (i.e. $k = 0$), then one obtains an expression for the *critical energy density*,

$$\rho_c(t) \;=\; 3M_{\rm Pl}^2 H^2\,. \tag{2.1.5}$$

One can then measure the densities of matter and energy relative to the critical density in terms of the *density parameter*

$$\Omega(t) \;:=\; \frac{\rho(t)}{\rho_c(t)}\,. \tag{2.1.6}$$

This is a useful quantity, as it can be used to determine the spatial geometry of the universe. Indeed, the value of $\Omega(t)$ determines the value of k, i.e. whether the universe is flat, open, or closed. Moreover, the density and pressure will generically be time dependent, such that the value of Ω, and thus the spacetime geometry changes over time. The evolution of the energy density $\rho(t)$ (and hence of $\Omega(t)$) can be determined from Eqs. (2.1.3a) and (2.1.3b). Indeed, they imply the following continuity equation

$$\dot{\rho} + 3H(\rho + p) \;=\; 0 \tag{2.1.7}$$

To fully close this equation one typically assumes that the density and pressure are related via an equation of state: $p = w\rho$, where w is the so-called equation of state parameter. Given this, and noting that $\mathrm{d}\ln a = H\mathrm{d}t$, one can recast eq. (2.1.7) as

$$\frac{\mathrm{d}\ln\rho}{\mathrm{d}\ln a} = -3(1+w)\,. \tag{2.1.8}$$

For fixed w, one can integrate this equation to get

$$\rho \propto a^{-3(1+w)}\,. \tag{2.1.9}$$

It is clear from eq. (2.1.9) how different cosmological fluids are diluted by the cosmological expansion. Indeed, for a matter dominated universe $w = 0$ and so we see that $\rho \propto a^{-3}$. For radiation domination $w = \frac{1}{3}$ and so $\rho \propto a^{-4}$, and for a universe dominated by a cosmological constant $w = -1$, such that $\rho \propto a^0$. Moreover, by inserting eq. (2.1.9) into eq. (2.1.3a) (in the case of a flat universe, $k = 0$) we then obtain the time evolution of the scale factor:

$$a(t) \propto \begin{cases} t^{2/(3(1+w))} & w \neq -1; \\ e^{Ht} & w = -1. \end{cases} \tag{2.1.10}$$

From this, we obtain the standard results that $a(t) \propto t^{2/3}$, $a(t) \propto t^{1/2}$ and $a(t) \propto e^{Ht}$, for the scale factor of a flat universe dominated by matter ($w = 0$), radiation ($w = \frac{1}{3}$) and a cosmological constant ($w = -1$), respectively. Moreover, we see that in the cases of matter or radiation domination, the Hubble parameter scales as $H \sim t^{-1}$. For reference, the present measured value of the Hubble parameter is $H_0 = H(t_0) \approx 67.6\,\mathrm{km}s^{-1}\mathrm{Mpc}^{-1} \sim 10^{-33}\,\mathrm{eV}$ [51].[3]

Before concluding our discussion on the HBB cosmology, we should emphasise the importance of the Hubble parameter $H(t)$ [eq. (2.1.4)] as a physical observable. Specifically, it describes the expansion rate of an FRW universe, thus setting its characteristic time and distance scales. One can see this by noting that $H(t)$ has units of inverse time, and as such one can construct associated time and length scales: the Hubble time $t_H = H^{-1}$ and the Hubble radius $d_H = H^{-1}$ (in units where $c = 1$). Accordingly, the time and distance scales of an FRW spacetime are characterised by the Hubble time t_H and radius d_H, respectively. By this we mean that t_H sets the

[3]Note that this is the value for H_0 as determined by the data from the Planck collaboration. It should be noted that there is currently a tension between the data sets of different experiments using different methods to measure the value of H_0 (see, e.g. refs. [71–73]). For the present discussion we shall adopt the value determined by the Planck data, and refer the reader to, e.g., ref. [74] for further details on the subject.

time-scale over which the scale factor changes appreciably ($a(t)$ roughly doubles over the course of one Hubble time). The Hubble radius d_H then sets (approximately) the distance that light can propagate over the course of one Hubble time t_H, and thus the length-scale over which causal interactions can occur within a time-scale of order t_H.[4] Given this, one can define the comoving particle (or cosmological) horizon as the maximum comoving distance $r_p(\tau)$ that a light ray can travel between the initial big-bang singularity ($t = 0$) and some later time t:

$$r_p(\tau) \;=\; \tau(t) - \tau(0) \;=\; \int_0^t \frac{\mathrm{d}t'}{a(t')} \;=\; \int_0^a \frac{\mathrm{d}a}{a^2 H} \;=\; \int_0^a \mathrm{dln}a\, \frac{1}{aH}\,. \tag{2.1.11}$$

where τ is the conformal time (defined via the relation $\mathrm{d}\tau = \frac{\mathrm{d}t}{a(t)}$). The physical size of the particle (or cosmological) horizon is then given by $d_p(t) = a(t) r_p$. Recall that in the framework of the HBB model, the 'origin of the universe' was at some finite time in the past, such that any time in its past the particle horizon was finite. An important consequence is that this limits the region of spacetime that could have been in causal contact.

We note here that one should be careful to distinguish between the comoving horizon $r_p(\tau)$ and the comoving Hubble radius $r_H = a^{-1} d_H = (aH)^{-1}$. Indeed, if particles are separated by a comoving distances greater than $r_p(\tau)$, then they never could have been in causal contact. However, if they are separated by distances greater than r_H, then they are not currently in causal contact (i.e. points that are separated by more than one Hubble radius have not been in causal contact for the last Hubble time or so). Consequently, it is possible for $r_p(\tau)$ to be much larger than r_H now, such that they cannot communicate in the present, but were in causal contact at much earlier times. This observation will be employed in our discussion of the inflationary paradigm in §2.1.2.

Having given a brief overview of the cosmology of the HBB model and highlighted its main successes, we should note that it is far from a complete model of the universe. At the very least, this is due to the fact that its description of the universe is limited to those epochs in which the temperature of the universe is sufficiently low for experimentally well understood physical processes to become well established

[4]Note that the Hubble time t_H and radius d_H are strictly *local* quantities, by which we mean that they depend on the instantaneous expansion rate at a given time t. This differs from the particle horizon d_p (cf. eq. (2.1.11)), whose value at a given time t depends on the entire expansion history of the universe up to that point.

($T \lesssim \mathcal{O}(\text{TeV})$) [66]. It is not able to account for the state of the universe at much earlier times, when it was significantly hotter, and furthermore requires very finely tuned intial conditions in order to proceed in the first place. Ultimately, a quantum theory of gravity will be required in order to understand the earliest moments of the universe in which the temperature of the universe is $\gtrsim \mathcal{O}(M_{\text{Pl}})$. Casting this issue aside for the time being, the HBB model already encounters several glaring problems when faced with observational data, and we shall briefly review them here.

We first observe that, by using eq. (2.1.6), we can recast eq. (2.1.3a) into the following form

$$\Omega - 1 \ = \ \frac{k}{(aH)^2} \,, \tag{2.1.12}$$

where Ω is the net density parameter, i.e. it is a sum of the matter, dark matter, radiation and dark energy density parameters. Now, from current observational data, we know that $|\Omega - 1| \lesssim 0.005$ [51], however, in the framework of the standard HBB model, the comoving Hubble radius $r_H = (aH)^{-1}$ is *increasing* as the universe evolves. As such, $\Omega = 1$ is an unstable critical point, which means that in order for $|\Omega - 1|$ to have its present observed value Ω must have been extremely close to unity at much earlier times, e.g. at the grand unified (GUT) scale, $|\Omega_{\text{GUT}} - 1| \lesssim \mathcal{O}(10^{-55})$. This is an incredible amount of fine-tuning and leaves one questioning why nature would have chosen parameters so precisely. This is the so-called *Flatness Problem.*

A further issue arises from the CMB data, which indicates that widely separated patches of the observable universe are almost the same temperature $\Delta T \sim \mathcal{O}(10^{-7}\text{eV})$ [51]. One could naively hope to explain this by positing that these regions have already interacted, attaining thermal equilibrium. However, this is not possible in the HBB model. Specifically, for a universe dominated by a cosmological fluid with equation of state $w = \frac{p}{\rho}$, one has that (for fixed w)

$$(aH)^{-1} \ \propto \ a^{\frac{1}{2}(1+3w)} \,. \tag{2.1.13}$$

Using this, one can then integrate eq. (2.1.11) to obtain the comoving horizon

$$r_p(\tau) \ = \ \tau \ \propto \ \frac{2}{1+3w} a^{\frac{1}{2}(1+3w)} \tag{2.1.14}$$

Note that for ordinary matter sources (i.e. those satisfying the strong energy condition $\rho + 3p > 0 \Rightarrow 1 + 3w > 0$), the initial big-bang singularity is at $\tau_{\text{intitial}} = \tau(0) = 0$, i.e. $\tau_{\text{intitial}} \propto a_{\text{intitial}}^{(1+3w)/2} = 0$, and consequently, the comoving horizon $r_p \propto a^{\frac{1}{2}(1+3w)}$ is finite. Importantly, we see that the region of the universe in causal contact increases with time. Moreover, for a matter or radiation dominated FRW universe, the particle horizon is of order the Hubble radius, i.e. $d_p \sim d_H$.[5] Given eq. (2.1.14), one can then estimate the ratio between the comoving particle horizons at the time $t_{\text{dec}} \sim 10^5$ yrs of the CMB decoupling and the present epoch $t_0 \sim 10^{10}$ yrs (noting that the universe is matter dominated in both epochs):

$$\frac{r_p(t_{\text{dec}})}{r_p(t_0)} \approx \left(\frac{t_{\text{dec}}}{t_0}\right)^{1/3} \approx \left(\frac{10^5}{10^{10}}\right)^{1/3} \approx 10^{-2}\,. \tag{2.1.15}$$

This means that comoving scales entering the horizon in the present would have been far outside the horizon at the time of the CMB decoupling. In other words, there would not have been enough time for these regions to interact before matter and radiation decoupled and the CMB was formed. The question is then, why are these a priori causally disconnected patches of spacetime almost the same temperature; why is the observable universe so homogeneous? More worryingly, this reasoning also prevents the creation of the observed fluctuations in the CMB. The HBB can be constructed such that the observed fluctuations are intrinsic to the surface of last scattering, however, these perturbations are much too large a scale for them to have been created between the big bang and the matter-radiation decoupling time. Thus the HBB model would require these perturbations to have been present in the initial conditions of the universe, which seems to be a highly contrived solution. This is the so-called *Horizon Problem*.

Finally, from the field theory sector, one also expects an abundance of 'relics', typically left over from the radiation epoch, e.g. magnetic monopoles, domain walls, etc. If such relics existed in the early universe, then their energy densities would have decreased as a matter component, thus they would have been diluted by cosmic expansion much more slowly than radiation (as $\sim a^{-3}$ opposed to $\sim a^{-4}$). As such, these massive relics can easily come to dominate the dynamics of the universe and would cause it to rapidly close in on itself ($k \to 1$). Clearly this is not what is observed, and furthermore, none of these relics have currently been experimentally observed. However, the HBB model has no way of disposing of them without also disturbing conventional matter in the universe. This is the so-called *Monopole* (or *Relic*) *Problem*.

[5] As are the comoving particle horizon and Hubble radius, i.e. $r_p \sim r_H$.

2.1.2 The inflationary paradigm

The conflicts between theory and data, present in the HBB model, motivate the need for (an) additional theory that provide(s) a more complete description. To date, the prevailing theory offering solutions (at least in part) to these problems is that of *inflation*, initially proposed by Guth in 1980 [48], whose original motivation was to explain the non-existence of magnetic monopoles, and further developed by Linde [75] and Albrecht and Steinhardt [76]. The idea of inflation is to introduce a phase of decreasing co-moving Hubble radius, $(aH)^{-1}$, in the very early universe,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[(aH)^{-1}\right] < 0\,. \tag{2.1.16}$$

Note that this implies that the universe was subject to an accelerated expansion during this period, i.e. $\ddot{a} > 0$. From Eqs. (2.1.12) and (2.1.16), we see that such an accelerated expansion dynamically drives Ω (regardless of its initial value) to unity, and thus the universe to spatial flatness. Accordingly, so long as the inflationary period is sufficiently long, Ω will be forced arbitrarily close to one, such that, despite its value being driven away from one for the remaining post-inflationary evolution of the universe, its value will remain extremely near to unity up to the present. We see then that inflation provides an elegant solution to the flatness problem (cf. fig. 2.1.1).

In fact, it is able to solve the horizon problem too. This is clear from the fact that eq. (2.1.16) implies that the comoving Hubble radius shrinks during inflation. As we discussed in §2.1.1, the Hubble radius provides a measure of how far particles can travel within the universe over a cosmological time-scale, i.e. over one Hubble time. If this is decreasing during inflation, then it implies that a much larger patch of the universe was in causal contact before inflation occurred than in the present. Hence, it would have been possible for a volume much larger than the current observable universe to have thermalised before becoming causally disconnected due to inflationary expansion.

Finally, inflation is also able to account for the lack of relic particles in the present observable universe; the accelerated expansion during inflation rapidly dilutes any initial particle densities such that they quickly become negligible. Of course, this requires that the temperature at the end of inflation must be much less than $T \sim 10^{14}$GeV, such that these relics are not produced during post-inflationary reheating.

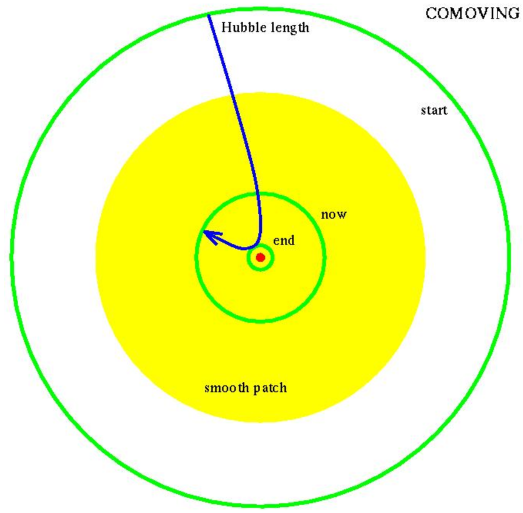


Figure 2.1.1: In the pre-inflationary universe the (comoving) Hubble radius is large, enabling a smooth patch to form via causal interactions. Inflation then precipitates a decrease in the (comoving) Hubble radius by such an amount that even the post-inflationary expansion leaves the present-day observable universe within this smoothed patch. [Source: Liddle, 1999 [66]].

To understand the dynamics of inflation in further detail, we first note that,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[(aH)^{-1}\right] \;=\; -\frac{1}{a}\left(1-\varepsilon\right)\,, \tag{2.1.17}$$

where,

$$\varepsilon \;\coloneqq\; -\frac{\dot{H}}{H^2}\,. \tag{2.1.18}$$

In order for the condition given by eq. (2.1.16) to be satisfied, we require that $\varepsilon < 1$, i.e. the fractional change in the Hubble parameter per Hubble time must be small. Hence we see that inflation corresponds to a period of quasi-de Sitter expansion, since for $\varepsilon = 0$, we have $H = \text{const.} \Rightarrow a(t) = e^{Ht}$. Of course, inflation should end in a finite amount of time and so spacetime cannot be exactly de Sitter, however, it is a good approximation when ε is small, but finite. In order to solve the problems faced by the HBB model, we need inflation to last for a sufficiently long period of time, typically measured in number of *e-folds*, $N = \int H\mathrm{d}t$. This corresponds to requiring the universe to inflate for a time-period of $N \gtrsim 60$ e-folds.[6] To see this, we note that in order for inflation to solve the horizon problem, the largest distance scales $\lambda(t_0)$ observed at the present time t_0 should be well within the horizon at the beginning of inflation t_i. Thus, upon noting that $\lambda(t_0) \sim H^{-1}(t_0)$ (cf. §2.1.1), where $H(t_0) \coloneqq H_0$ is the value of the Hubble parameter at $t = t_0$, we require that[7]

$$\lambda(t_i) \;\sim\; H_0^{-1}\left(\frac{a_f}{a_0}\right)\left(\frac{a_i}{a_f}\right) \;=\; H_0^{-1}\left(\frac{a_f}{a_0}\right)e^{-N} \;<\; H_I \tag{2.1.19}$$

where a_f is the value of the scale factor at the end of inflation, H_I is the value of the Hubble parameter during the inflationary epoch (which is approximately constant), and $N = \ln\left(\frac{a_f}{a_i}\right)$ is the number of e-folds of inflation. Let us assume that the post-inflationary transition to the radiation dominated epoch is effectively instantaneous (i.e. the conversion of inflationary energy into ultra-relativistic particles is almost instantaneous). The energy density of radiation is then $\rho_r \sim \rho_I \sim T_f^4$, where ρ_I is the inflationary energy density (which is approximately constant during inflation), T_f is the temperature at the end of inflation, and we have noted that the energy density of a radiation dominated universe scales as $T^4 \sim a^{-4}$ (cf. §2.1.1). As such, the bound on the number of e-folds is given by

$$N \;>\; \ln\left(\frac{T_0}{H_0}\frac{H_I}{T_f}\right) \;=\; \ln\left(\frac{T_0}{H_0}\right) + \ln\left(\frac{H_I}{T_f}\right) \;\sim\; 60 + \ln\left(\frac{H_I}{T_f}\right) \tag{2.1.20}$$

where $T_0 = 0.2348$ meV is the current measured temperature of the CMB [51]. Note that the precise value of the bound depends on the energy scale of inflation and on the details of reheating after inflation [68]. From this analysis, we see therefore that one requires ε to remain small for an appreciable number of Hubble times, a condition that is parametrised by

$$\eta \;\coloneqq\; \frac{\dot{\varepsilon}}{H\varepsilon}\,. \tag{2.1.21}$$

Inflation then persists for $|\eta| < 1$, since the fractional change in ε per Hubble time is small. For it to persist for a sufficiently long time, we require that $|\eta| \ll 1$.

Originally, inflation was conceived as occurring in a universe dominated by the constant energy density of a false metastable vacuum. However, this suffered from a

[6] In the following calculation we have made use of ref. [77].

[7] Here, we use that the physical distance scale $\lambda(t)$ is related to its corresponding comoving distance λ_c (which is constant with respect to Hubble expansion) as $\lambda(t) = a(t)\lambda_c$.

so-called *graceful-exit problem*, whereby classically, inflation never ends, and even though quantum mechanically it can (through tunnelling from the false to the true vacuum), in such a case the post-inflationary universe looks nothing like what we observe. To overcome this issue, modern inflationary models typically follow the so-called *slow-roll* paradigm, of which we shall elaborate on shortly.

A natural candidate for the driving force of inflation can be found though the introduction of a new scalar (i.e. spin zero) degree of freedom[8], minimally coupled to gravity. From a theoretical standing, the existence of such light scalar degrees of freedom in nature can be argued from the plethora of scalar fields present in string theory (see e.g. [81] for a discussion on moduli in string theory). Furthermore, the discovery of the Higgs boson [82] has provided the first experimental evidence for the existence of fundamental spin-zero particles, and in fact, models in which the Higgs boson plays the role of the inflaton have been constructed [83–85], and in doing so, are able to describe inflation within the framework of the SM (i.e. without needing to introduce additional degrees of freedom).[9] In this case, our new scalar field is the so-called *inflaton*, ϕ. The standard assumption in the inflationary paradigm is that, at the start of inflation, the inflaton is displaced sufficiently far from the minimum of its potential, such that its infrared modes have large occupancy numbers, and accordingly, its leading order behaviour is classical. This enables one to treat the inflaton in terms of a classical scalar field, the dynamics of which can be described by the following action,

$$S[g_{\mu\nu},\phi] \;=\; S_{\rm EH}[g_{\mu\nu}]+\int \mathrm{d}^4x\sqrt{-g}\left[-\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\mu\phi - V(\phi)\right] \;=\; S_{\rm EH}[g_{\mu\nu}]+S_\phi[g_{\mu\nu},\phi]\,, \tag{2.1.22}$$

where $S_{\rm EH}[g_{\mu\nu}]$ is the Einstein-Hilbert action [eq. (1.2.1)], and $S_\phi[g_{\mu\nu},\phi]$ is the action for the inflaton ϕ with canonical kinetic term and potential $V(\phi)$ describing the self-interactions of ϕ. The corresponding field equation and energ-momentum tensor for

[8]One can use vector degrees of freedom to describe inflation (see e.g. [78–80]), however, this typically tends to be more complex to construct such that it agrees with observational data. Fermionic fields do not tend to be used, as they are purely quantum in nature, and therefore cannot be used in a classical description (one could consider a condensate formed of Cooper pairs, however, this would simply be equivalent to using a complex scalar).

[9]We will briefly discuss the role of the Higgs boson as the inflaton at the end of this section.

ϕ are given by

$$\frac{\delta S_\phi}{\delta\phi} = \frac{1}{\sqrt{-g}}\partial_\mu\big(\sqrt{-g}\partial^\mu\phi\big) - V'(\phi) = 0\,, \tag{2.1.23a}$$

$$T^{(\phi)}_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S_\phi}{\delta g^{\mu\nu}} = \partial_\mu\phi\partial_\nu\phi - g_{\mu\nu}\Big(\frac{1}{2}\partial^\lambda\phi\partial_\lambda\phi + V(\phi)\Big)\,. \tag{2.1.23b}$$

Now, for inflation to start it must be seeded by at least one small region of homogeneous spacetime (at least of order the pre-inflationary Hubble radius, $\sim H^{-1}$ [86–89]). Within this region the classical field configuration is homogeneous. Moreover, since spatial gradients scale as $a^{-1}\nabla$ in an FRW spacetime, inflation will rapidly smooth out any spatial variations. Consequently, throughout the majority of inflation the inflaton field ϕ is well approximated by its homogeneous configuration (this will certainly be true after a couple of e-folds of inflation). Furthermore, by its very construction, inflation drives the observable universe towards an FRW background, and as such we shall take the metric $g_{\mu\nu}$ to be the FRW solution [eq. (2.1.1)]. Consequently, we can neglect any spatial gradients for ϕ such that its dynamics are governed by the homogeneous Klein-Gordon equation,

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0\,, \tag{2.1.24}$$

with the FRW geometry determined by the Friedmann equation

$$H^2 = \frac{1}{3M_{\rm Pl}^2}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right)\,. \tag{2.1.25}$$

Moreover, the energy-momentum tensor [eq. (2.1.23b)] for ϕ is that of a perfect fluid

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi)\,, \tag{2.1.26a}$$

$$p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi)\,. \tag{2.1.26b}$$

The corresponding equation of state is then

$$w_\phi = \frac{p_\phi}{\rho_\phi} = \frac{\frac{1}{2}\dot{\phi}^2 - V(\phi)}{\frac{1}{2}\dot{\phi}^2 + V(\phi)}\,, \tag{2.1.27}$$

from which we see, that if the potential of ϕ dominates over its kinetic energy, $\frac{1}{2}\dot{\phi}^2 \ll V(\phi)$, then $w_\phi \approx -1$ such that ϕ has negative pressure, and furthermore can drive an accelerated expansion, as required. Additionally, since $p_\phi \approx -\rho$ we see from eq. (2.1.7), that in this case $\dot{\rho}_\phi \approx 0 \Rightarrow \rho_\phi \approx V(\phi) \approx$ const., i.e. the energy density of ϕ is not diluted by the expansion of the universe whilst the condition $\frac{1}{2}\dot{\phi}^2 \ll V(\phi)$ holds.

With this information we can then determine the form of the parameters ε and η that set the conditions for the required inflation. From Eqs. (2.1.18) and (2.1.21), these are given by,

$$\varepsilon \;=\; \frac{\frac{1}{2}\dot{\phi}^2}{M_{\rm Pl}^2 H^2}\,, \quad \eta \;=\; 2\left(\varepsilon + \frac{\ddot{\phi}}{H\dot{\phi}}\right) \;=\; 2(\varepsilon - \delta)\,, \tag{2.1.28}$$

where $\delta \coloneqq -\frac{\ddot{\phi}}{H\dot{\phi}}$. For inflation to persist for long enough we require that $\{\varepsilon, |\delta|\} \ll 1$, such that both H and ε have only small fractional changes per e-fold, i.e. $\{\varepsilon, |\eta|\} \ll 1$. Given this, we note that the condition $\varepsilon \ll 1$ implies that the kinetic energy $\frac{1}{2}\dot{\phi}^2$ must only make a small contribution to the total energy density, i.e. $\frac{1}{2}\dot{\phi}^2 \ll V(\phi)$. This enables us to make the following approximation for the Friedmann equation (2.1.25),

$$H^2 \approx \frac{1}{3M_{\rm Pl}^2} V(\phi)\,. \tag{2.1.29}$$

Moreover, the condition $|\delta| \ll 1$ implies that $|\ddot{\phi}| \ll |H\dot{\phi}|$, leading to the following simplification of the Klein-Gordon equation (2.1.24),

$$3H\dot{\phi} \approx -V'(\phi)\,. \tag{2.1.30}$$

This simplification of the equations of motion is the so-called *slow-roll* approximation. This approximation then provides a useful way to assess whether or not a particular potential $V(\phi)$ can support inflation, via the introduction of so-called *potential slow-roll parameters*, ε_V and η_V. Motivated from our definitions of the slow-roll parameters ε and η, we then use the approximate equations of motion (2.1.29) and (2.1.30), to define ε_V and η_V,

$$\varepsilon \approx \frac{M_{\rm Pl}^2}{2}\left(\frac{V'}{V}\right)^2 \;\coloneqq\; \varepsilon_V\,, \quad \delta + \eta \approx M_{\rm Pl}^2 \frac{V''}{V} \;\coloneqq\; \eta_V\,, \tag{2.1.31}$$

which are related to ε and η during slow-roll, respectively, as follows: $\varepsilon_V \approx \varepsilon$ and $\eta_V \approx 2\varepsilon - \frac{1}{2}\eta$. We see, therefore, that slow-roll inflation occurs when $\{\varepsilon_V, |\eta_V|\} \ll 1$. Note from eq. (2.1.31), that the inflaton potential does not have to be (approximately) flat for inflation to occur, but rather its first and second derivatives have to be small relative to V itself. This is the case in so-called large-field inflationary models (see e.g. Refs. [90, 91]), in which one can have a steep potential, however the slow roll conditions can still be satisfied since the Hubble friction term $3H\dot{\phi}$ in eq. (2.1.24) damps the velocity of the scalar field sufficiently enough that the condition $\frac{1}{2}\dot{\phi}^2 \ll V(\phi)$ is satisfied. Note, however, that in such models, for inflation to persist for $N \sim 60$ e-foldings requires ϕ to undergo super-Planckian field excursions[10].

Qualitatively, slow-roll inflation can be visualised classically as a ball rolling down its potential (cf. fig. 2.1.2). The inflaton is initially displaced from the minimum of its potential at the start of inflation, and then as inflation proceeds it begins to slowly roll towards this minimum. Inflation ends when the slow-roll conditions are no longer satisfied, i.e. $\varepsilon(\phi_{\text{end}}) = 1$, $\varepsilon_V(\phi_{\text{end}}) \approx 1$. At this point ϕ will transition from being overdamped to underdamped and start to move rapidly on the Hubble time-scale, falling into the minimum of its potential, about which it oscillates. There are

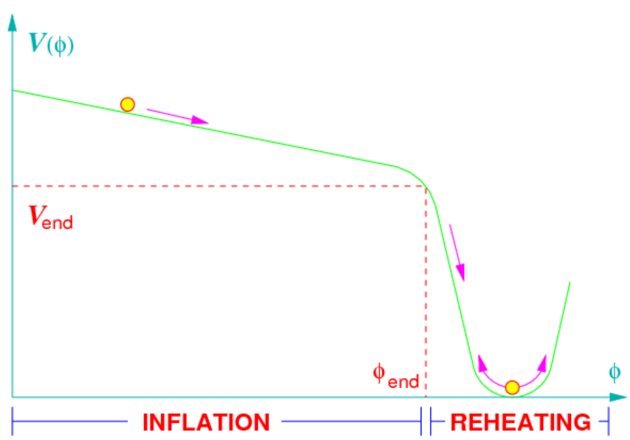


Figure 2.1.2: Schematic representation of the inflaton potential (the actual shape of the potential is model dependent). The inflaton ϕ (yellow circle) slowly rolls down its potential until it reaches a critical value ϕ_{end} (corresponding to $V_{\text{end}} = V(\phi_{\text{end}})$), at which point inflation ends and it starts to oscillate about its VEV. [Source: Dimopoulos, 2010 [92]].

[10]One may worry that the fact that ϕ can assume super-planckian values invalidates our semiclassical analysis, however, as long as the potential satisfies $V(\phi) \lesssim M_{\text{Pl}}$ we are safe to treat both gravity and the dynamics of the inflaton field classically.

a large number of different inflationary models, broadly classified in terms of *small-* and *large*-field inflation, corresponding to whether or not the amplitude of the inflaton field can take on super-Planckian values. For example, all large-field models with a sufficiently flat potential have a limit in which they reduce to the sub-class of so-called *chaotic inflation* [90, 91], where the inflaton potential is polynomial in form $V(\phi) \sim \frac{\lambda\phi^n}{M_{\rm Pl}^{n-4}}$, for $\phi \gtrsim M_{\rm Pl}$ (where $n > 0$, $0 < \lambda \ll 1$). Another example is Higgs inflation [83], where the SM Higgs boson plays the role of the inflaton. In this case, the potential is given by the Higgs effective potential $\sim \frac{\lambda}{4}\left(\phi^2 - v^2\right)^2$; one also requires a non-minimal coupling of the Higgs boson to gravity $\sim \xi\phi^2 R$ in order to tame the self-coupling of the Higgs, such that the matter perturbations it produces agree with observational data. Further examples include *k-inflation* [93] and *hybrid inflation* [94]. For a brief review of several of these models, we refer the reader to, e.g. [68] and [95]. Finally, it is worth noting that, the current best cosmology we have combines the HBB and inflationary models into the so-called Lambda Cold Dark Matter model, or ΛCDM, including a cosmological constant Λ to account for dark energy and a cold dark matter content in the matter distribution. ΛCDM is in strong agreement with current experimental observations, as evidenced by the most recent Planck data [51], and analogous to the SM, provides a standard cosmological model (see e.g. [96] for a brief review).

2.2 Post-inflationary reheating

In the slow-roll regime, the energy density of the inflaton field ϕ remains approximately constant, as opposed to the energy densities of all forms of matter, which are red-shifted to negligible amounts. As such, at the end of inflation, practically all of the energy density in the universe is stored in the inflaton field ϕ. Thus, we need a mechanism by which the vast amount of energy held in ϕ can be transferred back to the matter fields, thus producing the abundances of SM particles that we observe in the present. This post-inflationary stage is referred to as *reheating*.

Reheating begins as soon as the inflaton field ϕ begins to oscillate about its VEV - as it passes through the minimum of its potential $V(\phi)$, it transfers energy to the matter fields that it is coupled to. From a quantum mechanical perspective, this can be understood as follows. The inflaton is initially displaced from the minimum of its potential such that its infrared modes are densely populated (this ensures that the field can satisfy the slow-roll conditions - large occupations of shorter wavelength modes would violate these conditions). As such, the leading order behaviour of the field is classical. Moreover, as soon as inflation starts, the shorter wavelength modes of $\hat{\phi}$ are rapidly redshifted (as they scale as $a^{-1}|\mathbf{k}|$) and quickly become super-horizon length, at which point they become "frozen-in" as classical fluctuations. This leads to the build up of a condensate in the zero mode of $\hat{\phi}$, such that its background (classical) field value can be well approximated by its homogeneous configuration. As such, we can represent the quantum inflaton field ϕ as a sum of its background value $\varphi(t) = \langle\hat{\phi}(t,\mathbf{x})\rangle$ (corresponding to the expectation value of $\hat{\phi}$ in a translationally invariant vacuum state) and its quantum fluctuations $\delta\hat{\phi}(t,\mathbf{x})$, i.e. $\hat{\phi}(t,\mathbf{x}) = \varphi(t) + \delta\hat{\phi}(t,\mathbf{x})$. The quantum fluctuations of $\hat{\phi}$ correspond to small inhomogeneities in the pre-inflationary background. These lead to disparate regions of space inflating by different amounts. Such disparaties in the local expansion histories result in differences in the local energy densities post-inflation. These perturbations seed the formation of structure in the post-inflationary universe, observationally corresponding to small anisotropies in the CMB (see e.g. Refs. [68, 91, 97] for further details).

We shall now elaborate further on the details of post-inflationary reheating, first we briefly review the perturbative analysis and then move on to discuss the non-perturbative aspects of reheating that arise due to the coherent nature of the inflaton condensate. In the following two subsections, we have made use of Refs. [98], [99] and [100], and we refer the reader to them for further details.

2.2.1 The perturbative approach

At the end of inflation, we are left with a coherently oscillating condensate of zero momentum inflaton quanta. Naively, one then assumes that it is possible to treat the decay of the inflaton field perturbatively, i.e. that individual quanta in this condensate decay independently of each other into SM (or other intermediary) particles. Following this approach, one considers the inflaton Lagrangian with an effective potential $V(\phi)$ and additional interaction terms, describing interactions of the inflaton field ϕ with a scalar field χ and a fermion field ψ,

$$\mathscr{L} = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi - V(\phi) - \frac{1}{2}\partial^\mu\chi\partial_\mu\chi - \frac{1}{2}m_\chi^2\chi^2 + \bar{\psi}\left(\gamma^\mu\partial_\mu - m_\psi\right)\psi - \frac{1}{2}g\phi^2\chi^2 - h\bar{\psi}\psi\phi\,, \tag{2.2.1}$$

where m_χ and m_ψ are the bare masses of the χ and ψ fields, g and h are the coupling constants parameterising the strength of the interactions, and $\bar{\psi} = i\psi^\dagger\gamma^0$ is the Dirac adjoint spinor field, with γ^μ the Dirac matrices. For generality, it is assumed that the effective potential has a minimum at $\phi = \sigma$, which we assume to be quadratic in form near this minimum, i.e.,

$$V(\phi) \sim \frac{1}{2} m_\phi^2 (\phi - \sigma)^2 \,, \tag{2.2.2}$$

where m_ϕ is the effective mass of the inflaton field. We can make a field redefinition $\phi \to \phi + \sigma$ such that the Lagrangian [eq. (2.2.1)] now contains interactions between ϕ, χ and ψ that are linear in ϕ,

$$\Delta\mathscr{L} \;\supset\; -g\sigma\phi\chi^2 - h\bar{\psi}\psi\phi \,. \tag{2.2.3}$$

Given this, we will now study the effects of the Hubble expansion and particle creation, arising from interactions with other fields, on the dynamics of the inflaton field. In the early perturbative studies of reheating, typically the dissipation of energy from the inflaton condensate was treated from a phenomenological point of view, by including a dissipative term in its EOM [101–103]

$$\ddot{\phi} + 3H(t)\dot{\phi} + \Gamma_{\text{tot}}\dot{\phi} + V'(\phi) \;=\; 0 \,, \tag{2.2.4}$$

where Γ_{tot} is the total interaction (or decay) rate for interactions between the inflaton and scalar and fermionic fields arising from eq. (2.2.3). To obtain an approximate solution to eq. (2.2.4), let us consider the case where $V(\phi) = \frac{1}{2} m_\phi^2 \phi^2$. Given this, we first note that post-inflation $m_\phi \gg H \sim t^{-1}$, and so the condensate undergoes many oscillations over one Hubble time. As such, if $m_\phi \gg \Gamma$ also holds, and we neglect the time-dependence of H and Γ due to the expansion of the universe, then eq. (2.2.4) has the following approximate solution

$$\phi(t) \;\approx\; \phi_0 a^{-3/2}(t) e^{-\frac{1}{2}\Gamma_{\text{tot}} t} \cos(m_\phi t) \,, \tag{2.2.5}$$

where ϕ_0 is the amplitude of the inflaton condensate at the start of reheating. Such a solution describes damped oscillations of ϕ near the point $\phi = 0$. In particular, we see that the amplitude of ϕ decreases due to both the expansion of the universe and particle production. To estimate Γ_{tot}, we consider the flat-space limits of the contributing decay rates. Referring back to eq. (2.2.3), we see that ϕ can decay into two scalar particles or a fermion-anti-fermion pair. At tree-level, these interactions are given by the Feynman diagrams in fig. 2.2.1, with corresponding scattering am-

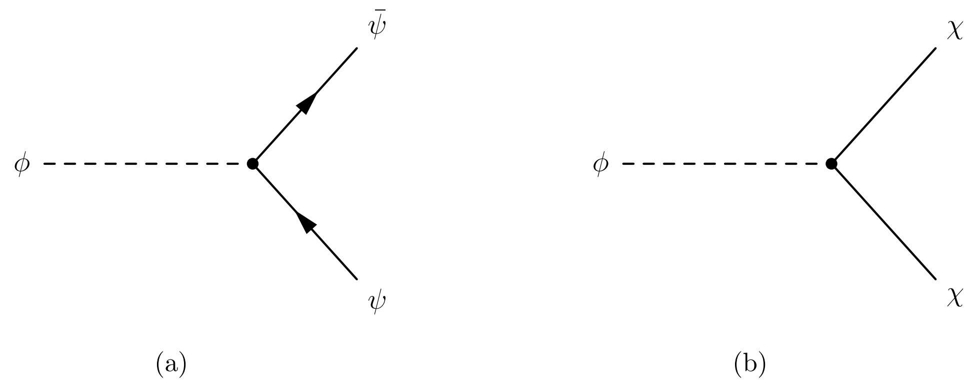


Figure 2.2.1: Tree-level Feynman diagrams for the perturbative decay of the an inflaton particle into: (a) two scalar particles and (b) a fermion-anti-fermion pair.

plitudes $\mathcal{M}_{\phi\to\chi\chi} = -2ig\sigma$ and $\mathcal{M}_{\phi\to\bar{\psi}\psi} = -ih\bar{u}v$, respectively, where u and v are (four-) spinor particle and anti-particle (momentum space) solutions to the Dirac equation, corresponding to a fermion-anti-fermion pair (where we have suppressed the spin indices and momentum dependence of $\bar{u}$ and v for brevity). The decay rates for $\phi \to \chi\chi$ and $\phi \to \bar{\psi}\psi$ are then

$$\Gamma\left(\phi \to \chi\chi\right) = \frac{g^2\sigma^2}{8\pi m_\phi}\,, \tag{2.2.6a}$$

$$\Gamma\left(\phi \to \bar{\psi}\psi\right) = \frac{h^2 m_\phi}{8\pi}\,, \tag{2.2.6b}$$

where we have assumed that $m_\phi \gg m_\chi, m_\psi$. Returning to the solution for ϕ, eq. (2.2.5), we see that the amplitude of the condensate $\Phi(t) \simeq \phi_0 a^{-3/2}(t) e^{-\frac{1}{2}\Gamma_{\rm tot} t}$ satifies,

$$\frac{1}{a^3}\frac{\mathrm{d}}{\mathrm{d}t}\left(a^3\Phi^2\right) = -\Gamma_{\rm tot}\Phi^2\,, \tag{2.2.7}$$

where $\Gamma_{\rm tot} = \Gamma(\phi \to \bar{\psi}\psi) + \Gamma(\phi \to \chi\chi)$. For time intervals larger than m_ϕ^{-1} one approximate $\dot{\phi}^2(t) \approx m_\phi^2\Phi^2(t)\sin^2(m_\phi t)$, such that the number density of inflaton quanta can be related to the amplitude $\Phi(t)$ of the condensate as $N_\phi = \frac{1}{m_\phi}\rho_\phi \approx \frac{1}{2}m_\phi\Phi^2$ [98]. Thus, upon multiplying both sides of eq. (2.2.7) by m_ϕ, we arrive at a rate equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(a^3 N_\phi\right) = -a^3\Gamma_{\rm tot}N_\phi\,, \tag{2.2.8}$$

describing the exponential decay of the total comoving number density of inflaton quanta, with decay rate $\Gamma_{\rm tot}$.

Note that, in theories without a non-zero minimum of the corresponding effective potential (where $\sigma = 0$), reheating can still can occur due to the presence of the oscillating inflaton condensate ϕ. Indeed, let us return to the Lagrangian [eq. (2.2.1)], and consider the quartic interaction $\mathscr{L} = -\frac{1}{2}g\phi^2\chi^2$, neglecting any fermion interactions for the time being. When the amplitude Φ of the oscillating condensate is sufficiently small ($\Phi \ll M_{\rm Pl}$) one may expect perturbation theory to be applicable. Given this, one can then consider the perturbative annihilation of two inflaton particles (at rest) into two χ particles $\phi\phi \to \chi\chi$. In so doing, it is found that the corresponding annihilation cross-section (at zero relative velocity) is given by $\left[\sigma_{\phi\phi\to\chi\chi}|\mathbf{v}_{\rm rel}|\right]\big|_{|\mathbf{v}_{\rm rel}|=0} \approx \frac{g^2}{64\pi {m_\phi}^2}$ [104]. One can then estimate the scattering rate for this process as $\Gamma(\phi\phi \to \chi\chi) \approx \left[\sigma_{\phi\phi\to\chi\chi}|\mathbf{v}_{\rm rel}|\right]\big|_{|\mathbf{v}_{\rm rel}|=0} N_\phi \sim \frac{g^2\Phi^2}{8m_\phi}$ (where we have used that $N_\phi \approx \frac{1}{2}m_\phi\Phi^2$). However, such a decay rate is problematic, since Φ^2 decreases as t^{-2}, whereas the Hubble parameter H decreases as t^{-1}. Consequently, the rate of decay of the inflaton field never catches up with the rate of expansion of the universe, and so reheating never completes. This would be a disaster, as an incomplete decay of the inflaton field would result in the matter and energy that was created during reheating being redshifted to the point where the present day universe would be cold, empty and devoid of life, something that certainly does not match observation. In fact, reheating will only complete if Γ decreases slower than t^{-1}, which typically requires spontaneous symmetry breaking, and/or coupling of the inflaton to fermions with $m_\psi < m_\phi/2$ [98].

The requirement that reheating does complete thus places important constraints on the structure of any theory that aims to describe it. It is also important to note that perturbative decays are not the only option for reheating. Indeed, if the initial amplitude of the inflaton condensate is large enough, then non-perturbative effects can play an important role in the reheating process. We shall discuss this in further detail in §2.2.2.

For now, we assume that reheating does complete solely due to perturbative decay of the inflaton field, and that the system thermalises almost instantaneously upon reheating terminating, at which point the universe enters an epoch of radiation domination. Given this, we can estimate the time at which reheating ends, and the corresponding temperature of the universe soon after this occurs. With this in mind, we note that during the post-inflationary oscillatory phase of the inflaton field, it still provides the dominant contribution to the total energy density of the universe. The behaviour of the universe during this phase is thus dictated by the inflaton condensate, which is that of a fluid of non-relativistic particles of mass m_ϕ, and as such $H(t) \sim \frac{2}{3t}$. From the assumption that $m_\phi \gg m_\chi$, m_ψ, the decay products of the inflaton are ultra-relativistic, and as such their energy densities decrease much faster ($\sim a^{-4}$) than that of the inflaton condensate ($\sim a^{-3}$). Consequently, reheating only terminates when $H < \Gamma_{\text{tot}}$ (otherwise most of the energy available will remain stored in the inflaton condensate). This occurs when the age of the universe is $t_{\text{rh}} \sim \frac{2}{3}\Gamma_{\text{tot}}^{-1}$ (referred to as the *reheat time*), at which point, the matter content in predominantly ultra-relativistic. Using the first Friedmann equation (2.1.3a), we can estimate the energy density of the universe at the reheating time,

$$\rho(t_{\text{rh}}) \simeq \frac{4M_{\text{Pl}}^2}{3t_{\text{rh}}^2} \; = \; 3M_{\text{Pl}}^2\Gamma_{\text{tot}}^2 \,. \tag{2.2.9}$$

The assumption is then, once the inflaton has completely decayed, these ultra-relativistic particles quickly equilibriate, forming a thermal bath of temperature T_{rh}, the so-called reheat temperature, with energy density

$$\rho(t_{\text{rh}}) \simeq 3M_{\text{Pl}}^2\Gamma_{\text{tot}}^2 \simeq \frac{\pi^2}{30} g_* T_{\text{rh}}^4 \,, \tag{2.2.10}$$

where g_* quantifies the total number of relativistic degrees of freedom. Thus, one arrives at an estimate for the reheat temperature,

$$T_{\text{rh}} \simeq \left(\frac{90M_{\text{Pl}}^2}{\pi^2 g_*}\Gamma_{\text{tot}}^2\right)^{1/4} \simeq 0.1\sqrt{M_{\text{Pl}}}\Gamma_{\text{tot}} \,, \tag{2.2.11}$$

where we have used $g_* \sim \mathcal{O}(100)$ (expected to be true in realistic models [98]). The reheat temperature is an important observable that one hopes to extract a prediction for from a given reheating theory, since it enables one to determine the thermal history of the universe. Indeed, its value has important consequences for leptogenesis (see, e.g., ref. [105]) and for the generation of dark matter relic densities (see, e.g., ref. [106]). Importantly, the reheat temperature should be larger than a few MeV to allow for the standard big-bang nucleosynthesis [107]. Requiring that thermalisation occurs before nucleosynthesis imposes an upper bound on the inflaton mass as a function of the reheat temperature [108]. In the case of supersymmetric theories, the potential over-production of gravitinos [109, 110] can spoil the generation of the light elements, providing an upper bound on the reheat temperature.

Before continuing, we should note that are some caveats to the above analysis. In particular, the heuristic EOM [eq. (2.2.4)], whilst looking physically intuitive, does not provide a consistent description of the inflaton dynamics during perturbative decay. This is because it violates the fluctuation-dissipation theorem, which states that dissipative processes necessarily lead to fluctuations within the system [100]. Such fluctuations would impact on the effective mass of ϕ, which is not accounted for in eq. (2.2.4). More careful perturbative analyses have been conducted, however, in which such fluctuations are included in the effective EOM for ϕ, see e.g. [111, 112]. A further issue with eq. (2.2.4), is that whilst it approximates the impact of the dissipative processes on the dynamics of the inflaton condensate, it neglects the fact that the decay products themselves are dynamical. To describe the full dynamics of the system would require one to solve a coupled set of self-consistent equations describing the evolution of both the inflaton condensate and the decay products.

Aside from the issue of correctly describing the effects of perturbative processes on the dynamics of the inflaton condensate, one is presented with the problem that a perturbative analysis of the inflaton decay completely neglects the effects of Bose condensation. These can greatly enhance the naive decay rates calculated above (as we will discuss in §2.2.2). Moreover, the fact that the inflaton condensate undergoes coherent oscillations gives rise to parametric resonances of the fields coupled to it, which can lead to explosive particle production and a corresponding exponential growth in particle number densities. Such resonant effects, which are particularly prevalent in cases where the amplitude of these coherent oscillations is large, cannot be accounted for within the elementary perturbative description. Through a non-perturbative analysis, however, it is possible to account for these phenomenona, and we shall elaborate on this further in the next subsection.

2.2.2 The non-perturbative nature of reheating

The perturbative approach to reheating is appealing, but is not sufficient to fully describe the post-inflationary reheating process. As we have mentioned, it neglects the effects of Bose condensation, in which the decay of an inflaton into bosonic particles is enhanced if the corresponding final states are already occupied. Furthermore, in the perturbative analysis, one treats the condensate as a collection of independent inflaton quanta decaying independently into bosonic, or fermion-anti-fermion pairs of particles. However, at the end of inflation, the inflaton condensate ϕ cannot be well-described by a superposition of asymptotically free single inflaton states, decaying independently into SM (or other intermediary) particles. Rather, it behaves as a coherently oscillating homogeneous field. This (classical) time-dependent background field can induce quantum mechanical production of matter particles, through driving parametric resonances of the fields it is coupled to. Such resonances are typically much more efficient than single-body decays at transferring energy from the inflaton condensate to coupled matter fields. The result is an extremely rapid and highly non-adiabatic production of particles across multiple bands of momenta. Moreover, this resonant production can be interpreted in terms of simultaneous collective decays of many inflaton quanta (which correspond to higher order Feynman diagrams). Therefore, to fully capture such behaviour necessitates a non-perturbative description[11]. A non-perturbative analysis of reheating was first conducted in detail by Kofman, Linde and Starobinsky [98, 113]. They found that the resonant particle production arising from the coherent oscillations of the inflaton condensate dominates over perturbative decays during the early stages of reheating.

Having mentioned the phenomenon of Bose enhancement, let us now give a heuristic analysis of how such effects arise. As reheating progresses the χ particle states will generally have non-zero number densities. The result of this is that the probability of the inflaton decaying into χ particles within these states is subject to so-called Bose enhancement. As an example, let us work in the perturbative regime and consider the process $\phi \to \chi\chi$. If the corresponding momentum states already have non-zero number density, then the effective decay rate of the inflaton Γ_{eff} is enhanced, i.e. $\Gamma_{\text{eff}} \simeq \Gamma_{\phi\to\chi\chi}(1 + 2N_{|\mathbf{k}|})$ (where $N_{\mathbf{k}}$ is the occupation number of the given momentum state) [114]. One can heuristically derive this by considering the decay of an inflaton particle into two χ particles, i.e. $\phi \to \chi\chi$. Since ϕ is at rest (i.e. has zero momentum) the produced χ particles will then have back-to-back momenta, with magnitude $|\mathbf{k}| \sim m_\phi/2$. Moreover, the inverse process $\chi\chi \to \phi$ can also occur,

[11] Note further, that the coherent oscillations of inflaton condensate can be large, and so even if the coupling between the condensate and other fields is small, this does not guarantee that perturbation theory is accurate.

with the rates of the two interactions being respectively proportional to

$$|\langle N_\phi - 1,\, N_{\mathbf{k}} + 1,\, N_{-\mathbf{k}} + 1|\hat{a}^\dagger_{\mathbf{k}}\hat{a}^\dagger_{-\mathbf{k}}\hat{a}^\phi_{\mathbf{0}}|N^\phi_{\mathbf{0}},\, N_{\mathbf{k}},\, N_{-\mathbf{k}}\rangle|^2 \;=\; (1+N_{\mathbf{k}})(1+N_{-\mathbf{k}})N^\phi_{\mathbf{0}} \tag{2.2.12a}$$

$$|\langle N^\phi_{\mathbf{0}} + 1,\, N_{\mathbf{k}} - 1,\, N_{-\mathbf{k}} - 1|\hat{a}_{\mathbf{k}}\hat{a}_{-\mathbf{k}}(\hat{a}^\phi_{\mathbf{0}})^\dagger|N^\phi_{\mathbf{0}},\, N_{\mathbf{k}},\, N_{-\mathbf{k}}\rangle|^2 \;=\; N_{\mathbf{k}}N_{-\mathbf{k}}(1+N^\phi_{\mathbf{0}}) \tag{2.2.12b}$$

where $\hat{a}^{(\dagger)}_{\mathbf{k}}$ and $(\hat{a}^\phi_{\mathbf{0}})^{(\dagger)}$ are the creation (annihilation) operators of χ particles and (the zero-mode) inflatons, respectively. Note that the and $N^\phi_{\mathbf{0}}$ is the occupation number of zero-momentum state of the inflaton field, and since the field has formed a condensate in this state at the end of inflation, its corresponding occupation number is large, i.e. $N^\phi_{\mathbf{0}} \gg 1$. Given this, one can compute the rate of change of (total) χ particle number density $N_\chi = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} N_{\mathbf{k}}$ per comoving volume V:

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(a^3 N_\chi\big) \;=\; \frac{2\,a^3}{V}\Gamma_{\phi\to\chi\chi}\left[(1+N_{\mathbf{k}})(1+N_{-\mathbf{k}})N^\phi_{\mathbf{0}} - N_{\mathbf{k}}N_{-\mathbf{k}}(1+N^\phi_{\mathbf{0}})\right] \;\approx\; 2\,a^3\Gamma_{\text{eff}}N_\phi \tag{2.2.13}$$

where $N_\phi = \frac{N^\phi_{\mathbf{0}}}{V} \approx \frac{m_\phi \Phi^2}{2}$, the factor of two accounts for the fact that a single inflaton is decaying into two χ particles, and we have noted that $N_{\mathbf{k}} = N_{-\mathbf{k}}$, with $|\mathbf{k}| = m_\phi/2$. Similarly, the comoving rate of change of inflaton quanta (considering $\phi \to \chi\chi$ processes only) is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(a^3 N_\phi\big) \;=\; -a^3\Gamma_{\text{eff}}N_\phi \tag{2.2.14}$$

We see then that Γ_{eff} can be significantly larger than the naive decay rate $\Gamma_{\phi\to\chi\chi}$, and our elementary treatment of reheating starts to break down as soon as $N_{\mathbf{k}} > 1$. This occurs very soon after reheating begins [114].

In the above discussion, we considered the case of bosonic particles. Fermions are not subject to Bose enhancement due to Pauli blocking, and following this line of thought, it was initially thought that parametric resonance of fermionic fields could not occur. However, Greene and Kofman have shown that fermions can be efficiently produced though parametric resonance whilst still adhering to the Pauli exclusion principle [115, 116]. The analysis was further developed by Berges et al., who showed that this efficiency is much higher if one takes into account backreactions of the fermions on the inflaton condensate [117]. Interestingly, one can produce super-heavy fermionic particles during parametric resonance, essentially as a consequence of the Yukawa coupling between the fermionic field and the condensate, enabling the effective mass of the fermions to become vanishingly small at certain points in the oscillation cycle of the inflaton condensate [118]. This is something that cannot occur for the bosonic case, since its effective mass cannot be reduced by an oscillating condensate, and so production of super-heavy bosonic particles is heavily suppressed [119].

Despite the importance of non-perturbative effects during the early stages of reheating, the elementary theory of perturbative inflaton decays is still applicable during its latter phase, once the amplitude of the inflaton oscillations becomes sufficiently small and parametric resonance has terminated. In fact, reheating never ends in the regime of parametric resonance. Due to the fact that the inflaton condensate decays (redshifted by Hubble expansion and transference of energy to the matter fields), eventually the bands of momentum space in which resonance can occur become vanishingly small. This occurs before the condensate is completely depleted of its energy density, such that particle production continues, however is subsequently dominated by perturbative decays of inflaton quanta. As this initial, non-perturbative stage of the reheating process precedes the perturbative phase, and the ensuing thermalisation of created particles, it is referred to as *preheating* [98, 113]. We shall return to the subject of preheating in §5, where we shall discuss it in more detail.

Chapter 3

The cosmological constant problem

3.1 Overview

As we briefly discussed in the introduction (cf. §1.4), the CCP is an issue that arises within the framework of semi-classical gravity, in which one treats the gravitational sector classically, but the matter sector quantum mechanically. More concretely, it is a result of combining the two cornerstones of our current understanding of fundamental physics, that is, GR and the SM. Indeed, QFT predicts that the vacuum state of a given theory is highly non-trivial, in particular, it has a non-zero energy density attributed to it, a so-called *vacuum energy* (density). In Minkowski spacetime this does not present an issue since in this case QFT is insensitive to absolute energies: two QFTs differing only by a constant energy are equivalent, since one can add an arbitrary constant the Lagrangian without affecting the equations of motion, and hence, the physics. In this sense, in flat spacetime the vacuum energy is not part of the physical content of any given QFT, rendering it impossible, even in principle, to make any predictions about it. This drastically changes, however, when one "switches on" gravity; GR tells us that gravity is sensitive to absolute energies, and from the equivalence principle we know that all forms of energy source spacetime curvature. As such there is no reason a priori why vacuum energy should be an exception to this, and accordingly, in the context of GR, vacuum energy is a physical quantity, since shifting it by some amount will affect the curvature of spacetime.

From an EFT perspective, we therefore expect that all massive particles, up to the cut-off energy scale of the theory (in this case $\sim \mathcal{O}(\text{TeV})$, but potentially as high as the Planck mass, M_{Pl}), should contribute. Given that the spacetime vacuum is observed to be Lorentz invariant to a high degree of accuracy [120], the only way we can introduce a vacuum energy density term into the energy-momentum tensor is for it to be proportional to the metric, i.e. $T^{\text{vac}}_{\mu\nu} = -\rho_{\text{vac}} g_{\mu\nu}$. We therefore expect the vacuum energy density contribution to the gravitational action to be of the form,

$$-\rho_{\text{vac}} \int \mathrm{d}^4x \sqrt{-g}\,. \tag{3.1.1}$$

There is a clear sense in which the cosmological constant manifests itself as a problem in the long wavelength, i.e., infrared sector of GR. The reason being is that gravitons probing dark energy are ultrasoft, with wavelengths of order the current Hubble length, i.e. $\lambda \sim H_0^{-1}$. As such, the effect of vacuum energy is most pronounced on cosmological scales. Moreover, at late times, the vacuum energy density will dominate over all other forms of energy density (ρ_{vac} does not redshift, whereas matter and radiation will have redshifted to negligible amounts). Accordingly, from eq. (2.1.3a), the late-time curvature will be $H^2 \approx \frac{1}{3M_{\text{Pl}}^2}\rho_{\text{vac}}$, and so if $\rho_{\text{vac}} > 0$, then this yields an accelerated de Sitter expansion at late times. Given that the value of the Hubble parameter in the present epoch is measured to be $H_0 \sim 10^{-33}$eV, this sets an upper-bound for the vacuum energy density, $\rho_{\text{vac}} \lesssim (\text{meV})^4$. If ρ_{vac} exceeded this bound, then it would have started to dominate long before the present epoch, and the corresponding cosmological horizon ($\sim H^{-1}$) would be much smaller than we observe today.

A natural question arising when considering these issues is, does vacuum energy truly exist, i.e. does it produce experimentally observable effects? To answer this, we first need to recall how vacuum fluctuations contribute to physical processes. In the standard framework of perturbative QFT, one can identify the contributions from the vacuum fluctuations of a particular theory with radiative, so-called *loop* corrections to tree level scattering amplitudes of physical processes. For example, consider a particular interacting QFT, quantum electrodynamics, described by the following Lagrangian,

$$\mathscr{L}_{\text{QED}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}\left(\not{\partial} + m\right)\psi - e\bar{\psi}\not{A}\psi\,, \tag{3.1.2}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field strength tensor for the vector potential A_μ, ψ is a fermion spinor field (with $\bar{\psi} = i\psi^\dagger\gamma^0$ its adjoint) of bare mass m, with the associated electric charge $e^2 = 4\pi\alpha$,[1] and we have adopted the Feynman *slash* notation $\not{\mathcal{O}} := \gamma^\mu\mathcal{O}_\mu$, with γ^μ the Dirac matrices. A familiar interaction in QED is the *electron self-energy*, arising from an electron interacting with its own EM field, and corresponds to radiative corrections to the electron mass m_e. At the one-loop level, this interaction is represented by the Feynman diagram given in fig. 3.1.1. The quantitative expression extracted from this diagram then corresponds to the electron

[1] Where $\alpha \approx 1/137$ is the fine structure constant at low energies $\lesssim \mathcal{O}(\text{MeV})$ [121].

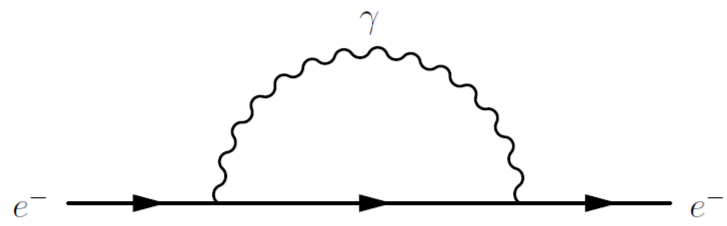


Figure 3.1.1: One-loop electron self-energy.

self-energy (at one-loop) [121]

$$i\Sigma_2(\not{p}) \ = \ \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \frac{(-ie\gamma^\mu)(-i\eta_{\mu\nu})i(\not{k}+m)(-ie\gamma^\nu)}{(k^2-m^2+i\varepsilon)((k-p)^2+i\varepsilon)} \, . \tag{3.1.3}$$

As is usually the case with loop corrections, the self-energy is divergent and so we must regularise eq. (3.1.3) to obtain an analytical solution to the integral. Upon using dimensional regularisation (cf. ref. [121] for details; we shall also carry out an explicit calculation in §3.1.1), one can isolate the divergent pieces of $\Sigma_2(\not{p})$. To avoid overcomplicating the discussion, we shall briefly comment that, in order to make sense of the theory and obtain physical predictions, we interpret the *bare* mass of the electron m as being formally infinite. In doing so, we can tame the infinities that arise in radiative corrections such as the self-energy contribution above. This is achieved through the introduction of counterterms, δ_m and δ_2 such that what we end up with are a renormalised mass m_{R}, defined via $m = (1+\delta_m)m_{\mathrm{R}}$, and self energy $\Sigma_{\mathrm{R}}(\not{p}) = \Sigma_2(\not{p}) + \delta_2\not{p} - (\delta_m + \delta_2)\, m_{\mathrm{R}} + \mathcal{O}(\alpha^2)$. These renormalised quantities are themselves finite, and can then be related to the *physical*, i.e. experimentally measured, (pole) mass m_e of the electron as $m_e = m_{\mathrm{R}} + \Sigma_{\mathrm{R}}(m_e)$. In this case, using the *modified minimal subtraction* ($\overline{\mathrm{MS}}$) renormalisation scheme, the result (to leading order in α) is [121],

$$m_e \ = \ m_{\overline{\mathrm{MS}}} + m_e \frac{\alpha}{4\pi}\left[5 + 3\ln\left(\frac{\mu^2}{m_e^2}\right)\right] + \mathcal{O}(\alpha^2)\, , \tag{3.1.4}$$

where $m_{\overline{\mathrm{MS}}}$ is the renormalised (so-called $\overline{\mathrm{MS}}$) mass of the electron, and μ is the renormalisation energy scale (this is the energy scale, the so-called *subtraction point*, where the theory is renormalised). Note that self-energy loop, fig. 3.1.1, provides a correction $\Delta m \ := \ m_e - m_{\overline{\mathrm{MS}}} = \frac{\alpha}{4\pi} m_e \left[5 + 3\mathrm{ln}\left(\frac{\mu^2}{m_e^2}\right)\right]$ to the $\overline{\mathrm{MS}}$ mass $m_{\overline{\mathrm{MS}}}$ relative to the physical mass m_e, however, it is evident that $\Delta m \lesssim \mathcal{O}(m_e)$.[2] It turns out that the correction to $m_{\overline{\mathrm{MS}}}$ at each order in perturbation theory, is proportional to m_e itself, i.e., it recieves *multiplicative* corrections. Consequently, the theoretically predicted mass of the electron is only very mildly sensitive to ultraviolet physics, as its corrections remain small as we increase the energy scale μ, such that its value never strays far from m_e. As such, we only ever need to fine-tune $m_{\overline{\mathrm{MS}}}$ by (at most) an $\mathcal{O}(1)$ amount at each order in perturbation theory in order to match the observed physical mass m_e. This property will be of importance when we come to discuss radiative instability in the next section, and we shall discuss it in further detail then.

Now that we have some appreciation for how vacuum fluctuations contribute to theoretical predictions of physical quantities, we shall now address whether or not such effects are experimentally observable. It appears to be the case that they do, as evidenced by the Lamb shift [122] and the Casimir effect [123]. Indeed, Willis Lamb observed a small splitting between two energy levels of the Hydrogen atom which are predicted to be degenerate at tree level in QED. This splitting can be accounted for theoretically by including one loop corrections, the dominant contributions of which are from the vacuum polarisation and the electron self-energy [fig. 3.1.1] [121]. The prediction agrees well with experiment [124–126], serving to highlight the importance of these contributions to physical processes, and leading to the development of modern renormalisation techniques. Moreover, the Casimir effect arises from considering the effects of the QED vacuum on two neutral, conducting plates (separated by a finite distance). Indeed, it is found that the relative difference between the vacuum energy of the EM field, inside and outside the two plates, generates a force between them [121, 123].

These phenomena seem to indicate that loop corrections do have observable effects on physical processes, however, we should note that there is a caveat here. The theoretical calculations pertaining to the Lamb shift and Casimir effect both involve external on-shell particle states, and thus do not correspond to pure vacuum effects, however, the fact that the contribution of vacuum fluctuations to these processes play an important role in the calculation of observables suggests that we should take them

[2]Note that, since the physical mass m_e must be independent of the renormalisation energy scale μ we choose, this implies that the $\overline{\mathrm{MS}}$ mass $m_{\overline{\mathrm{MS}}}$, and the fine-structure constant α are dependent on μ, and thus their values "run" with the energy scale. This leads to the concept of the renormalisation group, which provides a systematic structure for quantifying how coupling constants run as the renormalisation scale μ changes. See e.g. ref. [121] for a detailed discussion of the renormalisation group.

seriously when considering physical phenomena. With this in mind, we now consider what happens when we introduce gravity into the mix. Within the context of GR, we know that, by the equivalence principle, all forms of matter and energy gravitate, and do so with the same strength. As such, one expects that the contributions from vacuum fluctuations to the Lamb shift and Casimir effect will gravitate. Indeed, if we consider the Lamb shift, for example, it is found that the relevant loop corrections affect both the inertial and gravitational energy of an atom [53]. In particular, for heavy nuclei such as Aluminium and Platinium, it has been shown that the loop corrections arising from vacuum polarisation, to both their inertial energies, are of order 10^{-3}, but differ by a factor of 3. Nevertheless, the respective ratios between their gravitational and inertial energies remain the same up to order 10^{-12} [127].

We see then, that there is experimental evidence to suggest that the presence of vacuum fluctuations affect physical processes, and furthermore, at least in certain situations, gravitate. Given this, one is motivated to consider the effects that arise due to their presence when gravity is "switched on". Processes such as the example discussed above do not contribute to the cosmological constant, since they include SM particles on external legs. In vacuum, however, the diagams remaining are so-called bubble diagrams which do not contain external legs. Such bubble diagrams contribute to the vacuum energy of a given theory. With this in mind, in "switching on" gravity, and applying the principle of equivalence, one would expect these bubble diagrams to couple to external graviton legs, i.e. that the vacuum energy couples to gravity (just as all other forms of energy and matter do). This subject forms the basis of the next subsection, where we shall see that the CCP manifests itself as a consequence of taking the existence of vacuum energy seriously in the context of GR. For further details on the nature of the CCP, we refer the reader to the following reviews [53, 128–131].

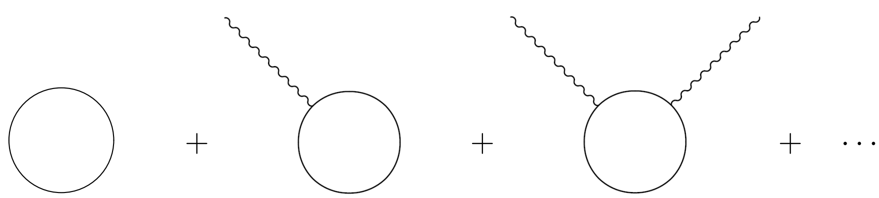
3.1.1 The radiative instability of Λ

The core of the issue manifests when one considers radiative corrections to the vacuum energy contributions from each massive particle species.[3] To highlight the problem we shall consider a scalar field ϕ of mass m, with a quartic self-interaction $\sim \lambda\phi^4$,

[3]Whilst it is true that all vacuum loop corrections for strictly massless particles are zero (since they are scale-free), as we are working in the framework of EFT, there is a caveat here, as we have integrated-out more massive particles, which could enter into loop corrections of naively massless Feynman diagrams. We shall neglect this subtlety here as a big problem already exists by considering explicitly massive particles.

that is minimally coupled to gravity. The results that we shall arrive at are qualitatively the same for other particles of different mass and spin, and so we therefore stick to a simple example to avoid over-complicating the analysis. In the following discussion we make use of Refs. [53, 130], and refer the reader to them for further details.

As a quantum theory of gravity currently eludes us, in order to calculate the vacuum energy contributions, we must adopt a semi-classical approximation in which we treat the gravitational sector classically and the matter sector quantum mechanically (in the framework of QFT). Although this means that we are ignoring all contributions involving virtual graviton exchanges, since they are sensitive to the quantum effects of gravity and as such not well understood presently, the loop contributions from the matter sector already present a significant problem in trying to match theoretical calculations of vacuum energy with the observed value of the cosmological constant.[4] Practically, this corresponds to perturbing the metric around a Minkowski background, $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, and in doing so fixes the graviton-scalar interaction vertex. The revelant Feynman diagrams are then given by a perturbative sum of all possible vacuum scalar loops coupled to external graviton legs. For example, at the one-loop level, we have,

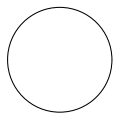


(3.1.5)

where the first diagram corresponds to the net sum of vacuum energy in the absence of gravity, and the second two diagrams correspond to the tadpole two-graviton contributions at one-loop order. The ellipsis denote all further one-loop contributions containing arbitary numbers of external graviton legs. Due to the diffeomorphism invariance of GR, it must be that the series of diagrams (3.1.5) resums, such that the net result is of the form $-\rho_{\text{vac}}^{\phi,\text{1-loop}} \int \mathrm{d}^4x\sqrt{-g}$. Consequently, since one expects $\rho_{\text{vac}}^{\phi,\text{1-loop}}$ to factor out (if it did not diffeomorphism invariance would be violated), we need only calculate one of the loop diagrams to obtain an estimate for the vacuum

[4] In a full solution to the CCP these would of course need to be dealt with.

energy contribution at one-loop. To this end, we consider the simplest vacuum loop,



$$\sim \frac{i}{2}\,\mathrm{Tr}\left[\ln\left(-\frac{\delta^2 S}{\delta\phi(x)\phi(y)}\right)\right] = \frac{i}{2}\mathrm{Tr}\left[\ln\Big(\big(-\Box_x + m^2\big)\delta^{(4)}(x-y)\Big)\right]$$

$$= \frac{i}{2}\int \mathrm{d}^4x\,\mathrm{d}^4y\,\delta^{(4)}(x-y)\int\frac{\mathrm{d}^4k}{(2\pi)^4}\,e^{ik\cdot(x-y)}\ln\left(-k^2+m^2\right)$$

$$= -\frac{1}{2}\int \mathrm{d}^4x\int\frac{\mathrm{d}^4k_{\scriptscriptstyle E}}{(2\pi)^4}\,\ln\left(k_{\scriptscriptstyle E}^2+m^2\right) \sim -\rho_{\mathrm{vac}}^{\phi,\text{1-loop}}\int d^4x\,. \tag{3.1.6}$$

Here we have diagonalised $\big(-\Box_x + m^2\big)\delta^{(4)}(x-y)$ by representing it in terms of a momentum basis, such that we can straighforwardly take its log and trace over this result. With the aim of computing this integral, we subsequently analytically continued the integral to Euclidean space via a Wick rotation $k^0 \to ik_E^0$. To obtain an expression for the momentum space integral, we shall make use of dimensional regularisation, analytically continuing the spacetime dimensions $D = 4 \to D = 4-\varepsilon$ (where $\varepsilon \ll 1$) to obtain a finite result,

$$\int\frac{\mathrm{d}^4k_{\scriptscriptstyle E}}{(2\pi)^4}\ln\left(k_{\scriptscriptstyle E}^2+m^2\right) = -\lim_{\varepsilon\to0}\frac{\partial}{\partial\alpha}\left(\mu^{4-D}\int\frac{d^Dk}{(2\pi)^D}\left(k^2+m^2\right)^{-\alpha}\right)$$

$$= -\lim_{\varepsilon\to0}\frac{m^4}{16\pi^2}\left(\frac{m^2}{4\pi\mu^2}\right)^{-\varepsilon/2}\Gamma\left(\frac{\varepsilon}{2}-2\right) = -\lim_{\varepsilon\to0}\frac{m^4}{32\pi^2}\left[\frac{2}{\varepsilon}-\frac{3}{2}+\ln\left(\frac{\tilde{\mu}^2}{m^2}\right)+\mathcal{O}(\varepsilon)\right] \tag{3.1.7}$$

where $\Gamma(z) = \int_0^\infty \mathrm{d}x\; x^{z-1}e^{-x}$ is the gamma function, and μ is an arbitrary (but regulator dependent) mass scale, which we must introduce by dimensional analysis as a consequence of our choice of regularisation. Indeed, the arbitrariness of μ allows us to absorb factors of 4π and the Euler-Mascheroni constant, $\gamma_{\scriptscriptstyle E}$, into a redefinition of μ, i.e. $\tilde{\mu}^2 = 4\pi e^{-\gamma_E}\mu^2$. We have kept the divergent piece and the leading order finite pieces, neglecting terms of order ε and above, since they will vanish in the limit as $D \to 4$ ($\varepsilon \to 0$). The final result is given by,

$$\rho_{\mathrm{vac}}^{\phi,\text{1-loop}} \sim -\frac{m^4}{(8\pi)^2}\left[\frac{2}{\varepsilon}+\ln\left(\frac{\tilde{\mu}^2}{m^2}\right)-\frac{3}{2}\right]. \tag{3.1.8}$$

where the limit ($\varepsilon \to 0$) is implicit. In order to obtain a sensible finite result in the limit $\varepsilon \to 0$, we must remove the divergence from this loop contribution using renormalisation. To do so, we recall from §1.2, that we are always free to add a bare cosmological constant, Λ_{bare}, to the EH action [eq. (1.2.1)]. This itself must contain a formally infinite piece, acting as a counter-term, to cancel off the divergence arising from the vacuum loop. Thus, if we separate Λ_{bare} up into a finite piece Λ_{C} and a formally divergent piece $\delta\Lambda_{\text{C.T.}}$, such that $\Lambda_{\text{bare}} = \Lambda_{\text{C}} + \delta\Lambda_{\text{C.T.}}$, then it is evident that we must choose (using a slightly modified $\overline{\text{MS}}$ scheme),

$$\delta\Lambda^{\phi,\text{1-loop}}_{\text{C.T.}} \;=\; \frac{m^4}{(8\pi)^2}\left[\frac{2}{\varepsilon} - \gamma_{_E} + \ln(4\pi) - \frac{3}{2}\right] . \tag{3.1.9}$$

Thus, what actually gravitates is the finite combination $\Lambda^{\phi,\text{1-loop}} = \Lambda^{\phi,\text{1-loop}}_{\text{bare}} + \rho^{\phi,\text{1-loop}}_{\text{vac, R}}$ (where $\rho^{\phi,\text{1-loop}}_{\text{vac, R}}$ is the renormalised vacuum energy at one loop), i.e. the renormalised one-loop vacuum energy contribution from the massive scalar field ϕ,

$$\Lambda^{\phi,\text{1-loop}} \;=\; \Lambda^{\phi,\text{1-loop}}_{\text{bare}} + \rho^{\phi,\text{1-loop}}_{\text{vac, R}} \;\sim\; \Lambda_{\text{C}} + \frac{m^4}{(8\pi)^2}\ln\left(\frac{m^2}{\mu^2}\right) . \tag{3.1.10}$$

Note that, in the $\overline{\text{MS}}$ scheme, one implicitly sets the arbitrary mass scale μ to the subtraction point at which we renormalise. Thus μ is interpreted as the renormalisation energy scale (itself still arbitrary, but *finite*). It is clear from this calculation that we have no way of predicting the magnitude of contributions to the vacuum energy from massive particles in the context of SM QFT. This is due to the fact that we are describing physics in the context of an EFT and thus after renormalising, the result always depends on an arbitrary mass scale, μ.

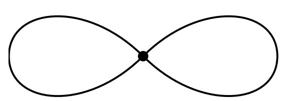
In practice we must experimentally measure its value and then adjust Λ_{C} accordingly, such that the theory matches observational data. For example, suppose that our massive scalar field is the SM Higgs boson, whose mass is $m = 125$ GeV [132]. Then at the one-loop level, the finite contributions to the renormalised vacuum energy must cancel to an accuracy of 1 part in 10^{57} in order to agree with the experimental upper bound of $\sim (\text{meV})^4$ [51] for the net cosmological constant.[5] This is already a considerable amount of fine-tuning which in itself presents a problem, why does such a precise cancellation seem to be occurring in nature?

The issue is further exacerbated when one takes into account all possible contributions to the cosmological constant. Indeed, the SM is valid at least up to the TeV

[5] Here, we have assumed that $\ln\left(\frac{m^2}{\mu^2}\right) \sim \mathcal{O}(1)$.

scale, and so one expects that Λ should receive contributions from heavy particles right up to this energy scale, requiring a fine-tuning of 1 part in 10^{60}, in order for theory to agree with experiment. Moreover, the SM is potentially valid up to the Planck scale [133], meaning that it is possible that Λ could receive contributions from radiative corrections right up to $M_{\text{Pl}} \sim 10^{18}$ GeV. The result is that there could be a difference of up to 120 orders of magnitude between the observed and theoretically predicted values for the cosmological constant. In this case, a fine-tuning of 1 part in 10^{120} is required for theory to match experiment. We could perhaps tolerate this extreme amount of fine-tuning to some extent by accepting an anthropic argument: that the vacuum we exist in is simply the one that is able to support evolution of intelligent life, and so such a fine-tuned value for the cosmological constant is, in some sense, necessary.[6] This does seem like a rather contrived solution, however.

Although the fine-tuning in itself is an issue to a certain extent, there is a much more severe problem that is really at the heart of the CCP, arising from the seemingly innocuous question: what happens when we consider higher-order loop corrections to the vacuum energy density from our massive scalar field? At two loops, the vacuum energy density ρ^{ϕ}_{vac} receives an additive correction from the following diagram



$$\begin{aligned} &\sim \frac{\lambda}{8}\int \mathrm{d}^4x \int \frac{\mathrm{d}^4 p_{_E}}{(2\pi)^4}\frac{\mathrm{d}^4 k_{_E}}{(2\pi)^4}\frac{1}{p_{_E}^2+m^2}\frac{1}{k_{_E}^2+m^2} \\ &\sim -\rho^{\phi,\text{2-loop}}_{\text{vac}}\int \mathrm{d}^4x \end{aligned} \tag{3.1.11}$$

After regularising and renormalising, as we did for the one-loop correction, we find that eq. (3.1.11) contributes a term that scales as $\rho^{\phi,\text{2-loop}}_{\text{vac}} \sim \lambda m^4$. Importantly, we see that this is an *additive* correction (unlike in the electron mass case, where they were multiplicative) and is proportional to the fourth power of the mass. As we did previously, assuming that scalar field is the Higgs boson, we find that even if we take the coupling constant $\lambda \sim 0.1$, the two-loop diagram still contributes a large correction to the cosmological constant relative to its observed value of $\sim (\text{meV})^4$. Now,

[6]By this, it is meant that if the cosmological constant were postive and large, i.e. $\Lambda \gg (\text{meV})^4$, then the early universe would have expanded so rapidly that it would not have been possible for matter to coalesce, forming stars and planets necessary for life. Similarly, if the cosmological constant were negative, and its magnitude large, i.e. $|\Lambda| \gg (\text{meV})^4$, then the early universe would have rapidly collapsed, again preventing life from taking a foothold.

the problem is that we have already fixed the value of Λ_C, such that Λ matches the experimental data at the one-loop level, and so in order to cancel off the additional contributions at the two-loop order, we must *re-tune* Λ_C. This is not necessarily a problem in itself; an artifact of perturbation theory is that we must re-tune the counter-terms at each order to match experiment. The issue here though, is that at each order in the loop expansion, the correction to the vacuum energy from the massive scalar is not significantly suppressed relative to either the lower order contributions or the observed value of Λ. Consequently, we are forced to re-tune Λ_C to almost the same level of accuracy, order-by-order in perturbation theory, such that our theoretical prediction matches experimental data.

The need for such a severe amount of re-tuning at each loop order indicates that the vacuum energy density is *radiatively unstable*. This instability to higher order loop corrections is the real reason as to why the CCP is so troubling. It indicates that our low-energy description of the cosmological constant is highly sensitive to physics in the UV, of which we are ignorant of at present. An alternative way of seeing this problem, is to simply focus on the one-loop correction [eq. (3.1.6)]. Now, consider a shift in the mass of the particle in the loop, $m \to m + \Delta m$, or keep everything fixed, and instead introduce a new heavy particle into the theory. In either case, it is clear that Λ has a power-law sensitivity to such a modification. We see then, that the cosmological constant is unstable against changes in our effective description of physics, either through introducing new energy scales, or by phase transitions in the early universe that give rise to constant shifts in the vacuum energy.

The fact that the observed value of the cosmological constant is so small signals a naturalness problem, i.e. why is its *natural* value at the meV scale and not at the cut-off scale of our EFT. This is analogous to the so-called "*Hierarchy Problem*" relating to the mass of the Higgs boson in the SM.[7] Note that this not the case for other SM parameters, such as the mass of the electron, whose value changes by an amount proportional to its mass at each loop order. For example, recall from §3.1 (cf. eq. (3.1.4)), that at the one-loop level, $\Delta m = \frac{\alpha}{4\pi} m_e \left[5 + 3\ln\left(\frac{\mu^2}{m_e^2}\right)\right]$, at leading order in α [121] (where $m_e \simeq 0.511\,\mathrm{MeV}$ is the pole mass of the electron). Hence its small value is radiatively stable under loop corrections, i.e., there is no need for significant fine tuning at each loop order to keep the electron mass small, since radiative

[7] See ref. [134] for a review of the Hierarchy Problem in the SM and some possible solutions to it.

corrections to its value satisfy $\Delta m \lesssim \mathcal{O}(m_e)$. This is a consequence of the fact that the mass of the electron is protected by the chiral symmetry possessed by massless fermions. Since in the massless limit there should be no loop corrections (otherwise their presence would generate a mass, thus violating the chiral symmetry), they must all be proportional to the symmetry breaking parameter, i.e., its mass. Any parameter whose small value arises from a softly broken symmetry is said to be *technically natural*.[8] The cosmological constant and the Higgs boson have no such 'custodial' symmetry and thus their small values are not protected from radiative corrections. As such the SM is unable to provide a technically natural explanation for the smallness of both of their observed values. This sensitivity of the cosmological constant to UV physics can also be understood in terms of Wilsonian effective actions [136, 137], and for a detailed discussion of the problem using this approach we refer the reader to ref. [131].

As the CCP arises from considering field theoretic effects, i.e. radiative (loop) corrections from massive particles in the matter sector, an immediate response is to search for a solution by refining one's quantum description of matter in such a way that the large loop corrections are tamed in some way. Indeed, it is already the case within the framework of the SM that radiative corrections from fermions and bosons have opposite signs, resulting in a partial cancellation of their contributions to the vacuum energy density. This naturally lends itself to the idea of supersymmetry, in which each fermion species has a partnering boson species with the same mass and interaction strengths (see e.g. ref. [138] for an introduction to supersymmetry). Thus, in the case where supersymmetry is exact, all the radiative corrections from each massive particle species cancel out, such that the net cosmological constant is zero. However, as nice an idea as this is, supersymmetry is *not* exact - the symmetry is broken at least up to the TeV scale [139], and accordingly this sets a natural scale for the value of the cosmological constant, i.e. $\Lambda \sim (\text{TeV})^4$. Clearly this is still much larger than the observed value of $\sim (\text{meV})^4$, and leaves us not much better off than we were before.

Given that the field theory sector is currently unable to offer much in terms of a solution to the CCP, this leads us to search for solutions in the gravitational sector. A popular approach has been to modify our description of gravity in such a way

[8] *Technical naturalness* is a criterion, originally introduced by t'Hooft, which states that a parameter of a theory is naturally small in value if setting it to zero enhances the symmetry of the theory [135]

as to prevent the vacuum energy density from sourcing curvature. In this way, we do not address the radiative instability of the vacuum energy, but we do provide a solution that allows for a small value for the cosmological constant, thus alleviating the issue, at least for the gravitational sector. As we shall see in the following two subsections, we cannot just make any modification that we please, there are some fairly strict constraints that one must satisfy in order to construct a physically viable (and theoretically "healthy") theory, particularly given the empirical success of GR at the scale of the solar system.

3.2 Solving the CCP: the pitfalls

3.2.1 Weinberg's no-go theorem

Perhaps the most minimal approach to resolving the CCP is to posit that there exists new local degrees of freedom in the matter sector, in the simplest case a collection of scalar fields ϕ_i, coupling minimally to gravity. We demand that their coupling to gravity is such that they can self-adjust, or "*self-tune*", to absorb the net vacuum energy and consequently prevent it from sourcing gravity. Of course, this would rule out vacuum energy as a source of dark energy, but it would offer a dynamical solution to a significant problem plaguing semi-classical gravity. However, this naive approach was shown by Weinberg to be flawed. Indeed, Weinberg's no-go theorem states that it is impossible to have such a self-tuning mechanism without transferring the fine-tuning of the bare cosmological constant to the potential of the self-tuning fields [128]. We shall give a review of this no-go theorem before proceeding (here we follow [53, 63] and refer the reader to them for further details).

Our initial assumption is that the gravitational action is the usual Einstein-Hilbert term with the addition of a term containing all contributions from the net cosmological constant, constructed from the spacetime metric $g_{\mu\nu}$ and a set of scalar fields $\{\phi_i\}_{i=1,\ldots,N}$, whose job it is to self-tune such that they absorb the vacuum energy contributions to the cosmological constant.[9] This theory is then described by the following action,

$$S \;=\; \int \mathrm{d}^4x\sqrt{-g}\left[\frac{M_{\rm Pl}^2}{2}R + \mathcal{L}(g_{\mu\nu},\phi_i)\right] \;=\; \int \mathrm{d}^4x\,\mathscr{L}(g_{\mu\nu},\phi_i)\,, \tag{3.2.1}$$

[9] Although Weinberg's argument is valid for fields of any tensor rank we stick to scalar fields here for simplicity as the main points of the argument are still present

where $\mathcal{L}$ in principle contains all possible interactions between the scalar fields and the metric. Assuming that a solution exists that corresponds to a Poincaré invariant vacuum geometry $g_{\mu\nu} = \eta_{\mu\nu}$, and field configuration $\phi_i = \bar{\phi}_i$ = constant, the corresponding equations of motion are

$$\frac{1}{\sqrt{-g}}\frac{\delta S}{\delta g^{\mu\nu}}\bigg|_{\substack{g^{\mu\nu} = \eta^{\mu\nu}, \\ \phi_i = \bar{\phi}_i}} = \frac{\partial \mathscr{L}}{\partial g^{\mu\nu}} = 0 \tag{3.2.2a}$$

$$\frac{1}{\sqrt{-g}}\frac{\delta S}{\delta \phi_i}\bigg|_{\substack{g^{\mu\nu} = \eta^{\mu\nu} \\ \phi_i = \bar{\phi}_i}} = \frac{\partial \mathscr{L}}{\partial \phi_i} = 0\,. \tag{3.2.2b}$$

If these equations are satisfied independently, then fine-tuning of the solutions is inevitable [128][10], and so we are left with the alternative: that Eqs. (3.2.2a) and (3.2.2b) do not hold independently. In this case, we can in principle avoid fine-tuned solutions by insisting that the trace of the gravitational EOM be automatically fulfilled as a consequence of the scalar equations. This corresponds to requiring the trace of the metric equation of motion to be a linear combination of the ϕ_i field equations, i.e.,

$$g^{\mu\nu}\frac{\partial \mathscr{L}}{\partial g^{\mu\nu}} = \sum_i f^i(\phi)\frac{\partial \mathscr{L}}{\partial \phi_i} \tag{3.2.3}$$

where $f^i(\phi)$ are generic functions of the scalar fields ϕ_i. By imposing this condition, we are implicitly demanding that the action possess a particular symmetry. Indeed, consider the variation of the action [eq. (3.2.1)] with respect to $g^{\mu\nu}$ and ϕ_i,

$$\delta S = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{1}{\sqrt{-g}}\frac{\delta \mathscr{L}}{\delta g^{\mu\nu}}\delta g^{\mu\nu} + \sum_i \frac{1}{\sqrt{-g}}\frac{\delta \mathscr{L}}{\delta \phi_i}\delta\phi_i\right]. \tag{3.2.4}$$

Given this, we observe that by choosing the infinitesimal variations,

$$\delta g^{\mu\nu} = 2\epsilon g^{\mu\nu} \qquad \delta\phi_i = \epsilon f^i(\phi)\,, \tag{3.2.5}$$

then eq. (3.2.3) implies that this variation is a symmetry of the action, when one takes the fields to be in their vacuum configuration, $\phi_i = \bar{\phi}_i$ = const. Following from this, we see that if we start out with an action invariant under the variations

[10] This is a consequence of the residual GL(4) symmetry remaining after fixing the vacuum solution. If one assumes that $\frac{\partial \mathscr{L}}{\partial \phi_i} = 0$ holds (without, for the time being, assuming $\frac{\partial \mathscr{L}}{\partial g^{\mu\nu}} = 0$), then using the GL(4) symmetry, one can solve this equation to give $\mathscr{L} = \sqrt{g}V(\phi_i)$. The remaining field equation $\frac{\partial \mathscr{L}}{\partial g^{\mu\nu}} = 0$ then yields $V(\phi_i)$, which corresponds to fine-tuning.

given by eq. (3.2.5), and admitting a vacuum solution $\phi_i = \bar{\phi}_i = \text{const.}$ such that $\frac{\partial \mathscr{L}}{\partial \phi_i}\Big|_{\bar{\phi}_i} = 0$, then eq. (3.2.3) will be automatically fulfilled, and the existence of a Minkowski solution is guaranteed. However, it turns out that this is impossible to achieve without having to fine-tune the scalar field potential.

Let us redefine the fields $\phi_i \to \psi_i$ such that the variations given by eq. (3.2.5) become,

$$\delta g^{\mu\nu} \;=\; 2\epsilon g^{\mu\nu}\,, \quad \delta\psi_{i\neq N} \;=\; 0\;(i \;=\; 1,\ldots,N-1), \quad \delta\psi_N \;=\; -\epsilon\,, \tag{3.2.6}$$

it can be seen that this corresponds to a conformal rescaling of the metric. Indeed, consider a conformal transformation of the metric $g_{\mu\nu} \to \tilde{g}_{\mu\nu} = e^{2\psi} g_{\mu\nu}$, with the corresponding transformation of the inverse metric given by $g^{\mu\nu} \to \tilde{g}^{\mu\nu} = e^{-2\psi} g^{\mu\nu}$. Now if $\delta\psi$ is taken to be *constant* and the transformation is infinitesimal, then we have that $\tilde{g}^{\mu\nu} = g^{\mu\nu} - 2\psi g^{\mu\nu}$, i.e. $\delta g^{\mu\nu} = -2\psi g^{\mu\nu}$. Hence, we see that the variation given by eq. (3.2.6) conformally rescales the metric. Consequently, the action will be invariant if it is constructed from the conformal metric $\tilde{g}_{\mu\nu} = e^{2\psi_N} g_{\mu\nu}$ and the scalar fields $\psi_{i\neq N}$. When all the fields are constant, the curvature invariants of this metric will vanish, and so diffeomorphism invariance then requires that the on-shell action has the form,

$$S \;=\; \int \mathrm{d}^4x\sqrt{-\tilde{g}}\,V(\psi_{i\neq N}) \;=\; \int \mathrm{d}^4x\sqrt{-g}\,e^{4\psi_N}V(\psi_{i\neq N})\,. \tag{3.2.7}$$

Accordingly the gravitational equation of motion becomes,

$$\frac{1}{\sqrt{-g}}\frac{\delta S}{\delta g^{\mu\nu}}\Bigg|_{\substack{g^{\mu\nu}\,=\,\eta^{\mu\nu}\\ \psi_{i\neq N}\,=\,\text{const.}}} \;=\; -\frac{1}{2}e^{4\psi_N}V(\psi_{i\neq N}) \;=\; 0 \quad\Longrightarrow\quad e^{4\psi_N}V(\psi_{i\neq N}) \;=\; 0\,. \tag{3.2.8}$$

Thus we are left we two choices: either we take $V(\psi_{i\neq N}) = 0$, or take the limit $e^{4\psi_N} \to 0$. The former case clearly corresponds to fine-tuning, since it would require us to tune the potential of $\psi_{i\neq N}$ such that it has a minimum at $V(\psi_{i\neq N})\big|_{\psi_{i\neq N}} = 0$. In the latter case, since the matter fields couple to $\tilde{g}_{\mu\nu}$, it follows that the physical masses in the matter sector would all scale as $e^{2\psi_N}$, and so taking the limit $e^{4\psi_N} \to 0$ corresponds to scale invariance in which all the physical masses vanish, a result that is clearly ruled out by observations.

We see then that Weinberg's no-go theorem presents a very large obstacle for any theory that hopes to employ such a self-tuning mechanism and remain physical. One might be left wondering whether this is even at all possible and perhaps we should abandon such an approach completely. However, if one carefully analyses the assumptions made in the no-go theorem it becomes apparent that there are options remaining to avoid it. For example, a simple, but elegant solution is to relax the assumption of on-shell Poincaré invariance at the level of the fields, instead simply requiring them to be spatially homogeneous. We shall explore this option in more detail in §4.

3.2.2 Ostrogradsky's theorem

A natural starting point in an attempt to cure the CCP is to construct a modified theory of gravity, i.e. a theory beyond GR. The reason for doing so is that GR in itself does not provide any wiggle room to evade the CCP, in particular, it is not possible to realise a Minkowski vacuum solution to Einstein's field equations without fine-tuning the net cosmological constant to zero. However, one has to be very careful in how such a modified theory of gravity is constructed, indeed there is good reason why Lagrangians tend to be dependent on (at most) first-order in time derivatives of the physical degrees of freedom. This reason is due to Ostrogradsky's theorem which states that there is a linear instability in the Hamiltonians associated with Lagrangians that depend non-degenerately on higher than first-order time derivatives [140]. In fact this result is so general that in order to review it one can consider the problem in the context of classical mechanics for simplicity (here we have made use of [141] and refer the reader to it for a more extensive discussion).

To highlight Ostrogradsky's result, we shall first briefly recapitulate on the construction of the Hamiltonian in the standard case where the Lagrangian depends on at most first-order time derivatives, then we will examine Ostrogradsky's construction of the problem for Lagrangians containing second-order time derivatives. Let us consider a one-dimensional system containing a single point-particle whose time-dependent position is given by $q(t)$. The system is described by a Lagrangian $L(q, \dot{q})$ that we assume depends non-degenerately on $\dot{q}$, i.e. $\det\left(\frac{\partial^2 L}{\partial \dot{q}^2}\right) \neq 0$. Given this the equation of motion is given by the usual Euler-Lagrange (EL) equation,

$$\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right) = 0\,, \tag{3.2.9}$$

and since $L(q,\dot{q})$ is assumed to be non-degenerate it is clear that the solutions to eq. (3.2.9) require two pieces of initial data: $q_0 = q(t_0)$ and $\dot{q}_0 = \dot{q}(t_0)$. Correspondingly, this means that we require two canonical coordinates, q and p in the phase space of our system, where we define the canonically conjugate momentum as $p = \frac{\partial L}{\partial \dot{q}}$. Since L is a convex function of $\dot{q}$, this is a one-to-one mapping and so we can (in principle) invert this transformation to find a function $v(q,p)$ such that,

$$\left.\frac{\partial L}{\partial \dot{q}}\right|_{\dot{q}\,=\,v(q,p)} \;=\; p\,. \tag{3.2.10}$$

The Hamiltonian corresponding to this system is then constructed by taking the Legendre transform of L with respect to $\dot{q}$,

$$H(q,p) \;=\; p\dot{q} - L(q,\dot{q}) \;=\; pv(q,p) - L(q, v(q,p))\,, \tag{3.2.11}$$

and, if it is not explicitly time-dependent (i.e. $\frac{\partial H}{\partial t} = 0$), H is a conserved quantity corresponding to the energy of the system. Furthermore, the Hamiltonian generates time evolution of the systems phase space coordinates $q(t)$ and $p(t)$. Indeed, their evolution is dictated by the following equations,

$$\dot{q} \;\coloneqq\; \frac{dq}{dt} \;=\; \{q, H\} \;=\; \frac{\partial H}{\partial p} \;=\; v + p\frac{\partial v}{\partial p} - \left.\frac{\partial v}{\partial p}\frac{\partial L}{\partial \dot{q}}\right|_{\dot{q}\,=\,v} \;=\; p\,, \tag{3.2.12a}$$

$$\dot{p} \;\coloneqq\; \frac{dp}{dt} \;=\; \{p, H\} \;=\; -\frac{\partial H}{\partial q} \;=\; -p\frac{\partial v}{\partial q} + \left.\frac{\partial v}{\partial q}\frac{\partial L}{\partial \dot{q}}\right|_{\dot{q}\,=\,v} + \frac{\partial L}{\partial q} \;=\; \frac{\partial L}{\partial q}\,, \tag{3.2.12b}$$

where $\{\,\cdot\,, H\} \;\coloneqq\; \frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p}$ is the Poisson bracket. We see that the evolution equations (3.2.12) are consistent with the EL equation (3.2.9) and the inverse phase space map [eq. (3.2.10)]. Importantly, it is clear that $H(q,p)$ is *stable*, by which we mean that it has quadratic dependence on the canonical momentum p and therefore is bounded from below such that the system has a well-defined minimal energy configuration, a so-called classical vacuum state. Furthermore, due to the existence of this vacuum state, the energy of a particle is strictly *positive* (or zero) and so its trajectories are finite. Of course, this is a desirable feature for any theory hoping to describe a physical system as it prevents the existence of so-called "ghost" instabilities in which the physical degrees of freedom can have negative kinetic energies. In this case the Hamiltonian becomes unbounded from below and consequently the physical trajectories of the particle can become singular in a finite amount of time (i.e. $q(t)$ and its derivatives become singular) [142]. At the quantum level this means that, if the system is interacting, the existence of negative energy states signals an instability, since the system can continue to lower its energy interminably by producing a cascade of negative energy states. This causes the Hamiltonian to become unbounded from below since such "ghost" states can have arbitrarily high negative energy, meaning that there is no stable vacuum state.[11]

Let us now consider the same system, but this time with a non-degenerate dependence on $\dot{q}$ and $\ddot{q}$, thus described by a Lagrangian $L(q,\dot{q},\ddot{q})$. Now the EL equation is given by,

$$\frac{\partial L}{\partial q}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right)+\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(\frac{\partial L}{\partial \ddot{q}}\right) \;=\; 0\,. \tag{3.2.13}$$

As $L(q,\dot{q},\ddot{q})$ depends non-degenerately on $\dot{q}$ and $\ddot{q}$, by which we mean $\det\left(\frac{\partial^2 L}{\partial \dot{q}^2}\right)\neq 0$ and $\det\left(\frac{\partial^2 L}{\partial \ddot{q}^2}\right)\neq 0$, we see that we now need four pieces of initial data, q_0, $\dot{q}_0$, $\ddot{q}_0=\ddot{q}(t_0)$ and $\dddot{q}_0=\dddot{q}(t_0)$, in order to find solutions to eq. (3.2.13). Thus, our phase space for this system must have four independent canonical coordinates, originally chosen by Ostrogradsky to be,

$$q_1 \;=\; q\,, \qquad p_1 \;=\; \frac{\partial L}{\partial \dot{q}}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \ddot{q}}\right)\,, \tag{3.2.14a}$$

$$q_2 \;=\; \dot{q}\,, \qquad p_2 \;=\; \frac{\partial L}{\partial \ddot{q}}\,. \tag{3.2.14b}$$

In principle we can then invert these transformations and find a solution for $\ddot{q}$, and since the Lagrangian depends on the coordinates q, $\dot{q}$ and $\ddot{q}$, we consequently need only three out of the four phase space coordinates to do so, a fact that has significant consequences. Indeed solving for $\ddot{q}$ in terms of q_1, q_2 and p_2 requires the existence of a function $a(q_1,q_2,p_2)$ such that,

$$\left.\frac{\partial L}{\partial \ddot{q}}\right|_{\substack{q\,=\,q_1\,,\\ \dot{q}\,=\,q_2\,,\\ \ddot{q}\,=\,a}} \;=\; p_2\,. \tag{3.2.15}$$

[11] One could of course redefine the inner product such that the energy is always positive, however, this would result in negative norm states which would no longer lend themselves to a probabilistic interpretation.

To obtain Ostrogradsky's Hamiltonian we then Legendre transform $L(q, \dot{q}, \ddot{q})$ with respect to both $\dot{q} = q^{(1)}$ and $\ddot{q} = q^{(2)}$,

$$\begin{aligned} H(q_1, q_2, p_1, p_2) &= \sum_{i=1}^{2} p_i q^{(i)} - L(q, \dot{q}, \ddot{q}) \,, \\ &= p_1 q_2 + p_2 a(q_1, q_2, p_2) - L(q_1, q_2, a(q_1, q_2, p_2)) \,. \end{aligned} \tag{3.2.16}$$

The corresponding equations of motion for the systems phase space coordinates q_i and p_i $(i = 1, 2)$ are,

$$\dot{q}_i = \{q_i, H\} = \frac{\partial H}{\partial p_i} \,, \qquad \dot{p}_i = \{p_i, H\} = -\frac{\partial H}{\partial q_i} \,. \tag{3.2.17}$$

One can check each equation to confirm that Ostrogradsky's Hamiltonian generates time evolution of the system:

$$\dot{q}_1 = \frac{\partial H}{\partial p_1} = q_2 \,, \tag{3.2.18a}$$

$$\dot{p}_1 = -\frac{\partial H}{\partial q_1} = -p_2 \frac{\partial a}{\partial q_1} + \frac{\partial L}{\partial q_1} + \frac{\partial a}{\partial q_1} \frac{\partial L}{\partial \ddot{q}} \bigg|_{\substack{q = q_1 \,, \\ \dot{q} = q_2 \,, \\ \ddot{q} = a}} = \frac{\partial L}{\partial q} \,, \tag{3.2.18b}$$

$$\dot{q}_2 = \frac{\partial H}{\partial p_2} = a + p_2 \frac{\partial a}{\partial p_2} - \frac{\partial a}{\partial p_2} \frac{\partial L}{\partial \ddot{q}} \bigg|_{\substack{q = q_1 \,, \\ \dot{q} = q_2 \,, \\ \ddot{q} = a}} = a \,, \tag{3.2.18c}$$

$$\dot{p}_2 = -\frac{\partial H}{\partial q_2} = -p_1 \frac{\partial a}{\partial q_2} + \frac{\partial L}{\partial q_2} + \frac{\partial a}{\partial q_2} \frac{\partial L}{\partial \ddot{q}} \bigg|_{\substack{q = q_1 \,, \\ \dot{q} = q_2 \,, \\ \ddot{q} = a}} = -p_1 + \frac{\partial L}{\partial \dot{q}} \,. \tag{3.2.18d}$$

Hence we see that Eqs. (3.2.18a) and (3.2.18d) reproduce the phase space transformations $\dot{q} = q_2$ and $p_1 = \frac{\partial L}{\partial \dot{q}} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \ddot{q}} \right)$, respectively. Furthermore, Eqs. (3.2.18c) and (3.2.18b) reproduce the inverse mapping [eq. (3.2.15)] and the EL equation (3.2.13), respectively, and thus confirm that $H(q_1, q_2, p_1, p_2)$ does indeed generate time evolution. Moreover, it can be seen from eq. (3.2.16) that if $H(q_1, q_2, p_1, p_2)$ is not explicity time-dependent then it is conserved under time translations, and therefore it constitutes the energy of the system. Notice from eq. (3.2.16), however, that the Ostrogradsky Hamiltonian has a *linear* dependence on the conjugate momentum p_1. Consequently, any system of this form possesses an instability. Indeed, as discussed earlier, the Hamiltonian is unbounded from below. The situation at the quantum level is more severe; it can be shown that the same Ostrogradsky-type canonical variable carries both positive and negative creation and annihilation operators, the result of which is that, if the system is interacting, then the ground-state can dynamically decay into a collection of positive and negative energy excitations [141]. Perhaps the most severe consequence of this instability is that degrees of freedom with large momenta do not decouple from low energy physics (due to the fact that high energy modes can be excited by also exciting modes with the opposite energy), rendering effective field theory useless.

One might ask whether this instability can be cured by simply adding higher-order time-derivatives to the Lagrangian, however, it can be shown that this simply makes the problem more acute. This result is very general and can be extended to the continuum limit such that it carries over to QFT. All is not lost, however, as the problem arises due to the Lagrangian having a non-degenerate dependence on second-order (and higher) time derivatives of the physical degrees of freedom, resulting in equations of motion that contain higher than second-order time-derivatives. Thus, there is a loop-hole in Ostrogradsky's argument: one can construct a Lagrangian that has higher- (than first) order time-dependence so long as it is degenerate such that it maintains second-order equations of motion. In fact, this loop-hole is readily exploited by GR; the Einstein-Hilbert Lagrangian contains second-order time derivatives of the metric, however, in can be shown that this dependence is degenerate and hence Einstein's field equations remain second-order [143]. This observation is of vital importance in the construction of modified theories of gravity such as those that will be discussed in §4.

3.2.3 Screening mechanisms

One of the more common approaches in attempts to solve the CCP in recent years, is to introduce new degrees of freedom into the gravitational sector, the simplest being an additional scalar field ϕ, of which we shall base our following discussion on (here we follow [63], and refer the reader to it for a more detailed review of screening mechanisms). Since we require these new degrees of freedom to negate the contribution of ρ_{vac} to the net cosmological constant to an accuracy of $\sim H_0^2 M_{\text{Pl}}^2 \sim (\text{meV})^4$, we generically require that ϕ are light scalars, i.e.,

$$m_\phi \lesssim H_0 \,, \tag{3.2.19}$$

where H_0 is the present-day Hubble parameter. This requirement is clear from an EFT point of view, as if they were any heavier then we could simply integrate them out such that their contribution is irrelevant to the low energy dynamics of the theory. Since the CCP is a problem arising in the infrared sector of GR, such a theory would clearly be useless as a solution to the CCP. A generic problematic feature arises in such approaches to modify gravity, due to the requirement that these new light scalars must couple to the the SM fields if they are to alleviate the problem of vacuum loop contributions to the cosmological constant. The result is that a given scalar ϕ mediates an additional *fifth force* between SM particles, besides gravity, electromagnetism and the weak and strong nuclear forces, the range of which is $\sim m_\phi^{-1}$. The requirement that ϕ is light [eq. (3.2.19)] implies that this range is of order the present Hubble radius, and as such the fifth force is mediated at both cosmological distances and within the solar system. However, within the solar system GR is well tested (see e.g. [7]), and there is no evidence to suggest the existence of such fifth forces. Thus, if a modified theory of gravity is to be physically viable there must be some mechanism by which any effects of such fifth forces are screened in high density regions, at least at distance scales of order the size of the solar system.

Fortunately, there do exist theoretical models that propose ways in which one can hide fifth forces, generically termed *screening mechanisms*. These can be classified in a fairly general manner though studying how fields present in a Lagrangian manifest themselves in our classical notions of force and potential. To this end, consider a general theory for a scalar field with a universal conformal coupling to matter,

$$\mathscr{L} \;=\; -\frac{1}{2}\tilde{Z}^{\mu\nu}\left(\phi, \partial\phi, \ldots\right)\partial_\mu\phi\partial_\nu\phi - V(\phi) + \frac{g(\phi)}{M_{\rm Pl}}T\,, \tag{3.2.20}$$

where $\tilde{Z}^{\mu\nu}$ encodes all the possible derivative self-interactions of ϕ, $V(\phi)$ is some scalar potential, $g(\phi)/M_{\rm Pl}$ is the (potentially ϕ dependent) dimensionless coupling constant between ϕ and the matter sector, and $T = T^\mu_\mu$ is the trace of the matter energy-momentum tensor $T_{\mu\nu}$. Let us then consider the scalar field in the presence of a non-relativistic, spherically symmetric, static point source, and expand eq. (3.2.20) around some background solution , i.e. $\phi = \bar{\phi} + \varphi$ and $T = \bar{T} + \delta T$. The Lagrangian [eq. (3.2.20)] at second-order in fluctuations, $\delta\phi = \varphi$ and δT, is then given by,

$$\mathscr{L}_{(2)} \;=\; -\frac{1}{2}\mathscr{Z}^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi - \frac{1}{2}m^2\varphi^2 + \varphi\frac{\tilde{g}}{M_{\rm Pl}}\delta T\,, \tag{3.2.21}$$

where $\mathscr{Z}^{\mu\nu}\left(\bar{\phi}, \partial\bar{\phi}\,\ldots\right)$ encodes the various derivative self-interactions of the scalar field (evaluated on its background configuration $\bar{\phi}$), and we have re-defined the coupling constant, such that $\tilde{g} = g'(\bar{\phi})$ is dimensionless. Furthermore, we have identified the coefficient of φ^2 with the mass, $m^2\left(\bar{\phi}, \partial\bar{\phi}\,\ldots\right)$ of the fluctuation. The corresponding field equations for φ are then schematically given by,

$$Z^{\mu\nu}\left(\bar{\phi}\right)\partial_\mu\partial_\nu\varphi - m^2\left(\bar{\phi}\right)\varphi + \frac{\tilde{g}}{M_{\rm Pl}}\delta T \;=\; 0\,. \tag{3.2.22}$$

Since the source is non-relativistic, we have that $\delta T \to -\rho$, and furthermore, it is possible to diagonalise $Z^{\mu\nu}$ [144], significantly simplifying the EOM. Considering a point-like source, i.e. $\rho = M^4\delta^3(\mathbf{x})$, eq. (3.2.22) reduces to,

$$Z(\bar{\phi})\Big(\ddot{\varphi} - c_s^2(\bar{\phi})\nabla^2\varphi\Big) + m^2(\bar{\phi})\varphi \;=\; -\frac{\tilde{g}}{M_{\rm Pl}}M^4\delta^{(3)}(\mathbf{x})\,, \tag{3.2.23}$$

where we have factored out $Z(\bar{\phi}) \;\coloneqq\; Z^{00}(\bar{\phi})$, and defined an effective speed of sound $c_s^2(\bar{\phi}) \;\coloneqq\; \frac{Z^{ii}(\bar{\phi})}{Z^{00}(\bar{\phi})}$ at which the fluctuations propagate.

For a static field configuration (neglecting the spatial variation of $\bar{\phi}$ over the scales of interest), we can solve eq. (3.2.23) using standard Green function techniques, such that the resulting potential is given by

$$V(r) \;=\; -\frac{\tilde{g}M^4}{c_s^2 Z M_{\rm Pl}^4}\frac{e^{-\frac{m}{\sqrt{Z}c_s}r}}{4\pi r}\,. \tag{3.2.24}$$

We thus see that the classical force arising from this potential is given by

$$\mathbf{F}_\varphi(r) \;=\; -\nabla V(r) \;=\; -\frac{\tilde{g}M^4}{c_s^2 Z M_{\rm Pl}^4}\frac{1}{4\pi r^2}\left(1 + \frac{mr}{c_s\sqrt{Z}}\right)e^{-\frac{m}{c_s\sqrt{Z}}r}\,\hat{\mathbf{r}}\,, \tag{3.2.25}$$

where $\hat{\mathbf{r}}$ is a radial unit 3-vector. We see then that $\mathbf{F}_\varphi(r)$ describes an attractive force. In particular, for a light scalar, the second term in the brackets is negligible and the exponential term approximately unity. If the remaining parameters are $\mathcal{O}(1)$, then it is clear that φ mediates a gravitational strength long range force $F_\varphi \;\coloneqq\; |\mathbf{F}_\varphi| \sim \frac{1}{r^2}$. The problem with this is that precision tests of GR on solar system scales strongly prohibit the presence of such a force, and so one needs a mechanism by which its effect is negligible in high density regions of spacetime, in doing so, rendering them undetectable at the current experimental precision. Fortunately, the fact that the remaining parameters, $\tilde{g}$, Z, c_s and m all depend on the background configuration of the scalar field provide ways by which one can generate a screening mechanism. Indeed, such mechanisms can be understood in terms of making one (or more) of these parameters dependent on the background environment:

1. **Weak coupling**: If one lets only the scalar-matter coupling $\tilde{g}$ depend on the environment, then it is possible to have scenarios in which for low density regions, $\tilde{g} \sim \mathcal{O}(1)$, such that a gravitational strength fifth force is present, however, for high density regions, $\tilde{g}$ can be made small enough such that the fifth force is sufficiently weak to satisfy constraints placed by local tests of GR. Examples of such cases include the *symmetron* [145–147] and *varying dilaton* models [148, 149].

2. **High mass**: An alternative is to allow the mass $m(\bar{\phi})$ depend on the local matter density, such that in high density regions, the field acquires a large mass, enabling a Yukawa-like suppression, and thus rendering its interaction short-ranged ($r \sim \frac{1}{m}$), to the extent that its effects are unobservable. In low density regions (e.g. interstellar space), its mass would be much lighter and thus the scalar would mediate a graviational fifth force. This is essential the assumption of the *Chameleon mechanism* [150, 151].

3. **Kinetic screening**: A third option is to allow the kinetic function $Z(\bar{\phi})$ to depend on the background environment - becoming large in high density regions. This corresponds to a *kinetic*-type screening, in which either the first or second derivatives become important. The case in which first derivatives of the field become relevant is employed by theoretical models such as *k-Mouflage* [152, 153], *DBI* and *k-essence* models [154, 155]. Where second derivatives of the scalar field become important corresponds to the *Vainshtein screening* mechanism [156, 157].

4. **Sound speed**: The final option is to consider cases in which the effective speed of sound $c_s(\bar{\phi})$ becomes very large in high density regions, however, this is not particulary useful since it manifestly relies on superluminality. Furthermore, for time dependent sources the screening will not be very efficient, since c_s will only multiply spatial gradients of the field (cancellations between $\dot{\bar{\phi}}$ and $\nabla\bar{\phi}$ terms contained in Z can occur, leading to smaller values of $|Z|$).

It is evident that, in general, a modified theory of gravity will require a screening mechanism to be empirically viable. That being said, it should be noted that such screening mechanisms are not without their own problems. For example, the

Chameleon mechanism came about through an attempt to explain the non-detection of light scalar fields, with gravitational strength couplings to matter fields, in local experiments. The idea being that the effective mass of such a scalar field is dependent on the local environment, such that it mediates a force long range force in low density regions, but is heavily suppressed in high density regions (such as the solar system), rendering it an extremely short range force on local scales, hence the name the "Chameleon". In principle, this would enable it to have non-trivial effects on large distance scales, but remain undetectable to current local gravitational experiments.

As such, it was orignally though that the Chameleon might explain the late-time acceleration of the universe [158]. Unfortunately, it has been shown that under certain assumptions, it is unable to account for dark energy by itself as a genunine modification of gravity, instead requiring a form of dark energy [159, 160]. Indeed, J. Wang, L. Hui and J. Koury argued in [160] that one can place an upper bound on the Compton wavelength of the Chameleon of $\mathcal{O}(1\text{Mpc})$, three orders of magnitude below the present Hubble scale, and thus restricting its impact to non-linear scales (due to Yukawa suppression). A further issue with the Chameleon model is that the coupling between the Chameleon and matter degrees of freedom, which is essential for successful screening on local scales, causes a breakdown in its classical description in the epoch of Big Bang Nucleosynthesis (BBN) due to the SM fields transferring from relativistic to non-relativistic [161, 162]. At these energy scales, one expects to have control over the theory if it is to present a viable alternative to GR. The primary cause of this breakdown of the classical description is due to the conformal coupling, $\beta \sim \mathcal{O}(1)$, between matter and the Chameleon. Recently, a possible solution to alleviate this problem was presented, which involves modifying the original Chameleon theory, such that it can support a Vainshtein mechanism, and in doing so, dynamically weakening the coupling to the matter degrees of freedom, and rendering the theory well-behaved during BBN [163].

A further example is the Vainshtein mechanism, originally constructed to resolve issues in massive gravity [156], in particular the *vDVZ discontinuity* [164, 165]. This mechanism is by no means exclusive to massive gravity and is in fact employable in other modifications of gravity, in particular those that introduce new scalar degrees of freedom that participate in non-linear derivative self-interactions (see e.g. [166–170]). Indeed, it has since been proven to provide a useful method to locally screen additional forces that inevitably arise as a result of introducing these new degrees of freedom. Such theories always possess a minimal shift symmetry of the scalar field, $\phi \to \phi + c$ (where c is some constant), to the extent that $m\left(\bar{\phi}\right) = 0$, and $\tilde{g} \sim \mathcal{O}(1)$ (cf. eq. (3.2.21)).

The Vainshtein mechanism relies on the theory containing higher-order non-linear self-interactions of the scalar field, which requires $Z^{\mu\nu}$ (in eq. (3.2.21)) to be dependent on second-order derivatives $\partial\partial\phi$ of the field. Specifically, on distance scales greater than the so-called *Vainshtein radius*, r_V, we have that $Z^{\mu\nu} \sim \mathcal{O}(1)$, and as such, the higher-derivative non-linear self-interactions of the scalar fluctuation φ are suppressed relative to the linear contributions to the scalar EOM (3.2.23). We therefore find, that φ mediates a long-range *fifth force* of gravitational strength on these scales. However, for $r \ll r_V$, we have that $Z^{\mu\nu} \gg 1$, such that the non-linear derivative self-interactions start to dominate the EOM (3.2.23). It is then found that, upon a canoncial normalisation of the scalar fluctuation, $\varphi \to Z^{-1/2}\varphi$, it couples to the source T with a strength $\sim \frac{1}{M_{\rm Pl}\sqrt{Z}} \ll \frac{1}{M_{\rm Pl}}$. Consequently, the fifth force sourced by φ is much weaker then the force propagated by the spin-2 graviton of GR. This relative suppression on local scales can be exploited to screen such fifth forces from local gravitational experiments, such that the theory mimics GR on local scales, and they remain undetectable. We refer the reader to the following reviews [62, 63, 171] for a detailed introduction to the Vainshtein mechanism.

In principle, this is a powerful tool at our disposal, enabling one quite a bit of freedom to construct a modified theory of gravity, whilst still being able to pass solar system tests of gravity. However, as was the case with the Chameleon mechanism, it is not without its issues. In particular, the main problem arises from positivity constraints placed on the coupling constants of the non-linear interactions, which suggest that for theories such as the Galileon, one cannot construct an EFT that implements the Vainshtein mechanism on local scales, whilst simultaneously describing the low-energy limit of a local, Lorentz invariant UV completion [172]. Furthermore, for theories where a UV completion is known, doubt has been cast on whether the corresponding low-energy EFTs, that are used to exhibit Vainshtein phenomena, can be trusted [173]. Having said this, given the ability of the Vainshtein mechanism to provide efficient screening, it still remains employed as a method to cloak the presence of additional degrees of freedom generically introduced in modified theories of gravity.

Chapter 4

Self-tuning solutions to the CCP

4.1 Horndeski theory and the Fab-Four

Having set the scene we see that, foregoing a modification to the field theory sector, one is not really left with much choice but to modify GR. The most popular route taken is to consider a minimal approach in which one introduces an additional degree of freedom into the gravitational sector in the form of a scalar field ϕ, and in doing so, constructing a so-called *scalar-tensor* theory of gravity. Indeed, this point of view has proven to be useful in a wide range of models, one of the earliest being Brans-Dicke theory, and more recently models inspired by Galileon theory [166] have been developed [167–170, 174–178]. Of course, as we have discussed in §3.2.2, one has to be very careful when formulating a new theory, importantly, ensuring that it does not propagate Ostrogradsky ghost instabilities, furthermore, given that GR has so far proven very successful phenomenologically, one must keep at least some of its structure in order to agree with experimental data. Given this, several desired features are most often required for a modified theory of gravity: that it is causal; it satisfies the EEP; and that the corresponding EOM are of second-order in derivatives. In fact, it was first discovered by G.W. Horndeski in 1974 [179], and more recently re-discovered by C. Deffayet et al. [180], that the most general class of scalar-tensor theories possessing second-order field equations, are described by the so-called *Horndeski* action,

$$S_{\mathrm{H}}\left[\phi, g_{\mu\nu}\right] \;=\; \sum_{i\,=\,2}^{5}\int \mathrm{d}^4x\sqrt{-g}\;\mathscr{L}_{(i)}\,, \tag{4.1.1}$$

where the component Lagrangians are defined as,

$$\mathscr{L}_{(2)} = K(\phi, X) \;, \tag{4.1.2a}$$

$$\mathscr{L}_{(3)} = -G_3(\phi, X)\,\Box\phi \,, \tag{4.1.2b}$$

$$\mathscr{L}_{(4)} = G_4(\phi, X)\,R + G_{4,X}\left[(\Box\phi)^2 - (\nabla_\mu\nabla_\nu\phi)^2\right] \,, \tag{4.1.2c}$$

$$\mathscr{L}_{(5)} = G_5(\phi, X)\,G^{\mu\nu}\nabla_\mu\nabla_\nu\phi - \frac{1}{6}G_{5,X}\left[(\Box\phi)^3 - 3\Box\phi\,(\nabla_\mu\nabla_\nu\phi)^2 + 2\,(\nabla_\mu\nabla_\nu\phi)^3\right] \,, \tag{4.1.2d}$$

where $X = -\frac{1}{2}\partial^\mu\phi\partial_\mu\phi$ is the canonical kinetic term for the scalar field $\phi = \phi(x)$. The corresponding field equations for the Horndeski action [eq. (4.1.1)] are then $\mathcal{E}^{\mu\nu} := \frac{1}{\sqrt{-g}}\frac{\delta S_{\mathrm{H}}}{\delta g_{\mu\nu}} = 0$ and $\mathcal{E}^\phi := \frac{1}{\sqrt{-g}}\frac{\delta S_{\mathrm{H}}}{\delta\phi} = 0$. Note that, as the theory is diffeomorphism invariant, the scalar field equation $\mathcal{E}_\phi$ and the metric field equation $\mathcal{E}^{\mu\nu}$ are related via $\nabla_\mu\mathcal{E}^{\mu\nu} = \frac{1}{2}\mathcal{E}^\phi\nabla^\nu\phi$, and so $\mathcal{E}^\phi$ can be readily determined given knowledge of $\mathcal{E}^{\mu\nu}$.

Introducing a matter action $S_{\mathrm{M}}[g_{\mu\nu}, \Psi]$, one can ensure that the EEP is satisfied by requiring that the matter fields are minimally coupled to gravity via the metric $g_{\mu\nu}$, i.e. there are no *direct* couplings between the scalar field ϕ and the matter fields Ψ. The full action for the theory is thus given by

$$S[\phi, \Psi, g_{\mu\nu}] = S_{\mathrm{H}}[\phi, g_{\mu\nu}] + S_{\mathrm{M}}[g_{\mu\nu}, \Psi] \;. \tag{4.1.3}$$

From which we obtain the following field equations: $\mathcal{E}^{\mu\nu} = \frac{1}{2}T^{\mu\nu}$ and $\mathcal{E}^\phi = 0$, where $T^{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S_{\mathrm{M}}}{\delta g_{\mu\nu}}$.

As mentioned briefly in §3.2.1, it is in principle possible to evade Weinberg's no-go theorem. Indeed, this possiblility was explored within the class of Horndeski theories, and in doing so, a general class of self-tuning theories, the so-called *Fab-Four*, was discovered [60, 181]. It was found that one can solve the CCP, at least at the classical level, through a viable self-tuning mechanism that completely screens the spacetime curvature from the cosmological constant. Moreover, the Fab-Four may also be able to solve the radiative instability that is at the heart of the CCP, as at least heuristically, it appears that the radiative corrections in this theory are manageable, given some not too restrictive conditions.

The essential idea of Fab-Four theory is to allow the cosmological constant to exist, and take on any value it may, but in such a way that its presence does not affect the spacetime geometry seen by matter, i.e. it simply does not source curvature. The starting point in the construction of this theory is to consider a scalar field ϕ that is able to self-tune such that it can screen the spacetime geometry from the cosmological constant. Furthermore, since the cosmological constant manifests itself in the infrared sector of gravity, we require that ϕ is a *light* degree of freedom, i.e. $m_\phi \sim H_0$, such that it can mediate long-range interactions.[1] Additionally, so as not to fall afoul of Weinberg's no-go theorem, one breaks Poincaré invariance at the level of the self-tuning field ϕ, that is, one allows the self-tuning vacuum configurations of ϕ to become time-dependent, $\phi = \phi(t)$. To proceed, we shall therefore consider Horndeski's theory on FRW backgrounds, for which we have a spatially homogeneous scalar field $\phi = \phi(t)$, and the metric is of the form given by eq. (2.1.1). Then, since we wish to determine whether Horndeski's theory admits self-tuning solutions, we need to define what it means for a theory to be able to self-tune in a relatively model independent manner. To do so, we define a so-called *self-tuning filter*:

S.1 The vacuum solution of the theory should always be Minkowksi, regardless of the value of the net cosmological constant;

S.2 This should remain true before and after any phase transition in which the cosmological constant jumps instantaneously by a finite amount;

S.3 The theory should admit a non-trivial cosmology, in other words, it does not self-tune for any other matter backgrounds other than vacuum energy (ensuring that Minkowksi spacetime is not the only solution, a condition that is certainly required by observational data).

This provides us with a template for viable cosmological field equations, i.e. those that describe dynamical evolution towards Minkowksi spacetime as some sort of fixed point. By this it is meant that, if we are on a Minkowksi solution, we stay there, otherwise we dynamically evolve to it as an attractor solution.[2] It is worth noting that *any* diffeomorphism invariant theory that passes through both the first two conditions of this filter will admit a Minkowksi solution in the presence of not just a

[1] If it were any heavier, then Yukawa suppression would reduce its range to below the Hubble scale, since $r \sim \frac{1}{m_\phi}$, and hence it would not be able to mediate interactions on cosmological scales.

[2] It has been shown in [182] that there are many examples within the Fab-Four class of self-tuning theories in which this is the case.

cosmological constant, but an arbitrary cosmological fluid. The reason being is that the vacuum energy density corresponds to a piece-wise constant function, with discontinuities arising at phase transitions, which in principle, can occur at any time. As such, a Minkowksi solution can be returned for all piece-wise constant energy densities. The third condition is then in place to ensure that a theory, satisfying the first two conditions, also permits (in some sense[3]) reasonable matter dominated cosmologies.

Remarkably, having passed Horndeski theory through the self-tuning filter, and applying the ensuing constraints, it was found that one can establish a self-tuning, minisuperspace Lagrangian on the FRW background. Furthermore, it was shown in [60], that it is possible to reconstruct curvature invariants from the corresponding cosmological field equations, leading to a set of base Lagrangians defining the Fab-Four theory:

$$S_{\text{Fab-Four}}[\phi, \Psi, g_{\mu\nu}] = \int d^4x\sqrt{-g}\left[\mathscr{L}_j + \mathscr{L}_p + \mathscr{L}_g + \mathscr{L}_r - \Lambda_{\text{bare}}\right] + S_{\text{M}}\left[g_{\mu\nu}, \Psi\right] , \tag{4.1.4}$$

where Λ_{bare} is a bare cosmological constant (which can always be absorbed into a renormalisation of the vacuum energy contained in S_{M}), and the base Lagrangians, denoted *John*, *Paul*, *Ringo* and *George*, are given by

$$\mathscr{L}_j = V_j(\phi)G^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi\,, \tag{4.1.5a}$$

$$\mathscr{L}_p = V_p(\phi)P^{\mu\nu\alpha\beta}\nabla_\mu\phi\nabla_\alpha\phi\nabla_\nu\nabla_\beta\phi\,, \tag{4.1.5b}$$

$$\mathscr{L}_g = V_g(\phi)R\,, \tag{4.1.5c}$$

$$\mathscr{L}_r = V_r(\phi)\hat{\mathcal{G}}\,, \tag{4.1.5d}$$

where $\{V_i\}$ $(i = \{j, p, g, r\})$ are arbitrary functions of the scalar field ϕ, R is the Ricci scalar, $G^{\mu\nu}$ the Einstein tensor, $\hat{\mathcal{G}}$ the Gauss-Bonnet combination, and $P^{\mu\nu\alpha\beta}$ the double-dual of the Riemann tensor.

This result implies that *any* self-tuning scalar-tensor theory (within the class of

[3]The caveat here is that it was found in [182], that the scalar field completely determines the dynamical evolution of the universe regardless of the dominant matter content. This is problematic, as one expects the dynamical evolution of the universe to be dictated by its matter content. In this sense, the Fab-Four self-tuning mechanism is too good. That being said, despite not providing the ultimate solution to the CCP, it is a significant step in the direction of one.

Horndeski theories), that satisfies the EEP, must be constructed from the Fab-Four. The four base Lagrangians are not all able to self-tune in isolation, in particular, it was found that $\mathscr{L}_r$ is unable to give rise to self-tuning "without a little help from [its] friends" [181], John and Paul (hence the name Fab-Four). In this particular case, Ringo has a non-trivial effect on the cosmological dynamics, but self-tuning is still possible. Furthermore, when $\{V_j = 0, V_p = 0, V_g = \text{const.}, V_r = 0\}$, the Fab-Four reduces to GR and is therefore unable to self-tune. If we allow V_g to vary (i.e. $V_g \neq \text{const.}$) then we recover a particular limit of Brans-Dicke gravity ($w = 0$), which is able to self-tune, but ruled out by solar system constraints. It follows that one should consider the Fab-Four Lagrangians, not in isolation as theories in their own right, but as combining to give a single self-tuning theory. Of course, in order for the Fab-Four to be a viable physical theory, it must permit some sort of screening mechanism such that it can pass solar system tests (since GR is well tested at these scales). From eq. (4.1.5), we see that John and Paul have non-trivial derivative interactions contained in them, and as such, these may be able to provide the require screening through Vainshtein or Chameleon effects [150, 156, 181].

Before proceeding, we remark that, although the Fab-Four certainly provides a solution to the CCP at the classical level, (as we discussed in §3.1.1) a full solution to the problem must resolve the issue of radiative instability of the vacuum. An initial heurisitic analysis tentatively suggests that it might be possible to render certain forms of Fab-four theory safe from large quantum corrections around a self-tuning background, provided that the cut-off of the effective theory Λ_{UV} satisfies $\sqrt{G_{\mathrm{eff}}\rho_\Lambda} < \Lambda_{\mathrm{UV}} < \rho_\Lambda^{1/4}$, where G_{eff} is the gravitational coupling strength to matter in the *linearised* regime [60]. It should be noted, however, that the radiative stability of the theory cannot be fully addressed without a better understanding of the preferred background solutions and potentials, since the corrections are sensitive to the cut-off, itself sensitive to the background.

4.2 Disformal self-tuning

Given the success of the Fab-Four at providing a potential viable solution to the CCP, one is left wondering if it is possible to generalise this result even further. Indeed, a minimal extension was found by Babichev *et al.* [183], in which they consider Fab-Four potentials V_i that depend on both the scalar field ϕ and its corresponding canonical kinetic term X. However, one can consider an alternative generalisation in which one allows the scalar field to enter the matter sector, interacting with matter directly. Before proceeding, we note that the class of Horndeski theories is invariant under conformal transformations of the metric, $g_{\mu\nu}(x) \to \bar{g}_{\mu\nu}(x) = A(\phi)g_{\mu\nu}(x)$ (where ϕ is some scalar field), by which we mean that if a given scalar-tensor theory is related to Horndeski theory by a conformal transformation, then it can always be put into Horndeski form (i.e. of the form given by eq. (4.1.1)) by appropriate field re-definitions $\phi \to \bar{\phi} = F(\phi)$ [184]. Accordingly, in order to achieve such a generalisation, one must consider a *disformal* coupling of matter to gravity. The presence of disformal couplings is theoretically well motivated, appearing in the low energy effective action of string theory [185], and manifesting in Galileon theories through the induced metric on probe branes within higher dimensional spacetimes [186, 187], as well as in the decoupling limit of massive gravity [188].

Following this approach, we consider two distinct (but related) geometries: one defining the geometry on which matter plays out its dynamics, and one describing gravitation. We refer to these as *physical* and *gravitational* geometries, described by the two metrics, $\bar{g}_{\mu\nu}$ and $g_{\mu\nu}$, respectively. It was shown by Bekenstein [189], that the most general relation between these two metrics, involving a scalar field $\phi = \phi(x)$ and its canonically conjugate kinetic term X, adhering to the WEP and causality, is given by the following *disformal* transformation,

$$\bar{g}_{\mu\nu}(x) \;=\; A^2(\phi, X)\left[g_{\mu\nu}(x) + B^2(\phi, X)\partial_\mu\phi\partial_\nu\phi\right] \,, \tag{4.2.1}$$

where $A(\phi, X)$ and $B(\phi, X)$ are arbitrary functions of ϕ and X. We further note that it has been shown that the Horndeski action is invariant (up to field redefinitions) under disformal transformations of the metric [eq. (4.2.1)] in which A and B depend solely on ϕ (and *not* X) [184], and so in order to fully generalise beyond any subset of Horndeski theory, one must consider cases in which A and B are non-trivially dependent on X as well as ϕ.

The inverse metric $\bar{g}^{\mu\nu}$ transformation can be determined using the *Sherman-Morrison formula* [190, 191] for the inverse of the sum of an invertible matrix M and the outer product $\mathbf{u}\mathbf{v}^T$, of two vectors $\mathbf{u}$ and $\mathbf{v}$:

$$\left(M + \mathbf{u}\mathbf{v}^T\right)^{-1} \;=\; M^{-1} - \frac{M^{-1}\mathbf{u}\mathbf{v}^T M^{-1}}{1 + \mathbf{v}^T M^{-1}\mathbf{u}} \,, \tag{4.2.2}$$

In our case, the role of M is played by the metric $g_{\mu\nu}$, and the roles of $\mathbf{u}$ and $\mathbf{v}$ are both played by the gradient of the scalar field $\partial_\mu\phi$ (appropriately multiplied by factors of A an B). As such, we find that the inverse metric $\bar{g}^{\mu\nu}$ is of the form,

$$\bar{g}^{\mu\nu} \;=\; \frac{1}{A^2(\phi,X)}\left[g^{\mu\nu} - \frac{B^2(\phi,X)}{1-2B^2(\phi,X)X}\,g^{\mu\lambda}g^{\nu\sigma}\partial_\lambda\phi\partial_\sigma\phi\right]. \tag{4.2.3}$$

Moreover, we can determine the relationship between the integration measures $\sqrt{-g}$ and $\sqrt{-\bar{g}}$ (where $g := \det(g_{\mu\nu})$ and likewise for $\bar{g}$) from the determinant relation $\det\left(M+\mathbf{u}\mathbf{v}^T\right) = \left(1+\mathbf{v}^T M^{-1}\mathbf{u}\right)\det M$. Indeed, we have that,

$$\bar{g} \;=\; A^8\left(1-2B^2X\right)g \quad\Longrightarrow\quad \sqrt{-\bar{g}} \;=\; A^4\sqrt{1-2B^2X}\sqrt{-g}\,. \tag{4.2.4}$$

From eq. (4.2.3), we are also lead to an expression for the canonical kinetic term X, in terms of the physical metric,

$$\bar{X}(\phi,X) \;=\; -\frac{1}{2}\bar{g}^{\mu\nu}\partial_\mu\phi\partial_\nu\phi \;=\; \frac{X}{A^2(\phi,X)\left(1-2B^2(\phi,X)X\right)}\,. \tag{4.2.5}$$

In principle, we can invert this equation such that $X = X(\phi,\bar{X})$, and with this in mind, we can rearrange eq. (4.2.1) to arrive at an expression for the inverse disformal transformation, from the physical to the gravitational metric,

$$\begin{aligned} g_{\mu\nu}(x) \;&=\; \frac{1}{A^2(\phi,X)}\left[\bar{g}_{\mu\nu}(x) - A^2(\phi,X)B^2(\phi,X)\partial_\mu\phi\partial_\nu\phi\right] \\ &:=\; \bar{A}^2(\phi,\bar{X})\left[\bar{g}_{\mu\nu}(x) - \bar{B}^2(\phi,\bar{X})\partial_\mu\phi\partial_\nu\phi\right], \end{aligned} \tag{4.2.6}$$

where X is implicitly dependent on ϕ and $\bar{X}$. From this, we can imply that $\bar{A}$ and $\bar{B}$ are related to A and B in the following manner,

$$\bar{A}^2(\phi,\bar{X}) \;=\; \frac{1}{A^2(\phi,X)}\,, \qquad \bar{B}^2(\phi,\bar{X}) \;=\; -A^2(\phi,X)B^2(\phi,X)\,. \tag{4.2.7}$$

Equipped with this knowledge, we are able to delve into the details of constructing a disformal generalisation of the Fab-Four.

Our aim is to construct a self-tuning theory of gravity in which matter is disformally coupled to the gravitational metric, described by an action of the form

$$S \;=\; S_H[g_{\mu\nu},\phi] + S_{\mathrm{M}}[\bar{g}_{\mu\nu},\Psi]\,, \tag{4.2.8}$$

where, as usual, Ψ collectively denotes the matter fields. Note that, in this representation of the theory, there is a direct coupling of both gravity and matter to the scalar field ϕ, due to the non-trivial dependence of $\bar{g}_{\mu\nu}$ on ϕ and its gradient $\partial_\mu\phi$. The reason this is the case is that, in this representation, we treat $g_{\mu\nu}$, ϕ and Ψ as our fundamental dynamical variables, with the physical geometry $\bar{g}_{\mu\nu}$, determined via eq. (4.2.1). An alternative representation of the theory can be found by expressing the action, eq. (4.2.8), in terms of $\bar{g}_{\mu\nu}$, ϕ and Ψ, treating these as our fundamental dynamical variables, with $g_{\mu\nu}$ determined via eq. (4.2.6),

$$S \;=\; S_H[\bar{g}_{\mu\nu}, \phi] + S_{\mathrm{M}}[\bar{g}_{\mu\nu}, \Psi]\,. \tag{4.2.9}$$

In doing so we eliminate any direct coupling of ϕ to the matter sector at the level of the action, however, the gravitational sector remains directly coupled to to the scalar field ϕ. In principle, there is an infinite number of ways in which we could represent our theory, however, these two particular representations are often the most useful. Conventionally, these different representations are referred to as "*frames*", and adopting this terminology, we shall refer to the representation in which the action has the form given by eq. (4.2.8), as the *Horndeski frame*, which is the analogue of the Einstein frame.[4] The representation in which we can express the action in the form given by eq. (4.2.9) is referred to as the *Jordan frame*. At least at the classical level, all physical observables are independent of the frame we calculate them in, and so we are free to work in whichever frame is most suitable.

It is in the Jordan frame that matter is minimally couple to the (physical) metric $\bar{g}_{\mu\nu}$, and follows the geodesics defined by it. This is in stark contrast to the Horndeski frame, in which the matter geodesics defined by the metric are also influenced by variations in the scalar field ϕ. Furthermore, in the Jordon frame, the matter energy-momentum tensor is locally conserved, i.e. $\bar{\nabla}_\mu \bar{T}^{\mu\nu} = 0$, whereas, in the Horndeski frame this is not true, $\nabla_\mu T^{\mu\nu} \neq 0$, due to the direct coupling between the scalar field ϕ and the matter fields. It is for these reasons that we consider the Jordan frame to be the *physical* frame: matter follows geodesics defined purely by the metric $\bar{g}_{\mu\nu}$, and its energy-momentum tensor $\bar{T}^{\mu\nu}$ is locally conserved.

[4]The Einstein frame is a representation of a theory in which the gravitational action is of the form of the EH action [eq. (1.2.1)].

4.2.1 Towards a disformally self-tuning theory of gravity

As our aim here is to construct a generalisation of Fab-Four theory, we shall adopt the self-tuning structure laid out in its original derivation. Thus, by a self-tuning theory, we assume the existence of a spatially homogeneous, light scalar field $\phi = \phi(t)$, that is able to evolve in such a way as to dynamically screen any vacuum energy density contributed by the net cosmological constant from the spacetime curvature seen by matter. In other words, the presence of the scalar field effectively degravitates the net cosmological constant. Moreover, in order for the theory to be a physically legitimate self-tuning theory of gravity, we require it to satify the requirements laid out by the self-tuning filter, **S.1 - S.3**.

To proceed, we shall set-up the cosmological structure in which we shall carry out our self-tuning analysis, and to do so, we require that the geometry in both the Horndeski and the Jordan frame are both FRW. To this end, in the Horndeski frame, where we treat $g_{\mu\nu}$, ϕ and Ψ as our fundamental dynamical variables, with the Jordan-frame (physical) metric, $\bar{g}_{\mu\nu}$, defined via eq. (4.2.1), and foliate spacetime into a set of space-like hypersurfaces, Σ_t, such that the spatial "slice" at each given instant in time t is homogeneous and isotropic (this corresponds to making a *minisuperspace* approximation[5]). The geometry defined by the Horndeski-frame (gravitational) metric $g_{\mu\nu}$ is then,

$$\mathrm{d}s^2 \;=\; g_{\mu\nu}(x)\mathrm{d}x^\mu \mathrm{d}x^\nu \;=\; -N^2(t)\mathrm{d}t^2 + a^2(t)\gamma_{ij}(\mathbf{x})\mathrm{d}x^i\mathrm{d}x^j \,, \tag{4.2.10}$$

where $N(t)$ is the lapse function, $a(t)$ the scale factor, and $\gamma_{ij}(\mathbf{x})$ the (maximally symmetric) metric on the plane ($k = 0$), sphere ($k = 1$), or hyperboloid ($k = -1$) given by eq. (2.1.2).

In the Jordan frame, where we instead treat $\bar{g}_{\mu\nu}$, ϕ and Ψ as our fundamental dynamical variables, with the Horndeski-frame metric, $g_{\mu\nu}$ defined via eq. (4.2.6), we similarly stipulate that the geometry defined by the Jordan-frame metric $\bar{g}_{\mu\nu}$ is asymptotically Minkowksi in form, and as such,

$$\mathrm{d}\bar{s}^2 \;=\; \bar{g}_{\mu\nu}(x)\mathrm{d}x^\mu \mathrm{d}x^\nu \;=\; -\mathrm{d}t^2 + \bar{a}^2(t)\gamma_{ij}(\mathbf{x})\mathrm{d}x^i\mathrm{d}x^j \,. \tag{4.2.11}$$

[5] *Superspace* is the configuration space in the Hamiltonian formulation of GR; it is the set of equivalence classes of 3-dimensional metrics $\gamma_{ij}(\mathbf{x})$ on the hypersurfaces Σ_t foliating spacetime. The minisuperspace approximation corresponds to neglecting spatial dependence in the metric, taking our gravitational degrees of freedom to be purely time-dependent.

An immediate consequence of imposing the two geometries Eqs. (4.2.10) and (4.2.11), along with the assumption that the scalar field is spatially homogeneous, is that the canonical kinetic term X takes on the following explicit form in the Horndeski frame

$$X \;=\; -\frac{1}{2}g^{00}\partial_0\phi\partial_0\phi \;=\; \frac{1}{2}\left(\frac{\dot{\phi}}{N}\right)^2 , \tag{4.2.12}$$

and

$$\bar{X} \;=\; -\frac{1}{2}\bar{g}^{00}\partial_0\phi\partial_0\phi \;=\; \frac{1}{2}\dot{\phi}^2 , \tag{4.2.13}$$

in the Jordan frame. Moreover, we can determine a mapping between the Horndeski and Jordan frames for the lapse function $N(t)$ and the scale factor $a(t)$ by noting that the geometry defined by $g_{\mu\nu}$, eq. (4.2.10), can be expressed in the Jordan frame via the inverse disformal transformation, eq. (4.2.6),

$$\mathrm{d}s^2 \;=\; -\bar{A}^2(\phi,\bar{X})\left[1-2\bar{B}^2(\phi,\bar{X})\bar{X}\right]\mathrm{d}t^2+\bar{a}^2(t)\bar{A}^2(\phi,\bar{X})\gamma_{ij}(\mathbf{x})\mathrm{d}x^i\mathrm{d}x^j . \tag{4.2.14}$$

Since this expression must be equivalent to eq. (4.2.10), we can thus imply the following relations,

$$N^2(t) \;=\; \bar{A}^2(\phi,\bar{X})\left[1-2\bar{B}^2(\phi,\bar{X})\bar{X}\right] , \tag{4.2.15a}$$

$$a^2(t) \;=\; \bar{a}^2(t)\bar{A}^2(\phi,\bar{X}) , \tag{4.2.15b}$$

providing a useful check for the relation between X and $\bar{X}$, eq. (4.2.5).

Having set the cosmological scene, we remark that, since we are considering a spatially homogeneous scalar field, propagating on an FRW background, the Horndeski Lagrangian is homogeneous, and thus the we can define a corresponding cosmological minisuperspace action of the form

$$\tilde{S}_{\mathrm{H}} \;\coloneqq\; \frac{S_{\mathrm{H}}[g_{\mu\nu}|_{\mathrm{FRW}},\phi]}{\int \mathrm{d}^3x\sqrt{\gamma}} \;=\; \frac{\int \mathrm{d}^4x\sqrt{-g}\,\mathscr{L}_{\mathrm{H}}\big|_{\mathrm{FRW}}}{\int \mathrm{d}^3\mathbf{x}\sqrt{\gamma}} \;=\; \int \mathrm{d}t\sqrt{-\tilde{g}}\,\mathscr{L}_{\mathrm{H}}\big|_{\mathrm{FRW}} \;\coloneqq\; \int \mathrm{d}t\,\mathcal{L}(t) , \tag{4.2.16}$$

where we have used that $\sqrt{-g}=N(t)a^3(t)\sqrt{\gamma}$ (noting that γ_{ij} is maximally symmetric), and defined an effective metric determinant, $g\to\tilde{g}$, such that $\sqrt{-\tilde{g}}=N(t)a^3(t)$. Furthermore, we have defined the cosmological minisuperspace Lagrangian as $\mathcal{L}(t)=\sqrt{-\tilde{g}}\mathscr{L}_{\mathrm{H}}(t)\big|_{\mathrm{FRW}}$.

Given that the Horndeski Lagrangian encapsulates the dynamics of the gravitational sector, we shall proceed to evaluate the appropriate curvature terms on the metric defined in this sector, $g_{\mu\nu}$. We can subsequently map to their corresponding descriptions in terms of the physical metric $\bar{g}_{\mu\nu}$ using the inverse disformal transformation [eq. (4.2.6)] and in doing so, obtain an expression for $\mathcal{L}(t)$ in the Jordan frame:

$$\mathcal{L}(t) \;=\; \sqrt{-\tilde{g}}\sum_{i\,=\,2}^{5}\mathscr{L}_{(i)}\Big|_{\text{FRW}} \;=\; \bar{a}^3\sum_{j\,=\,0}^{3} Z_j(\bar{a},\phi,\dot{\phi},\ddot{\phi})\bar{H}^j\,, \tag{4.2.17}$$

where $\bar{H} = \frac{\dot{\bar{a}}}{\bar{a}}$ is the Hubble parameter evaluated in the Jordan frame, and the functions $Z_j = Z_j(\bar{a},\phi,\dot{\phi},\ddot{\phi})$ are defined as follows

$$\begin{aligned} Z_0 \;=\;& N\bar{A}^3K + \left(\bar{A}^3\right)^{\bullet}\frac{\dot{\phi}}{N}G_3 - \dot{\phi}\tilde{G}_{3,\phi} - 6\frac{\bar{A}\dot{\bar{A}}^2}{N}G_4 - 6\bar{A}^2\dot{\bar{A}}\frac{\dot{\phi}}{N}G_{4,\phi} \\ &+ 6\frac{\bar{A}\dot{\bar{A}}^2}{N}\left(\frac{\dot{\phi}}{N}\right)^2 G_{4,X} + \frac{\dot{\bar{A}}^3}{N^2}\left(\frac{\dot{\phi}}{N}\right)^3 G_{5,X} - 3\frac{\bar{A}\dot{\bar{A}}^2}{N}\left(\frac{\dot{\phi}}{N}\right)^2 G_{5,\phi} \\ &- \frac{k}{\bar{a}^2}\left(3\dot{\phi}\tilde{G}_{5,\phi} - 3\dot{\bar{A}}\frac{\dot{\phi}}{N}G_5 - 6N\bar{A}G_4\right) \\ =\;& \mathscr{X}_0(\phi,\dot{\phi},\ddot{\phi}) - \frac{k}{\bar{a}^2}\mathscr{Y}_0(\phi,\dot{\phi},\ddot{\phi})\,, \end{aligned} \tag{4.2.18a}$$

$$\begin{aligned} Z_1 \;=\;& 3\bar{A}^3\frac{\dot{\phi}}{N}G_3 - 3\tilde{G}_3 + 12\frac{\bar{A}^2\dot{\bar{A}}}{N}\left(\frac{\dot{\phi}}{N}\right)^2 G_{4,X} - 6\bar{A}^3\frac{\dot{\phi}}{N}G_{4,\phi} - 12\frac{\bar{A}^2\dot{\bar{A}}}{N}G_4 \\ &+ 3\frac{\bar{A}\dot{\bar{A}}^2}{N^2}\left(\frac{\dot{\phi}}{N}\right)^3 G_{5,X} - 6\frac{\bar{A}^2\dot{\bar{A}}}{N}\left(\frac{\dot{\phi}}{N}\right)^2 G_{5,\phi} - \frac{k}{\bar{a}^2}\left(3\tilde{G}_5 - 3\bar{A}\frac{\dot{\phi}}{N}G_5\right) \\ =\;& \mathscr{X}_1(\phi,\dot{\phi},\ddot{\phi}) - \frac{k}{\bar{a}^2}\mathscr{Y}_1(\phi,\dot{\phi},\ddot{\phi}) \end{aligned} \tag{4.2.18b}$$

$$\begin{aligned} Z_2 \;=\;& 6\frac{\bar{A}^3}{N}\left(\frac{\dot{\phi}}{N}\right)^2 G_{4,X} - 6\frac{\bar{A}^3}{N}G_4 + 3\frac{\bar{A}^2\dot{\bar{A}}}{N^2}\left(\frac{\dot{\phi}}{N}\right)^3 G_{5,X} - 3\frac{\bar{A}^3}{N}\left(\frac{\dot{\phi}}{N}\right)^2 G_{5,\phi} \\ =\;& \mathscr{X}_2(\phi,\dot{\phi},\ddot{\phi})\,, \end{aligned} \tag{4.2.18c}$$

$$Z_3 \;=\; \frac{\bar{A}^3}{N^2}\left(\frac{\dot{\phi}}{N}\right)^3 G_{5,X} \;=\; \mathscr{X}_3(\phi,\dot{\phi},\ddot{\phi})\,, \tag{4.2.18d}$$

where we have implicitly defined two auxilliary functions $\tilde{G}_3$ and $\tilde{G}_5$,

$$\tilde{G}_{3,X} \;\coloneqq\; \frac{\partial \tilde{G}_3}{\partial X} \;=\; \frac{N\bar{A}^3}{\dot{\phi}} G_3 \;=\; \frac{\bar{A}^3}{\sqrt{2X}} G_3\,, \tag{4.2.19a}$$

$$\tilde{G}_{5,X} \;\coloneqq\; \frac{\partial \tilde{G}_5}{\partial X} \;=\; \frac{N\bar{A}}{\dot{\phi}} G_5 \;=\; \frac{\bar{A}}{\sqrt{2X}} G_5\,. \tag{4.2.19b}$$

Upon inspection of the functional forms of the component functions Z_j, we have further noted that they can be expressed in the following form

$$Z_j(\bar{a}, \phi, \dot{\phi}, \ddot{\phi}) \;=\; \mathscr{X}_j(\phi, \dot{\phi}, \ddot{\phi}) - \frac{k}{\bar{a}^2} \mathscr{Y}_j(\phi, \dot{\phi}, \ddot{\phi}) \tag{4.2.20}$$

where $\mathscr{Y}_2 = \mathscr{Y}_3 = 0$. The corresponding Hamiltonian density $\mathscr{H}$ for the gravitational sector can then be ascertained by taking the Legendre transform of eq. (4.2.17) with respect to $\dot{\bar{a}}$, $\dot{\phi}$ and $\ddot{\phi}$,

$$\begin{aligned}
\mathscr{H} \;&=\; \mathscr{H}(\bar{a}, \dot{\bar{a}}, \ddot{\bar{a}}, \phi, \dot{\phi}, \ddot{\phi}, \dddot{\phi}) \;=\; \dot{\bar{a}} p_{\dot{\bar{a}}} + \dot{\phi} p_{\dot{\phi}} + \ddot{\phi} p_{\ddot{\phi}} - \mathcal{L} \\
&=\; \bar{a}^3 \sum_{j\,=\,0}^{3} \Bigg[\Big((j-1)\, Z_j + \dot{\phi} Z_{j,\dot{\phi}} - \dot{\phi}^2 Z_{j,\phi,\ddot{\phi}} - \dot{\phi}\ddot{\phi} Z_{j,\dot{\phi},\ddot{\phi}} + \ddot{\phi} Z_{j,\ddot{\phi}} - \dot{\phi}\dddot{\phi} Z_{j,\ddot{\phi},\ddot{\phi}} \Big) \\
&\qquad + \Big((j-3)\, \dot{\phi} Z_{j,\ddot{\phi}} - \bar{a}\dot{\phi} Z_{j,\bar{a},\ddot{\phi}} \Big) \bar{H} - j \dot{\phi} \frac{\ddot{\bar{a}}}{\bar{a}} Z_{j,\ddot{\phi}} \bar{H}^{-1} \Bigg] \bar{H}^j\,,
\end{aligned} \tag{4.2.21}$$

where we have defined the "Ostrogradsky" canonical momenta $p_{\dot{\bar{a}}} = \frac{\partial \mathcal{L}}{\partial \dot{\bar{a}}}$, $p_{\dot{\phi}} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \ddot{\phi}}\right)$ and $p_{\ddot{\phi}} = \frac{\partial \mathcal{L}}{\partial \ddot{\phi}}$.

The full Jordan frame Hamiltonian density $\mathscr{H}_{\text{total}}$ of the theory can be constructed from the contribution from the gravitational sector $\mathscr{H}$, and a source term from the matter sector in the form of a homogeneous cosmological fluid of energy density ρ_{M}, and pressure p, minimally coupled to the metric, i.e. $\mathscr{H}_{\text{total}} = \mathscr{H} + \rho_{\text{M}}$. Due to the invariance of the theory under temporal (and spatial) diffeomorphisms, the full Hamiltonian satisfies the constraint $\mathscr{H}_{\text{total}} = 0$.

Finally, to obtain the EOM for the scalar field ϕ, we can take advantage of the fact that we are working in the Jordan frame, and as such, matter is decoupled from the scalar field. Thus, the EOM for ϕ can be determined exclusively from eq. (4.2.17),

$$\begin{aligned}
\mathcal{E}^{\phi} &= \mathcal{E}^{\phi}(\bar{a},\dot{\bar{a}},\ddot{\bar{a}},\dddot{\bar{a}},\phi,\dot{\phi},\ddot{\phi},\dddot{\phi},\ddddot{\phi}) \;=\; \frac{\partial \mathcal{L}}{\partial \phi} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\right) + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(\frac{\partial \mathcal{L}}{\partial \ddot{\phi}}\right) \\
&= \bar{a}^3 \sum_{j\,=\,0}^{3} \Bigg[2\bar{a}\,(2-j)\,Z_{j,\bar{a},\ddot{\phi}}\bar{H}^2 + \bar{a}^2 Z_{j,\bar{a},\bar{a},\ddot{\phi}}\bar{H}^2 + 3jZ_{j,\ddot{\phi}}\frac{\ddot{\bar{a}}}{\bar{a}}\bar{H} - \bar{a}Z_{j,\bar{a},\dot{\phi}}\bar{H} \\
&\qquad + 2\Big(\dddot{\phi}Z_{j,\bar{a},\ddot{\phi},\ddot{\phi}} + \ddot{\phi}Z_{j,\bar{a},\dot{\phi},\ddot{\phi}} + \dot{\phi}Z_{j,\bar{a},\phi,\ddot{\phi}}\Big)\bar{H} - (3-j)\,Z_{j,\dot{\phi}}\bar{H} \\
&\qquad + 2\,(3-j)\left(\dddot{\phi}Z_{j,\dot{\phi},\ddot{\phi}} + \ddot{\phi}Z_{j,\dot{\phi},\ddot{\phi}} + \dot{\phi}Z_{j,\phi,\ddot{\phi}}\right)\bar{H} \\
&\qquad + (j+1)\,(3-j)\,Z_{j,\ddot{\phi}}\left(1-\bar{H}\right)\frac{\ddot{\bar{a}}}{\bar{a}} + \ddddot{\phi}\,Z_{j,\ddot{\phi},\ddot{\phi}} + \dddot{\phi}^2 Z_{j,\ddot{\phi},\ddot{\phi},\ddot{\phi}} \\
&\qquad + \ddot{\phi}^2 Z_{j,\dot{\phi},\dot{\phi},\ddot{\phi}} + 2\Big(\dot{\phi}\dddot{\phi}Z_{j,\phi,\ddot{\phi},\ddot{\phi}} + \ddot{\phi}\dddot{\phi}\,Z_{j,\dot{\phi},\ddot{\phi},\ddot{\phi}} + \dot{\phi}\ddot{\phi}Z_{j,\phi,\dot{\phi},\ddot{\phi}}\Big) \\
&\qquad + \dot{\phi}^2 Z_{j,\phi,\phi,\ddot{\phi}} + \ddot{\phi}Z_{j,\bar{a},\phi,\ddot{\phi}} + \dot{\phi}Z_{j,\phi,\dot{\phi}} - \ddot{\phi}Z_{j,\dot{\phi},\dot{\phi}} + Z_{j,\phi} \\
&\qquad + \bar{a}\Big(\,(2j+1)\,Z_{j,\bar{a},\ddot{\phi}} - jZ_{j,\ddot{\phi}}\Big)\frac{\ddot{\bar{a}}}{\bar{a}} + jZ_{j,\ddot{\phi}}\frac{\dddot{\bar{a}}}{\bar{a}}\bar{H}^{-1} \\
&\qquad + 2j\left(\dot{\phi}Z_{j,\phi,\ddot{\phi}} + \ddot{\phi}Z_{j,\dot{\phi},\ddot{\phi}} + \dddot{\phi}\,Z_{j,\ddot{\phi},\ddot{\phi}}\right)\frac{\ddot{\bar{a}}}{\bar{a}}\bar{H}^{-1} - jZ_{j,\dot{\phi}}\frac{\ddot{\bar{a}}}{\bar{a}}\bar{H}^{-1}\Bigg]\bar{H}^{j}\,. \qquad (4.2.22)
\end{aligned}$$

The Jordan frame field equations for our disformally coupled Horndeski theory are thus given by

$$\mathscr{H} \;=\; -\rho_{\mathrm{M}}\,, \qquad \mathcal{E}^{\phi} \;=\; 0\,, \qquad \dot{\rho}_{\mathrm{M}} + 3\bar{H}\left(\rho_{\mathrm{M}} + p\right) \;=\; 0\,. \qquad (4.2.23)$$

4.2.2 Sieving out the self-tuning solutions

So far, we have only applied minimal constraints to our disformally coupled Horndeski theory; in order to determine if it admits self-tuning solutions, we need to apply the self tuning filter, **S.1 - S.3**, and establish the restrictions placed on the theory. To do so, we shall continue to work in the Jordan frame, in which the spacetime geometry is described by $\bar{g}_{\mu\nu}$. Since $\bar{g}_{\mu\nu}$ is what matter couples to, we wish to determine how the theory must self-tune in order for the effects of the cosmological constant to be screened from this metric.

Our starting point is to apply the self-tuning filter in the case where our cosmological background is in vacuo and consider the implications following from this scenario. Given this, the matter sector is expected to contribute a constant vacuum energy density that we identify with the cosmological constant, $\Lambda = \langle\rho_{\rm M}\rangle_{\rm vac}$ (as is the case in the derivation of the Fab-Four [60]). According to the first filter **S.1**, the vacuum energy density should have no effect on the spacetime curvature seen by matter, therefore we require a (portion of) Minkowski spacetime regardless of the value of Λ. It also follows from the second filter **S.2**, that this should remain true even in the case where the matter sector undergoes a phase transition, and in doing so, alters the net value of Λ by a constant amount (over an effectively infinitesimal time interval). This translates to requiring the scalar field ϕ to completely absorb any abrupt changes in the matter sector, leaving the geometry unaffected. Consequently, the scalar field tunes itself to each change in the vacuum energy density Λ, and this must be permitted independently of the time of transition.

Focussing our attention on the first filter **S.1**, we observe that in order to be consistent with it, our theory must admit cosmological vacuum solutions that are *Ricci flat*, i.e. $\bar{R}_{\mu\nu} = 0$. Note that, in the Jordan frame, the Ricci tensor is given by

$$\bar{R}^{\mu}_{\ \nu} \ = \ \mathrm{diag}\left(3\dot{\bar{H}}+3\bar{H}^2,\ \dot{\bar{H}}+3\bar{H}^2+2\frac{k}{\bar{a}^2},\ \dot{\bar{H}}+3\bar{H}^2+2\frac{k}{\bar{a}^2},\ \dot{\bar{H}}+3\bar{H}^2+2\frac{k}{\bar{a}^2}\right). \tag{4.2.24}$$

Thus, enforcing the constraint of Ricci flatness provides us with our so-called *on-shell-in-$\bar{a}$* condition

$$\bar{H}^2 \ = \ -\frac{k}{\bar{a}^2} \ = \ \left(\frac{s}{\bar{a}}\right)^2 , \tag{4.2.25}$$

where $k = 0$ and $k = -1$ correspond to a flat, and a Milne slicing of spacetime, respectively (for $k = 1$, we see that $-\frac{k}{\bar{a}^2} < 0$, and so no flat spacetime slicing is possible in this case). For brevity, we have also defined $s \ \coloneqq \ \sqrt{-k}$.

To proceed, we shall assume that ϕ is a *continuous* function, but that $\dot{\phi}$, $\ddot{\phi}$ and $\dddot{\phi}$ can be *discontinuous*. With this in mind, we then go *on-shell-in-$\bar{a}$* at the level of the field equations, imposing (4.2.25) by inserting $\bar{a} = \bar{a}_k = \bar{a}_0 + st$, whilst leaving ϕ to be determined dynamically. In doing so, we find that

$$\mathscr{H}(\bar{a}, \dot{\bar{a}}, \ddot{\bar{a}}, \phi, \dot{\phi}, \ddot{\phi}, \dddot{\phi}) \qquad \rightarrow \qquad \mathscr{H}_k(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi}, \dddot{\phi}) , \tag{4.2.26a}$$

$$\mathcal{E}^{\phi}(\bar{a}, \dot{\bar{a}}, \ddot{\bar{a}}, \dddot{\bar{a}}, \phi, \dot{\phi}, \ddot{\phi}, \dddot{\phi}, \ddddot{\phi}) \qquad \rightarrow \qquad \mathcal{E}^{\phi}_k(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi}, \dddot{\phi}, \ddddot{\phi}) , \tag{4.2.26b}$$

such that the *on-shell-in-$\bar{a}$* field equations are

$$\mathscr{H}_k = -\Lambda\,, \qquad\qquad \mathcal{E}^{\phi}_k = 0\,, \tag{4.2.27}$$

where, in accordance with the second filter, the matter sector contributes Λ to the vacuum energy density, where Λ is a piece-wise constant function of time. From here on in, a sub-(super) script k on a variable denotes that it is *on-shell-in-$\bar{a}$*. Further note that $\mathscr{H}_k$ and $\mathcal{E}^{\phi}_k$ contain no *explicit* time dependence.

From eq. (4.2.21), we see that the gravitational Hamiltonian is constructed from a set of functions $Z_j(\bar{a}, \phi, \dot{\phi}, \ddot{\phi})$ and their derivatives, with terms depending on $\dot{\phi}$, $\ddot{\phi}$ and $\dddot{\phi}$. As such, the requirement that it satisfy the condition given by eq. (4.2.27), imposes constraints on how $\dot{\phi}$, $\ddot{\phi}$ and $\dddot{\phi}$ appear in $\mathscr{H}$. Furthermore, from eq. (4.2.22), we see that the scalar EOM is similarly constructed from the functions Z_j; we can therefore use these restrictions on the forms of $\dot{\phi}$, $\ddot{\phi}$ and $\dddot{\phi}$ to similarly impose constraints on the functional form of $\mathcal{E}^{\phi}$. In particular, we note that since Λ is piece-wise continuous, there must be some discontinuity in $\mathscr{H}_k$ to account for this, and as $\bar{a}$ and ϕ are continuous, this means that $\mathscr{H}_k$ must (at least) have some non-trivial dependence on $\dot{\phi}$.

One might worry about the validity of imposing constraints on the higher-order derivatives of ϕ, as they arise from transforming between the Horndeski and Jordan frames. The point here, however, is that in the Jordan frame the matter sector is minimally coupled to metric, whereas in the Horndeski frame it is not. Were we to work in the Horndeski frame, complicated functions of the metric, and the scalar field (and its derivatives) would be present, and we would have to apply constraints to these instead. By working in the Jordan frame, the information contained in these complicated functions manifests as higher-order derivatives of the scalar fields. They therefore have a physical interpretation, rendering it legitimate to apply the self-tuning constraints to them.

Given these observations, and the form of $\mathscr{H}$, eq. (4.2.21), it is sufficient to consider the functional dependence of the functions, Z_j, on $\ddot{\phi}$ in order to determine the constraints we must impose on $\dot{\phi}$, $\ddot{\phi}$ and $\dddot{\phi}$, and subsequently, $\mathcal{E}^{\phi}$. Thus, we are left with three possible cases to consider:

1. $Z_j(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi})$ is *non-linear* in $\ddot{\phi}$;

2. $Z_j(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi})$ is *linear* in $\ddot{\phi}$;

3. $Z_j(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi})$ is *independent* of $\ddot{\phi}$.

Since we require $\mathscr{H}_k$ to possess a discontinuity, we can refer back to eq. (4.2.21) to derive the constraints this imposes on the forms of $\ddot{\phi}$ and $\dddot{\phi}$. We note that at a phase transition, localised at some time $t = t_*$, the right-hand side of the Hamiltonian constraint $\mathscr{H}_k = -\Lambda$ is discontinuous, and therefore, there must be support for this on the left-hand side. Furthermore, the support for this discontinuity must be provided by the highest derivative of ϕ in each case; if it were not, then terms proportional to a δ-function would appear on the left-hand side of eq. (4.2.27), which is not consistent with the right-hand side. Thus, proceeding in a case-by-case fashion, we find that:

(Ia) If $Z_j(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi})$ is *non-linear* in $\ddot{\phi}$, then $Z_{j,\ddot{\phi},\ddot{\phi}} \neq 0$, and thus $\dddot{\phi} \propto \theta(t - t_*)$ (where $\theta(t)$ is the Heaviside step-function), implying that $\ddddot{\phi} \propto \delta(t - t_*)$.

(Ib) Alternatively, if $Z_j(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi})$ is *linear* in $\ddot{\phi}$, then $Z_{j,\ddot{\phi},\ddot{\phi}} = 0$, and so it follows that $\ddot{\phi} \propto \theta(t - t_*)$, such that $\dddot{\phi} \propto \delta(t - t_*)$. Note that this condition automatically implies that $\mathscr{H}_k$ is *linear* in $\ddot{\phi}$.

(Ic) Finally, if $Z_j(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi})$ is *independent* of $\ddot{\phi}$, then all derivatives of Z_j with respect to $\ddot{\phi}$ vanish. In this case, it follows that $\dot{\phi} \propto \theta(t-t_*)$, and so $\ddot{\phi} \propto \delta(t-t_*)$. Moreover, it is evident from eq. (4.2.21) that $\mathscr{H}_k$ is independent of $\ddot{\phi}$ and $\dddot{\phi}$.

The implications of this analysis can then be applied to the on-shell EOM, $\mathcal{E}^\phi_k$ for ϕ [eq. (4.2.22)], placing constraints on the (4.2.27) forms of the functions Z_j:

(IIa) If $Z_j(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi})$ is *non-linear* in $\ddot{\phi}$, then $Z_{j,\ddot{\phi},\ddot{\phi}} \neq 0$ and accordingly, $\mathcal{E}^\phi_k$ is (at most) *linear* in $\ddddot{\phi}$. However, we know that in this case $\ddddot{\phi} \propto \delta(t - t_*)$, but we require that $\mathcal{E}^\phi_k = 0$. There is no support for a δ-function on the right-hand side of this equation and so we must conclude that, actually, $Z_{j,\ddot{\phi},\ddot{\phi}} = 0$.

(IIb) If instead $Z_j(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi})$ is *linear* in $\ddot{\phi}$, then clearly $Z_{j,\ddot{\phi},\ddot{\phi}} = 0$ and it follows that $\mathcal{E}^\phi_k$ will be (at most) *linear* in $\ddot{\phi}$. This can be seen by referring back to eq. (4.2.22), and noting that $Z_{j,\dot{\phi},\dot{\phi}} \sim \alpha(\bar{a}, \phi, \dot{\phi})\ddot{\phi}$ (since Z_j is linear in $\ddot{\phi}$), hence $Z_{j,\dot{\phi},\dot{\phi},\ddot{\phi}} \sim \alpha(\bar{a}, \phi, \dot{\phi})$, implying that $\ddot{\phi}^2 Z_{j,\dot{\phi},\dot{\phi},\ddot{\phi}} \sim \alpha(\bar{a}, \phi, \dot{\phi})\ddot{\phi}^2$. We see, therefore, that this term will cancel with the term $-\ddot{\phi} Z_{j,\dot{\phi},\dot{\phi}}$, and so any non-linear terms in $\ddot{\phi}$ appearing in $\mathcal{E}^\phi_k$ cancel out.

(IIc) In the final case, if $Z_j(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi})$ is *independent* of $\ddot{\phi}$, then it is immediately clear that $\mathcal{E}^{\phi}_k$ is (at most) *linear* in $\ddot{\phi}$.

Strictly speaking, these arguments only apply in the neighbourhood of the transition time $t = t_*$, however, the transition (or transitions) can occur at *any* time, so we can extend these results to include *all* times. It is evident then, that in each case Z_j can be (at most) *linear* in $\ddot{\phi}$, and consequently, the *on-shell-in-$\bar{a}$* Lagrangian, $\mathcal{L}_k$, will be also. This suggests that $\mathcal{L}_k$ should be of the form

$$\mathcal{L}_k \;=\; \alpha(\bar{a}_k, \phi, \dot{\phi}) + \ddot{\phi}\,\beta_{,\dot{\phi}}(\bar{a}_k, \phi, \dot{\phi})\,, \tag{4.2.28}$$

where the derivative of β with respect to $\dot{\phi}$ has been introduced for later convenience (α and β are arbitrary at this point and so we are always free to arrange things such that this is the case). Note that $\dot{\bar{a}} = \text{const.}$ when *on-shell-in-$\bar{a}$* and so, whilst α and β can be dependent on $\bar{a}$, they will not be dependent on any of its derivatives.

From eq. (4.2.17), it is clear that $\mathcal{L}$ only depends on the form of the functions Z_j (since the form of $\bar{H}$ is fixed as $\bar{H} = \frac{\dot{\bar{a}}}{\bar{a}}$). We can thus deduce that in cases 2 and 3, the *on-shell-in-$\bar{a}$* Lagrangian can always be cast into the form given by eq. (4.2.28); case 1 is less obvious. To prove that $\mathcal{L}_k$ can be expressed in the form given by eq. (4.2.28) even in the case where, $Z_j(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi})$ is *non-linear* in $\ddot{\phi}$, we refer to the *on-shell-in-$\bar{a}$* scalar EOM, $\mathcal{E}^{\phi}_k = 0$, and concentrate on the terms proportional to $\ddddot{\phi}$, i.e. $\bar{a}^3 \sum_{j=0}^{3} \ddddot{\phi}\, Z_{j,\ddot{\phi},\ddot{\phi}} \bar{H}^j_k \subset \mathcal{E}^{\phi}_k$. Now, we know that in case 1, $\ddddot{\phi} \propto \delta(t - t_*)$, however, there is no support for a δ-function on the right-hand side of $\mathcal{E}^{\phi}_k = 0$, and so it must be the case that, in actual fact, $Z_{j,\ddot{\phi},\ddot{\phi}}(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi}) = 0$, i.e. Z_j cannot be non-linear in $\ddot{\phi}$ (for all j). Therefore, we see that in all three cases, the *on-shell-in-$\bar{a}$* Lagrangian can be expressed in the form given by eq. (4.2.28).

From our ansatz for $\mathcal{L}_k$, eq. (4.2.28), we can subsequently ascertain an expression for the *on-shell-in-$\bar{a}$* scalar EOM, $\mathcal{E}^{\phi}_k$

$$\mathcal{E}^{\phi}_k \;=\; \ddot{\phi} f_k(\bar{a}_k, \phi, \dot{\phi}) + g_k(\bar{a}_k, \phi, \dot{\phi}) \;=\; 0\,, \tag{4.2.29}$$

where

$$f_k(\bar{a}_k, \phi, \dot{\phi}) \;=\; -\alpha_{,\dot{\phi},\dot{\phi}} + 2\beta_{,\phi,\dot{\phi}} + \dot{\phi}\beta_{,\phi,\dot{\phi},\dot{\phi}} + s\beta_{,\bar{a},\dot{\phi},\dot{\phi}}\,, \tag{4.2.30a}$$

$$g_k(\bar{a}_k, \phi, \dot{\phi}) \;=\; -\alpha_{,\phi} - \dot{\phi}\alpha_{,\phi,\dot{\phi}} - s\alpha_{,\bar{a},\dot{\phi}} + 2s\dot{\phi}\beta_{,\bar{a},\phi,\dot{\phi}} + \dot{\phi}^2\beta_{,\phi,\phi,\dot{\phi}} + s^2\beta_{,\bar{a},\bar{a},\dot{\phi}}\,, \tag{4.2.30b}$$

and we note that there is no *explicit* time dependence contained in either f_k or g_k.

Given these expressions for f_k and g_k, we can work through each case, 1 - 3, applying the ensuing constraints that they must satisfy. Since the latter two cases are less computationally heavy we shall start with those first, working backwards towards case 1.

We first look at case 3, in which $Z_j(\bar{a}_k, \phi, \dot{\phi}, \ddot{\phi})$ is independent of $\ddot{\phi}$. We now know that this implies $\ddot{\phi} \propto \delta(t - t_*)$, however, there is no support for a δ-function on the right-hand side of the on-shell field equation[6], eq. (4.2.29), and therefore, we are forced to conclude

$$f_k(\bar{a}_k, \phi, \dot{\phi}) \;=\; 0\,, \qquad \Rightarrow \qquad g_k(\bar{a}_k, \phi, \dot{\phi}) \;=\; 0\,. \tag{4.2.31}$$

If we then take the time derivative of these constraints,

$$\frac{\mathrm{d}f_k}{\mathrm{d}t} \;=\; \ddot{\phi}f_{k,\dot{\phi}} + \dot{\phi}f_{k,\phi} + sf_{k,\bar{a}} \;=\; 0\,, \qquad \frac{\mathrm{d}g_k}{\mathrm{d}t} \;=\; \ddot{\phi}g_{k,\dot{\phi}} + \dot{\phi}g_{k,\phi} + sg_{k,\bar{a}} \;=\; 0\,, \tag{4.2.32}$$

(where we have noted that $\dot{\bar{a}}_k = s$ *on-shell-in-$\bar{a}$*) we can use the same argument to arrive at the conclusion that $f_{k,\dot{\phi}} = 0 = g_{k,\dot{\phi}}$, i.e. both f_k and g_k are independent of $\dot{\phi}$. Moreover, referring back to eq. (4.2.32), with $f_{k,\dot{\phi}} = 0 = g_{k,\dot{\phi}}$, we see that the right-hand side contains a discontinuity, due to the presence of $\dot{\phi}$, however, this is inconsistent with the left-hand side (of which there is no discontinuity), and hence we must further have that $f_{k,\phi} = 0 = g_{k,\phi}$. Thus, both *on-shell-in-$\bar{a}$* functions, f_k and g_k, depend on (at most) only $\bar{a}$, i.e.

$$f_k \;=\; f_k(\bar{a}_k)\,, \qquad\qquad g_k \;=\; g_k(\bar{a}_k)\,. \tag{4.2.33}$$

It can be readily seen, through following the same set of steps (up to taking an additional time derivative), that this result also holds for case 2, in which Z_j is linear

[6] Note that this argument relies on the fact that there is no explicit time dependence in f_k and g_k, and so there is nothing to absorb the discontinuity in $\ddot{\phi}$. The same is true for $\dot{\phi}$.

in $\ddot{\phi}$. Case 1 , however, requires a little more care due to the non-linear dependence of Z_j in $\ddot{\phi}$. We first recall that in this case, $\dddot{\phi} \propto \theta(t-t_*)$ implying that $\ddddot{\phi} \propto \delta(t-t_*)$. Then, taking the time derivative of the *on-shell-in-$\bar{a}$* field equation [eq. (4.2.29)], we have

$$\dddot{\phi} f_k + \ddot{\phi}^2 f_{k,\dot{\phi}} + \dot{\phi}\ddot{\phi} f_{k,\phi} + s\ddot{\phi} f_{k,\bar{a}} + \ddot{\phi} g_{k,\dot{\phi}} + \dot{\phi} g_{k,\phi} + s g_{k,\bar{a}} \;=\; 0\,, \tag{4.2.34}$$

and thus $f_k = 0$, due to the unsupported discontinuity in $\dddot{\phi}$ on the left-hand side, further implying (by inserting $f_k = 0$ back into eq. (4.2.29)) that $g_k = 0$. Differentiating $f_k = 0$ twice with respect to t,

$$\dddot{\phi} f_{k,\dot{\phi}} + \ddot{\phi}^2 f_{k,\dot{\phi},\dot{\phi}} + 2\dot{\phi}\ddot{\phi} f_{k,\phi,\dot{\phi}} + 2s\ddot{\phi} f_{k,\bar{a},\dot{\phi}} + \ddot{\phi} f_{k,\phi} + \dot{\phi}^2 f_{k,\phi,\phi} + 2s\dot{\phi} f_{k,\bar{a},\phi} + s^2 f_{k,\bar{a},\bar{a}} \;=\; 0\,, \tag{4.2.35}$$

it is evident that $f_{k,\dot{\phi}} = 0$, i.e. f_k is independent of $\dot{\phi}$, due to the discontinuity in $\dddot{\phi}$ and the lack of support for this on the right-hand side of the equation. Returning to eq. (4.2.35), and taking a further time derivative, we find that the resulting equation has a term of the form $\dddot{\phi} f_{k,\phi}$. Again, the discontinuity in $\dddot{\phi}$ implies that $f_{k,\phi} = 0$, that is, f_k is independent of ϕ and hence depends only on $\bar{a}_k$. Following the same procedure, it can be shown that $g_{k,\dot{\phi}} = 0$ and $g_{k,\phi} = 0$ also, and so both f_k and g_k are dependent only on $\bar{a}_k$, $f_k = f_k(\bar{a}_k)$ and $g_k = g_k(\bar{a}_k)$. Since it is true that, in all three cases, 1 - 3, f_k and g_k are dependent only on $\bar{a}_k$, which is fixed, we see that the *on-shell-in-$\bar{a}$* scalar EOM $f_k(\bar{a}_k) = 0$ and $g_k(\bar{a}_k) = 0$ contain no dynamics. In other words, the scalar field equation E^k_ϕ places no further constraints on the evolution of ϕ, as it vanishes identically when *on-shell-in-$\bar{a}$*.

It is clear that the conditions placed on f_k and g_k inevitably impact on the admissable form of the *on-shell-in-$\bar{a}$* Lagrangian, eq. (4.2.28). Indeed, referring back to our expressions relating α and $\beta_{,\dot{\phi}}$ to f_k and g_k, Eqs. (4.2.30a) and (4.2.30b), and setting $f_k = 0$, we can integrate eq. (4.2.30a) twice with respect to $\dot{\phi}$ to obtain an expression for α

$$\alpha \;=\; \alpha(\bar{a}_k, \phi, \dot{\phi}) \;=\; \dot{\phi}\beta_{,\phi} + s\beta_{,\bar{a}} + \dot{\phi}\zeta_{,\phi} + \xi\,, \tag{4.2.36}$$

where $\zeta_{,\phi} = \zeta_{,\phi}(\bar{a}_k, \phi)$ and $\xi = \xi(\bar{a}_k, \phi)$ are constants of integration (with respect to $\dot{\phi}$). We can then insert this result back into eq. (4.2.30b), and upon setting it to zero, we obtain a differential equation relating $\zeta_{,\phi}$ and ξ, of the form $\xi_{,\phi} = s\zeta_{,\phi,\bar{a}}$, which can trivially integrate to give $\xi = \xi(\bar{a}_k, \phi) = s\zeta_{,\bar{a}}$, where we have absorbed the constant of integration through a redefinition of ζ. Therefore, upon substituting this result back into eq. (4.2.36), we find that

$$\alpha(\bar{a}_k, \phi, \dot{\phi}) \;=\; \dot{\phi}\beta_{,\phi} + s\beta_{,\bar{a}} + \dot{\zeta}\,. \tag{4.2.37}$$

Finally, inserting this into our *on-shell-in-$\bar{a}$* Lagrangian, eq. (4.2.28), we arrive at the following result

$$\mathcal{L}_k \;=\; \alpha + \ddot{\phi}\beta_{,\dot{\phi}} \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\big(\beta + \zeta\big)\,. \tag{4.2.38}$$

That is, the *on-shell-in-$\bar{a}$* Lagrangian is simply a total derivative.

It remains for us to apply the third self-tuning filter, **S.3**, which requires that our self-tuning theory admit a non-trivial cosmology. We can determine the criterion for this to occur by referring back to the scalar EOM [eq. (4.2.22)] before going *on-shell-in-$\bar{a}$*. For the theory to admit self-tuning solutions, the first filter requires Ricci flatness, which we now know requires $\mathcal{E}^{\phi}$ to vanish identically (i.e. it is satisfied trivially) when *on-shell-in-$\bar{a}$*. This can only happen in either one of two ways: (1) either $\mathcal{E}^{\phi} = 0$ is an algebraic equation in $\bar{H} - \frac{s}{\bar{a}}$, i.e. of the form $\sum_n \chi_n \left(\bar{H} - \frac{s}{\bar{a}}\right)^n = 0$ (where χ_n are arbitrary coefficients); or (2) $\mathcal{E}^{\phi} = 0$ is a dynamical equation in $\bar{H} - \frac{s}{\bar{a}}$. Now, if it is the case that option (1) is satisfied, then it is clear that only a trivial cosmology is ever permitted - the scalar EOM enforces Minkowksi spacetime for all time, in direct conflict with the third filter, **S.3**. We therefore adopt option (2), which requires that the scalar EOM contains derivatives of $\bar{H} - \frac{s}{\bar{a}}$, implying that it must be non-trivially dependent on $\ddot{\bar{a}}$ and $\dddot{\bar{a}}$. In other words, the theory must satisfy the field equations (4.2.23) in such a way that it evolves asympotically towards a Ricci flat solution, i.e. its late time solution to the EOM is *on-shell-in-$\bar{a}$*.

Let us briefly return to the full minisuperspace Lagrangian [eq. (4.2.17)], and pass it through the self-tuning filter, taking into account the results we have obtained above:

(IIIa) the *on-shell-in-$\bar{a}$* Lagrangian must be zero (up to a total derivative);

(IIIb) the *on-shell-in-$\bar{a}$* Hamiltonian must *not* be independent of $\dot{\phi}$;

(IIIc) the full scalar EOM must *not* be independent of $\ddot{\bar{a}}$.

In doing so, we can infer the following set of constraints:

(IVa) $\bar{a}^3 \sum_{j=0}^{3} Z_j \left(\frac{s}{\bar{a}}\right)^j \Big|_{\bar{a}=\bar{a}_k} = \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{G}(\bar{a},\phi,\dot{\phi})\Big|_{\bar{a}=\bar{a}_k} \cong 0$;

(IVb) $\bar{a}^3 \sum_{j=1}^{3} \Big[jZ_{j,\dot{\phi}} + (j-1)\big(Z_{j,\ddot{\phi}} + \dot{\phi} Z_{j,\dot{\phi},\ddot{\phi}}\big)\Big] \left(\frac{s}{\bar{a}}\right)^j \Big|_{\bar{a}=\bar{a}_k} \neq 0$;

(IVc) *Cannot* simultaneously have

$$3Z_{0,\ddot{\phi}} - Z_{1,\dot{\phi}} + \bar{a} Z_{0,\bar{a},\ddot{\phi}} + 2\left(\dot{\phi} Z_{1,\phi,\ddot{\phi}} + \ddot{\phi} Z_{1,\dot{\phi},\ddot{\phi}} + \dddot{\phi} Z_{1,\ddot{\phi},\ddot{\phi}}\right) = 0\,,$$

$$4Z_{1,\ddot{\phi}} - 3Z_{0,\ddot{\phi}} - 2Z_{2,\dot{\phi}} + \bar{a}\left(3Z_{1,\bar{a},\ddot{\phi}} - Z_{1,\ddot{\phi}}\right) + 4\left(\dot{\phi} Z_{2,\phi,\ddot{\phi}} + \ddot{\phi} Z_{2,\dot{\phi},\ddot{\phi}} + \dddot{\phi} Z_{2,\ddot{\phi},\ddot{\phi}}\right) = 0\,,$$

$$3Z_{2,\ddot{\phi}} - Z_{1,\ddot{\phi}} - 3Z_{3,\dot{\phi}} + \bar{a}\left(5Z_{2,\bar{a},\ddot{\phi}} - 2Z_{2,\ddot{\phi}}\right) + 6\left(\dot{\phi} Z_{3,\phi,\ddot{\phi}} + \ddot{\phi} Z_{3,\dot{\phi},\ddot{\phi}} + \dddot{\phi} Z_{3,\ddot{\phi},\ddot{\phi}}\right) = 0\,,$$

$$3Z_{2,\ddot{\phi}} + \bar{a}\left(7Z_{3,\bar{a},\ddot{\phi}} - 3Z_{3,\ddot{\phi}}\right) = 0\,,$$

$$Z_{3,\ddot{\phi}} = 0\,.$$

Note that (IVa) implies that $\sum_{j=0}^{3} Z_{j,\dot{\phi}} \left(\frac{s}{\bar{a}}\right)^j \Big|_{\bar{a}=\bar{a}_k} = 0 = \sum_{j=0}^{3} Z_{j,\ddot{\phi}} \left(\frac{s}{\bar{a}}\right)^j \Big|_{\bar{a}=\bar{a}_k}$, which we have made use of to simplify (IVb) (we have also used that $\ddot{\bar{a}}_k = 0$). In particular, (IVb) rules out $k = 0$. This is an important result, as it shows that self-tuning is not possible within this class of scalar-tensor theories for a homogeneous scalar field and a spatial flat cosmology. Nevertheless, as was the case for the original Fab-Four, there is no obvious obstruction to self-tuning with a homogeneous scalar and a spatially hyperbolic cosmology ($k = -1$).

4.2.3 Constructing a self-tuning theory

Having determined the required features of our theory in order for it to pass through the self-tuning filters, we are now at the point in which we can construct our putative self-tuning Lagrangian. To this end, we note that we are working within an equivalence class of Lagrangians, $[\mathcal{L}, \cong]$, where two Lagrangians are equivalent if and only if they differ by a total derivative

$$\tilde{\mathcal{L}} \cong \mathcal{L} \qquad \Longleftrightarrow \qquad \tilde{\mathcal{L}} = \mathcal{L} + \frac{\mathrm{d}\mathcal{G}}{\mathrm{d}t}\,. \tag{4.2.39}$$

This is a consequence of the fact that any two such Lagrangians describe the same dynamical theory, i.e. they lead to identical EOM. Then, given that $\bar{H} = \frac{s}{\bar{a}}$ when *on-shell-in-*$\bar{a}$, and recalling that $\mathcal{L}_k$ must be equal to a total derivative, we have from

Eqs. (4.2.17) and (4.2.38), that

$$\mathcal{L}_k \;=\; \bar{a}^3 \sum_{j\,=\,0}^{3} Z_j(\bar{a},\phi,\dot{\phi},\ddot{\phi}) \left(\frac{s}{\bar{a}}\right)^j \bigg|_{\bar{a}\,=\,\bar{a}_k} \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{G}(\bar{a},\phi,\dot{\phi})\bigg|_{\bar{a}\,=\,\bar{a}_k} \cong 0\,. \qquad (4.2.40)$$

Given this analysis, let us consider a Horndeski-like theory of the form

$$\begin{aligned} \tilde{\mathcal{L}} \;&=\; \bar{a}^3 \sum_{j\,=\,0}^{3} \tilde{Z}_j \bar{H}^j \;=\; \bar{a}^3 \sum_{j\,=\,0}^{3} \tilde{Z}_j \left(\frac{s}{\bar{a}}\right)^j + \bar{a}^3 \sum_{j\,=\,1}^{3} \tilde{Z}_j \left[\bar{H}^j - \left(\frac{s}{\bar{a}}\right)^j\right] \\ &\cong\; \bar{a}^3 \sum_{j\,=\,1}^{3} \tilde{Z}_j \left[\bar{H}^j - \left(\frac{s}{\bar{a}}\right)^j\right], \end{aligned} \qquad (4.2.41)$$

where $\tilde{Z}_j = \tilde{Z}_j(\bar{a},\phi,\dot{\phi},\ddot{\phi})$.

Such a theory certainly conforms to the self-tuning constraints laid out in this section, passing through the self-tuning filter, **S.1-S.3**. It is, in this sense, *sufficient* for self-tuning, but to what extent is it *necessary*? Indeed, a priori, it is certainly not necessary that the theory takes the form of eq. (4.2.41), as there could be other equivalent Lagrangians, with $Z_j = \tilde{Z}_j + \Delta Z_j$, that admit the same set of self-tuning solutions. To establish whether this is the case, we need to demand that the "tilded" and "untilded" systems each have EOM that give rise to the same dynamics. That is, we require that when on-shell,

$$\mathscr{H} \;=\; -\rho_{\mathrm{M}}\,, \quad \mathcal{E}^\phi \;=\; 0 \qquad \Longleftrightarrow \qquad \tilde{\mathscr{H}} \;=\; -\rho_{\mathrm{M}}\,, \quad \tilde{\mathcal{E}}^\phi \;=\; 0\,. \qquad (4.2.42)$$

In general, we cannot imply from this statement that $\tilde{\mathcal{E}}^\phi \equiv \mathcal{E}^\phi$, nor even that $\tilde{\mathcal{E}}^\phi \propto \mathcal{E}^\phi$, as the relevant equations could well be related to one another non-linearly. Notwithstanding these observations, a detailed analysis reveals that the self-tuning limit of the Horndeski Lagrangian enforces the conclusion that

$$\mathscr{H} \;=\; \tilde{\mathscr{H}}\,, \qquad \mathcal{E}^\phi \;=\; \tilde{\mathcal{E}}^\phi\,. \qquad (4.2.43)$$

In other words, our putative self-tuning Lagrangian $\tilde{\mathcal{L}}$, given by eq. (4.2.41), furnishes a general self-tuning theory (up to a total derivative). To see that this is the case, we consider the following.

In principle, the Hamiltonian's of two different self-tuning Horndeski theories (defined by Eqs. (4.2.17) and (4.2.41)), $\mathscr{H}$ and $\tilde{\mathscr{H}}$, differ by a function $\Delta\mathscr{H} = \Delta\mathscr{H}(\bar{a}, \dot{\bar{a}}, \ddot{\bar{a}}, \phi, \dot{\phi}, \ddot{\phi}, \dddot{\phi})$, such that

$$\mathscr{H} + \rho_{\rm M} \equiv \tilde{\mathscr{H}} + \rho_{\rm M} + \Delta\mathscr{H}\,, \tag{4.2.44}$$

where the functional dependence of $\Delta\mathscr{H}$ arises from the fact that matter couples the same way in both theories (by assumption)[7]. If we now go on-shell, such that $\mathscr{H} + \rho_{\rm M} = 0 = \tilde{\mathscr{H}} + \rho_{\rm M}$ and $\mathcal{E}^\phi = 0 = \tilde{\mathcal{E}}^\phi$, then it follows from eq. (4.2.44) that $\Delta\mathscr{H} = 0$. Since $\Delta\mathscr{H}$ does not depend on $\rho_{\rm M}$, it cannot vanish by virtue of the on-shell equation $\tilde{\mathscr{H}} = -\rho_{\rm M}$. Can it instead vanish by virtue of $\tilde{\mathcal{E}}^\phi = 0$? Let us consider cases 1 and 2, where $\tilde{Z}_i$ depends *non-linearly* and *linearly* on $\ddot{\phi}$, respectively. From eq. (4.2.22), it is clear that $\tilde{\mathcal{E}}^\phi$ contains a term proportional to $\dddot{\bar{a}}$, however $\tilde{\mathscr{H}}$ does not, and so we cannot use $\tilde{\mathcal{E}}^\phi = 0$ to enforce $\Delta\mathscr{H} = 0$ (as in both cases there would remain non-trivial terms with no corresponding terms to cancel with). In case 3, in which $\tilde{Z}_i$ is independent of $\ddot{\phi}$, we see that $\tilde{\mathcal{E}}^\phi$ contains $\ddot{\bar{a}}$, but $\tilde{\mathscr{H}}$ does not, and as such, we cannot use $\tilde{\mathcal{E}}^\phi = 0$ to enforce $\Delta\mathscr{H} = 0$ in this case either. Therefore, we are unable to use the on-shell equations, $\tilde{\mathscr{H}} = -\rho_{\rm M}$ and $\tilde{\mathcal{E}}^\phi = 0$, to dynamically enforce $\Delta\mathscr{H} = 0$, we must conclude that $\Delta\mathscr{H}$ is identically zero. In other words,

$$\mathscr{H} \;=\; \tilde{\mathscr{H}}\,. \tag{4.2.45}$$

This constraint provides us with a useful relation. Indeed, given that $\Delta Z_j = Z_j - \tilde{Z}_j$, we have that

$$\begin{aligned}
\Delta\mathscr{H} \;&=\; \bar{a}^3 \sum_{j\,=\,0}^{3} \Bigg[(j-1)\,\Delta Z_j + \dot{\phi}\Delta Z_{j,\dot{\phi}} - \dot{\phi}^2 \Delta Z_{j,\phi,\ddot{\phi}} - \dot{\phi}\ddot{\phi}\Delta Z_{j,\dot{\phi},\ddot{\phi}} + \ddot{\phi}\Delta Z_{j,\ddot{\phi}} \\
&\qquad - \dot{\phi}\dddot{\phi}\Delta Z_{j,\ddot{\phi},\ddot{\phi}} + \Big((j-3)\,\dot{\phi}\Delta Z_{j,\ddot{\phi}} - \bar{a}\dot{\phi}\Delta Z_{j,\bar{a},\ddot{\phi}} \Big)\bar{H} - j\dot{\phi}\frac{\ddot{\bar{a}}}{\bar{a}}\Delta Z_{j,\ddot{\phi}}\bar{H}^{-1} \Bigg] \bar{H}^j \\
&=\; 0\,.
\end{aligned} \tag{4.2.46}$$

We first observe that there is only one term proportional to $\dddot{\phi}$ for each power of $\bar{H}$, and thus it cannot cancel with any other terms. Hence, it must vanish identically, i.e.

$$\Delta Z_{j,\ddot{\phi},\ddot{\phi}} \;=\; 0 \;\;\forall\, j\,. \tag{4.2.47}$$

[7] As we are working in the Jordan frame, this means that matter does not directly couple to the scalar field ϕ, and as such, neither $\mathscr{H}$ ($\tilde{\mathscr{H}}$) nor $\Delta\mathscr{H}$ can depend on $\rho_{\rm M}$.

That is, $\Delta Z_{j,\ddot{\phi},\ddot{\phi}}$ can be at most *linear* in $\ddot{\phi}$, regardless of the dependence $\tilde{Z}_j$ has on $\ddot{\phi}$. Moreover, there is also only one term proportional to $\ddot{a}$ for each power in $\bar{H}$, and therefore, by the same reasoning it must vanish identically,

$$j\dot{\phi}\frac{\ddot{\bar{a}}}{\bar{a}}\Delta Z_{j,\ddot{\phi}} \;=\; 0 \qquad \Longrightarrow \qquad \Delta Z_{j,\ddot{\phi}} \;=\; 0 \quad \text{for } j \;=\; 1,2,3\,, \tag{4.2.48}$$

and this must hold for whatever dependence $\tilde{Z}_j$ has on $\ddot{\phi}$. Referring back to eq. (4.2.46), and equating the remaining powers in $\bar{H}$ (noting that any terms containing derivatives of ΔZ_j must vanish on account of eq. (4.2.48)) , we find that

$$(j-1)\Delta Z_j + \dot{\phi}\Delta Z_{j,\dot{\phi}} \;=\; 0 \quad \text{for } j \;=\; 1,2,3\,, \tag{4.2.49}$$

leaving us with a first-order differential equation for ΔZ_j, which at most, must be a function of $\bar{a}$, ϕ and $\dot{\phi}$, i.e. $\Delta Z_j = \Delta Z_j(\bar{a},\phi,\dot{\phi})$ by virtue of the constraint given by eq. (4.2.48). Upon integration of eq. (4.2.49) with respect to $\dot{\phi}$, we have

$$\Delta Z_j(\bar{a},\phi,\dot{\phi}) \;=\; \sigma_j(\bar{a},\phi)\dot{\phi}^{1-j} \quad \text{for } j \;=\; 1,2,3\,, \tag{4.2.50}$$

where $\sigma_j(\bar{a},\phi)$ is an arbitary function of $\bar{a}$ and ϕ.

Note that eq. (4.2.50) is only the solution for ΔZ_j in the cases where $j \neq 0$. For $j = 0$ the situation is a little more complicated, since $j\dot{\phi}\frac{\ddot{\bar{a}}}{\bar{a}}\Delta Z_{j,\ddot{\phi}}$ vanishes by virtue of j. In this case, we refer back to the constraint given by eq. (4.2.47), which implies that

$$\Delta Z_0 \;=\; \mathcal{A}(\bar{a},\phi,\dot{\phi}) + \ddot{\phi}\mathcal{B}(\bar{a},\phi,\dot{\phi})\,. \tag{4.2.51}$$

With this information at our disposal, we now turn our attention to the scalar EOM. Analogously to the Hamiltonian's of the two different self-tuning Horndeski theories, the scalar equations of motion will differ by a function
$\Delta\mathcal{E}^\phi = \Delta\mathcal{E}^\phi(\bar{a},\dot{\bar{a}},\ddot{\bar{a}},\dddot{\bar{a}},\phi,\dot{\phi},\ddot{\phi},\dddot{\phi},\ddddot{\phi})$ such that

$$\mathcal{E}^\phi \;=\; \tilde{\mathcal{E}}^\phi + \Delta\mathcal{E}^\phi\,. \tag{4.2.52}$$

This implies that, on-shell, $\Delta\mathcal{E}^\phi = 0$. However, as in the Hamiltonian case above, $\Delta\mathcal{E}^\phi$ is independent of $\rho_{\rm M}$ and therefore cannot vanish by virtue of the on-shell equation $\tilde{\mathscr{H}} = -\rho_{\rm M}$. Therefore, at best, it can vanish by virtue of $\tilde{\mathcal{E}}^\phi = 0$. To proceed, we note that

$$\begin{aligned}\Delta\mathcal{E}^{\phi} &= \frac{\partial \Delta\mathcal{L}}{\partial\phi} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \Delta\mathcal{L}}{\partial\dot{\phi}}\right) + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(\frac{\partial \Delta\mathcal{L}}{\partial\ddot{\phi}}\right) \\ &= \sum_{j=0}^{3}\left[\bar{a}^3\Delta Z_{j,\phi}\bar{H}^j - \frac{\mathrm{d}}{\mathrm{d}t}\left(\bar{a}^3\Delta Z_{j,\dot{\phi}}\bar{H}^j\right)\right] + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(\bar{a}^3\Delta Z_{0,\ddot{\phi}}\right) . \end{aligned} \tag{4.2.53}$$

Upon calculating the time derivatives, we see then that $\Delta\mathcal{E}^{\phi}$ has the following form

$$\begin{aligned}\Delta\mathcal{E}^{\phi} &= \bar{a}^3\Big[\Delta Z_{0,\phi} + \frac{\ddot{\bar{a}}}{\bar{a}}\big(3\Delta Z_{0,\ddot{\phi}} + \bar{a}\Delta Z_{0,\bar{a},\ddot{\phi}}\big) - \dot{\phi}\big(\Delta Z_{0,\phi,\dot{\phi}} - \Delta Z_{0,\phi,\phi,\ddot{\phi}}\big) \\ &\qquad + \ddot{\phi}\big(\Delta Z_{0,\phi,\ddot{\phi}} - \Delta Z_{0,\dot{\phi},\dot{\phi}} + 2\dot{\phi}\Delta Z_{0,\phi,\dot{\phi},\ddot{\phi}} + \ddot{\phi}\Delta Z_{0,\dot{\phi},\dot{\phi},\ddot{\phi}}\big)\Big] \\ &\quad + \bar{a}^3\Big[\Delta Z_{1,\phi} - 3\Delta Z_{0,\dot{\phi}} - \bar{a}\big(\Delta Z_{0,\bar{a},\ddot{\phi}} - 3\Delta Z_{0,\bar{a},\ddot{\phi}}\big) - 2\frac{\ddot{\bar{a}}}{\bar{a}}\Delta Z_{2,\dot{\phi}} \\ &\qquad + 2\dot{\phi}\big(\bar{a}\Delta Z_{0,\bar{a},\phi,\ddot{\phi}} + 3\Delta Z_{0,\phi,\dot{\phi}}\big) + 2\ddot{\phi}\big(\bar{a}\Delta Z_{0,\bar{a},\dot{\phi},\ddot{\phi}} + 3\Delta Z_{0,\dot{\phi},\ddot{\phi}}\big)\Big]\bar{H} \\ &\quad + \bar{a}^3\Big[6\Delta Z_{0,\ddot{\phi}} + 3\bar{a}\Delta Z_{0,\bar{a},\ddot{\phi}} + \bar{a}^2\Delta Z_{0,\bar{a},\bar{a},\ddot{\phi}} + \Delta Z_{2,\phi} - 3\frac{\ddot{\bar{a}}}{\bar{a}}\Delta Z_{3,\dot{\phi}}\Big]\bar{H}^2 \\ &\quad - \bar{a}^3\Big[\Delta Z_{3,\phi} + \Delta Z_{2,\dot{\phi}} + \bar{a}\Delta Z_{3,\bar{a},\dot{\phi}} + \dot{\phi}\Delta Z_{3,\phi,\dot{\phi}} + \ddot{\phi}\Delta Z_{2,\dot{\phi},\dot{\phi}}\Big]\bar{H}^3 \\ &\quad - \bar{a}\Delta Z_{3,\bar{a},\dot{\phi}}\bar{H}^4 \end{aligned} \tag{4.2.54}$$

From this expression, it is clear that $\Delta\mathcal{E}^{\phi}$ is independent of $\dddot{\bar{a}}$ and $\ddddot{\phi}$, i.e
$\Delta\mathcal{E}^{\phi} = \Delta\mathcal{E}^{\phi}(\bar{a}, \dot{\bar{a}}, \ddot{\bar{a}}, \phi, \dot{\phi}, \ddot{\phi}, \dddot{\phi})$.

Following the same procedure as for the Hamiltonian, we again consider cases 1 and 2, where $\tilde{Z}_i$ depends *non-linearly* and *linearly* on $\ddot{\phi}$, respectively. Now, from eq. (4.2.22), we see that $\tilde{\mathcal{E}}^{\phi}$ contains a term proportional to $\dddot{\bar{a}}$, hence $\tilde{\mathcal{E}}^{\phi} = 0$ cannot be used to enforce $\Delta\mathcal{E}^{\phi} = 0$ (if we use it to substitute in for $\ddot{\bar{a}}$ in $\Delta\mathcal{E}^{\phi}$, there will always be non-trivial terms proportional to $\dddot{\bar{a}}$ remaining). It is evident from this, that in these two cases, $\Delta\mathcal{E}^{\phi}$ must vanish identically.

Now, collecting the terms proportional to $\ddot{\bar{a}}$ for each power in $\bar{H}$, we see that there are no other contributions to cancel with them, hence they must each vanish identically. As such, we find that

$$-2\Delta Z_{2,\dot{\phi}} = 2\sigma_2\dot{\phi}^{-2} = 0 \qquad \Rightarrow \qquad \sigma_2(\bar{a},\phi) = 0\,, \tag{4.2.55a}$$

$$-3\Delta Z_{3,\dot{\phi}} = 6\sigma_3\dot{\phi}^{-3} = 0 \qquad \Rightarrow \qquad \sigma_3(\bar{a},\phi) = 0\,, \tag{4.2.55b}$$

$$-\Delta Z_{1,\dot{\phi}} + 3\Delta Z_{0,\ddot{\phi}} + \bar{a}\Delta Z_{0,\bar{a},\ddot{\phi}} \;=\; 3\Delta Z_{0,\ddot{\phi}} + \bar{a}\Delta Z_{0,\bar{a},\ddot{\phi}} \;=\; 3\mathcal{B} + \bar{a}\mathcal{B}_{,\bar{a}} \;=\; 0$$

$$\Rightarrow \qquad \mathcal{B} \;=\; \bar{a}^{-3}\tilde{\mathcal{B}}(\phi,\dot{\phi})\,, \tag{4.2.55c}$$

where we have made use of eq. (4.2.50). Given this information, we proceed to equate the remaining powers in $\bar{H}$. Concentrating on the coefficients of $\bar{H}^2$, it follows that,

$$6\Delta Z_{0,\ddot{\phi}} + 3\bar{a}\Delta Z_{0,\bar{a},\ddot{\phi}} + \bar{a}^2\Delta Z_{0,\bar{a},\bar{a},\ddot{\phi}} \;=\; 0$$

$$\Rightarrow \qquad 6\mathcal{B} + 3\bar{a}\mathcal{B}_{,\bar{a}} + \bar{a}^2\mathcal{B}_{,\bar{a},\bar{a}} \;=\; 6\mathcal{B} - 9\mathcal{B} + 12\mathcal{B} \;=\; 9\mathcal{B} \;=\; 0$$

$$\Rightarrow \qquad \mathcal{B}(\bar{a},\phi,\dot{\phi}) \;=\; 0\,. \tag{4.2.56}$$

Therefore, in actual fact, ΔZ_0 is also independent of $\ddot{\phi}$, and as such eq. (4.2.50) holds for $j = 0$ as well.

Turning our attention to case 3, in which $\tilde{Z}_i$ is independent of $\ddot{\phi}$, it is apparent from eq. (4.2.22) that $\tilde{\mathcal{E}}^{\phi}$ no longer contains any terms proportional to $\dddot{a}$ and so we must be more careful with our analysis. To proceed, we note that for $\tilde{Z}_{j,\ddot{\phi}} = 0$, the scalar EOM has the form,

$$\mathcal{E}^{\phi} \;=\; \mathscr{A}(\bar{a},\dot{\bar{a}},\phi,\dot{\phi})\ddot{\bar{a}} + \mathscr{B}(\bar{a},\dot{\bar{a}},\phi,\dot{\phi})\ddot{\phi} + \mathscr{C}(\bar{a},\dot{\bar{a}},\phi,\dot{\phi})\,, \tag{4.2.57}$$

and similarly for $\tilde{\mathcal{E}}^{\phi}$ (with "tilded" functions $\tilde{\mathscr{A}}$, $\tilde{\mathscr{B}}$ and $\tilde{\mathscr{C}}$ replacing $\mathscr{A}$, $\mathscr{B}$ and $\mathscr{C}$).[8] It follows then, that

$$\ddot{\bar{a}} \;=\; \frac{1}{\tilde{\mathscr{A}}}\left[\tilde{\mathcal{E}}^{\phi} - \tilde{\mathscr{B}}\ddot{\phi} - \tilde{\mathscr{C}}\right]\,, \tag{4.2.58}$$

which leads us to the expression

$$\Delta\mathcal{E}^{\phi} \;=\; \mathcal{E}^{\phi} - \tilde{\mathcal{E}}^{\phi} \;=\; \Delta\mathscr{A}\ddot{\bar{a}} + \Delta\mathscr{B}\ddot{\phi} + \Delta\mathscr{C} \;=\; \frac{\Delta\mathscr{A}}{\tilde{\mathscr{A}}}\left[\tilde{\mathcal{E}}^{\phi} - \tilde{\mathscr{B}}\ddot{\phi} - \tilde{\mathscr{C}}\right] + \Delta\mathscr{B}\ddot{\phi} + \Delta\mathscr{C}$$

$$= \; \frac{\Delta\mathscr{A}}{\tilde{\mathscr{A}}}\tilde{\mathcal{E}}^{\phi} + \frac{\tilde{\mathscr{A}}\Delta\mathscr{B} - \tilde{\mathscr{B}}\Delta\mathscr{A}}{\tilde{\mathscr{A}}}\ddot{\phi} + \frac{\tilde{\mathscr{A}}\Delta\mathscr{C} - \tilde{\mathscr{C}}\Delta\mathscr{A}}{\tilde{\mathscr{A}}}\,, \tag{4.2.59}$$

where $\Delta\mathscr{A} = \mathscr{A} - \tilde{\mathscr{A}}$, and likewise for $\Delta\mathscr{B}$ and $\Delta\mathscr{C}$. Given that $\Delta\mathcal{E}^{\phi}$ ought to vanish by virtue of $\tilde{\mathcal{E}}^{\phi} = 0$, it is immediately clear that the following relationship must hold

$$\Delta\mathcal{E}^{\phi} \;=\; \frac{\Delta\mathscr{A}}{\tilde{\mathscr{A}}}\tilde{\mathcal{E}}^{\phi}\,, \qquad \tilde{\mathscr{A}}\Delta\mathscr{B} \;=\; \tilde{\mathscr{B}}\Delta\mathscr{A}\,, \qquad \tilde{\mathscr{A}}\Delta\mathscr{C} - \tilde{\mathscr{C}}\Delta\mathscr{A} \;=\; 0\,. \tag{4.2.60}$$

[8]Note that $\mathscr{A} \neq 0$ ($\tilde{\mathscr{A}} \neq 0$) due to the third self-tuning filter, **S.3**.

Moreover, upon comparison of eq. (4.2.59) with eq. (4.2.54), and making use of eq. (4.2.50), we can infer explicit expressions for $\Delta\mathscr{A}$, $\Delta\mathscr{B}$ and $\Delta\mathscr{C}$,

$$\Delta\mathscr{A} = -\bar{a}^3 \sum_{j=0}^{3} j\Delta Z_{j,\dot{\phi}} \frac{\bar{H}^{j-1}}{\bar{a}} = -\bar{a}^3 \sum_{j=0}^{3} j(1-j)\sigma_j \frac{\bar{H}^{j-1}}{\bar{a}\dot{\phi}^j}\,, \tag{4.2.61a}$$

$$\Delta\mathscr{B} = -\bar{a}^3 \sum_{j=0}^{3} \Delta Z_{j,\dot{\phi},\dot{\phi}} \bar{H}^j = \bar{a}^3 \sum_{j=0}^{3} j(1-j)\sigma_j \frac{\bar{H}^j}{\dot{\phi}^{j+1}} = -\frac{\bar{a}\bar{H}}{\dot{\phi}}\Delta\mathscr{A}\,, \tag{4.2.61b}$$

$$\begin{aligned}\Delta\mathscr{C} &= -\bar{a}^3 \sum_{j=0}^{3} \left[\dot{\phi}\Delta Z_{j,\phi,\dot{\phi}} + (3-j)\Delta Z_{j,\dot{\phi}}\bar{H} + \bar{a}\Delta Z_{j,\bar{a},\dot{\phi}}\bar{H} - \Delta Z_{j,\phi}\right]\bar{H}^j \\ &= -\bar{a}^3 \sum_{j=0}^{3} \left[\big((3-j)\sigma_j + \bar{a}\sigma_{j,\bar{a}}\big)(1-j)\bar{H} - j\dot{\phi}\sigma_{j,\phi}\right]\frac{\bar{H}^j}{\dot{\phi}^j}\,. \end{aligned}\tag{4.2.61c}$$

Given that we require $\tilde{\mathscr{A}}\Delta\mathscr{B} = \tilde{\mathscr{B}}\Delta\mathscr{A}$ in order for $\Delta\mathcal{E}^\phi$ to vanish by virtue of $\tilde{\mathcal{E}}^\phi = 0$ (otherwise $\Delta\mathcal{E}^\phi$ vanishes identically), it must be the case that $\bar{a}\tilde{\mathscr{A}}\bar{H} = -\dot{\phi}\tilde{\mathscr{B}}$, and this leads to the following solution for $\tilde{Z}_j$,

$$\begin{aligned} \bar{a}^3 \sum_{j=0}^{3} j\tilde{Z}_{j,\dot{\phi}}\bar{H}^j &= -\bar{a}^3 \sum_{j=0}^{3} \dot{\phi}\tilde{Z}_{j,\dot{\phi},\dot{\phi}}\bar{H}^j \\ \implies \qquad j\tilde{Z}_{j,\dot{\phi}} &= -\dot{\phi}\tilde{Z}_{j,\dot{\phi},\dot{\phi}} \\ \implies \qquad \tilde{Z}_j &= u_j(\bar{a},\phi)\mathscr{I}_j(\dot{\phi}) + v_j(\bar{a},\phi)\,, \end{aligned}\tag{4.2.62}$$

where

$$\mathscr{I}_j(\dot{\phi}) = \begin{cases} \ln(\dot{\phi}) & \text{if } j = 1 \\ \dot{\phi}^{1-j} & \text{if } j \neq 1 \end{cases} \tag{4.2.63}$$

From this, and the definition of $\tilde{\mathcal{L}}$ given by eq. (4.2.41), we see that,

$$\bar{a}^3 \sum_{j=0}^{3} \tilde{Z}_j \left(\frac{s}{\bar{a}}\right)^j = \bar{a}^3 \sum_{j=0}^{3} \Big(u_j(\bar{a},\phi)\mathscr{I}_j(\dot{\phi}) + v_j(\bar{a},\phi)\Big)\left(\frac{s}{\bar{a}}\right)^j \cong 0\,. \tag{4.2.64}$$

By equating powers in $\dot{\phi}$, it is clear that $u_j = 0$ for all j, and this evidently implies that $\tilde{Z}_{j,\dot{\phi}} = 0$ for all j. However, we know that in order for the theory to be self-tuning, it must have a non-tivial dependence on $\dot{\phi}$ (such that its *on-shell-in-$\bar{a}$* Hamiltonian is not independent of $\dot{\phi}$). Therefore, we are forced to conclude that $\Delta\mathcal{E}^\phi$ vanishes identically (i.e. $\Delta A = 0$), and so,

$$- \bar{a}^3 \sum_{j=0}^{3} j(1-j)\sigma_j \frac{\bar{H}^{j-1}}{\bar{a}\dot{\phi}^j} \;=\; 0 \qquad \Longrightarrow \qquad j(1-j)\sigma_j \bar{a}^2 \frac{\bar{H}^{j-1}}{\dot{\phi}^j} \;=\; 0\,. \tag{4.2.65}$$

We see that, for $i = \{0,1\}$ this condition is satisfied by $j(1-j)$, however, for $i = \{2,3\}$, $j(1-j) \neq 0$ and so it must be that,

$$\sigma_2 \;=\; 0\,, \qquad \sigma_3 \;=\; 0\,. \tag{4.2.66}$$

Therefore, in all three cases, $\Delta\mathcal{E}^\phi$ vanishes identically, implying that $\sigma_2 = 0 = \sigma_3$, and furthermore, that $\mathcal{E}^\phi = \tilde{\mathcal{E}}^\phi$.

We can utilise these results to determine explicit expressions for the remaining non-zero functions, σ_0 and σ_1. To this end, we refer back to eq. (4.2.59); noting that $\Delta\mathcal{E}^\phi$ must vanish identically (in all three cases), it is clear that $\Delta\mathscr{A} = \Delta\mathscr{B} = \Delta\mathscr{C} = 0$ identically. Now, from $\Delta\mathscr{A} = 0$ and using eq. (4.2.61a) with ΔZ_j given by eq. (4.2.50), we can immediately imply that $\sigma_2 = 0 = \sigma_3$. Moreover, referring back to eq. (4.2.61c) and requiring that $\Delta\mathscr{C} = 0$ identically, we find that

$$\sigma_{1,\phi} \;=\; 3\sigma_0 + \bar{a}\sigma_{0,\bar{a}} \;=\; \frac{1}{\bar{a}^2}\left(\bar{a}^3\sigma_0\right)_{,\bar{a}} \qquad \Longrightarrow \qquad \bar{a}^3\sigma_0 \;=\; \mu_{,\phi}\,, \qquad \bar{a}^2\sigma_1 \;=\; \mu_{,\bar{a}}\,, \tag{4.2.67}$$

where $\mu = \mu(\bar{a},\phi)$. Following this extensive analysis, we finally consider the difference $\Delta\mathcal{L}$ between the Lagrangians, $\mathcal{L}$ and $\tilde{\mathcal{L}}$, of the two different self-tuning Horndeski theories,

$$\begin{aligned} \Delta\mathcal{L} \;&=\; \mathcal{L} - \tilde{\mathcal{L}} \;=\; \bar{a}^3 \sum_{j=0}^{3} \Delta Z_j \bar{H}^j \;=\; \bar{a}^3\Delta Z_0 + \bar{a}^3\Delta Z_1 \bar{H} \\ &=\; \dot{\phi}\mu_{,\phi} + \dot{\bar{a}}\mu_{,\bar{a}} \;=\; \dot{\mu}\,. \end{aligned} \tag{4.2.68}$$

We see, therefore, that any two self-tuning Horndeski theories differ by a total derivative, and as such a general self-tuning Lagrangian is equivalent to eq. (4.2.41), up to a total derivative $\frac{\mathrm{d}}{\mathrm{d}t}\mu(\bar{a},\phi)$.

With this information at our disposal, we are in a position to determine the forms of the functions $\mathscr{X}_j$ and $\mathscr{Y}_j$ for our self-tuning Lagrangian. To this end, we shall compare the general form of the self-tuning Lagrangian, $\mathcal{L} = \tilde{\mathcal{L}} + \frac{\mathrm{d}}{\mathrm{d}t}\mu(\bar{a},\phi)$ (where

$\tilde{\mathcal{L}}$ is given by eq. (4.2.41)), with our expressions for the component functions $\tilde{Z}_j$ in terms of the functions $\mathscr{X}_j$ and $\mathscr{Y}_j$, as defined by eq. (4.2.20). We find that,

$$\begin{aligned}\tilde{\mathcal{L}} + \dot{\mu} &= \bar{a}^3\Bigg\{\tilde{Z}_1\Big[\bar{H} - \frac{s}{\bar{a}}\Big] + \tilde{Z}_2\Big[\bar{H}^2 - \Big(\frac{s}{\bar{a}}\Big)^2\Big] + \tilde{Z}_1\Big[\bar{H}^3 - \Big(\frac{s}{\bar{a}}\Big)^3\Big]\Bigg\} + \dot{\phi}\mu_{,\phi} + \dot{\bar{a}}\mu_{,\bar{a}} \\ &= \bar{a}^3\Bigg\{\Big(\mathscr{X}_0 - \frac{s^2}{\bar{a}^2}\mathscr{Y}_0\Big) + \Big(\mathscr{X}_1 - \frac{s^2}{\bar{a}^2}\mathscr{Y}_1\Big)\bar{H} + \mathscr{X}_2\bar{H}^2 + \mathscr{X}_3\bar{H}^3\Bigg\}, \end{aligned} \tag{4.2.69}$$

which, upon equating powers in $\bar{H}$ we obtain the following set of equations,

$$-\frac{s}{\bar{a}}\tilde{Z}_1 - \left(\frac{s}{\bar{a}}\right)^2\tilde{Z}_2 - \left(\frac{s}{\bar{a}}\right)^3\tilde{Z}_3 + \bar{a}^{-3}\dot{\phi}\mu_{,\phi} = \mathscr{X}_0 - \frac{s^2}{\bar{a}^2}\mathscr{Y}_0\,, \tag{4.2.70a}$$

$$\tilde{Z}_1 + \bar{a}^{-2}\mu_{,\bar{a}} = \mathscr{X}_1 - \frac{s^2}{\bar{a}^2}\mathscr{Y}_1\,, \tag{4.2.70b}$$

$$\tilde{Z}_2 = \mathscr{X}_2 \tag{4.2.70c}$$

$$\tilde{Z}_3 = \mathscr{X}_3\,. \tag{4.2.70d}$$

Substituting Eqs. (4.2.70b), (4.2.70c) and (4.2.70d) into eq. (4.2.70a) leads to the relation,

$$\bar{a}^{-3}\dot{\phi}\mu_{,\phi} + s\bar{a}^{-3}\mu_{,\bar{a}} - \mathscr{X}_0 - \frac{s}{\bar{a}}\mathscr{X}_1 - \left(\frac{s}{\bar{a}}\right)^2\Big[\mathscr{X}_2 + \mathscr{Y}_0\Big] - \left(\frac{s}{\bar{a}}\right)^3\Big[\mathscr{X}_3 + \mathscr{Y}_1\Big] = 0\,. \tag{4.2.71}$$

To proceed, we now restrict attention to $s \neq 0$, and expand $\bar{a}^{-3}\dot{\phi}\mu$ as a power series in $\frac{s}{\bar{a}}$,

$$\bar{a}^{-3}\mu(\bar{a},\phi) = \sum_{j\,=\,-\infty}^{\infty} V_j(\phi)\left(\frac{s}{\bar{a}}\right)^j\,, \tag{4.2.72}$$

where $V_j(\phi)$ are arbitrary functions of ϕ. Inserting this into eq. (4.2.71), we have,

$$\sum_{j\,=\,-\infty}^{\infty}\dot{\phi}V_{j,\phi}\left(\frac{s}{\bar{a}}\right)^j + \sum_{j\,=\,-\infty}^{\infty}(3-j)V_{j,\phi}\left(\frac{s}{\bar{a}}\right)^{j+1} - \sum_{j\,=\,0}^{3}\mathscr{X}_j\left(\frac{s}{\bar{a}}\right)^j - \sum_{j\,=\,0}^{1}\mathscr{Y}_j\left(\frac{s}{\bar{a}}\right)^{j+2} = 0\,, \tag{4.2.73}$$

and so, upon equating powers of $\frac{s}{\bar{a}}$, first in the cases where $j \leq -1$ and $j \geq 4$, we find that,

$$\dot{\phi}V_{j,\phi} + (4-j)V_{j-1} = 0 \qquad j \leq -1 \text{ or } j \geq 4\,. \tag{4.2.74}$$

Since V_j does not depend on $\dot{\phi}$, it follows that,

$$V_{-1} = \text{const.}\,,\; V_{-2} = V_{-3} = \cdots = 0\,,\; V_4 = V_5 = \cdots = 0\,. \tag{4.2.75}$$

As such, equating the remaining powers of $\frac{s}{\bar{a}}$, i.e. for $0 \leq j \leq 3$, we arrive at a set of equations for $\mathscr{X}_j$ and $\mathscr{Y}_j$,

$$\mathscr{X}_0 = 4V_{-1} + \dot{\phi}V_{0,\phi} = -\Lambda_{\text{bare}} + \dot{\phi}V_{0,\phi}\,, \tag{4.2.76a}$$

$$\mathscr{X}_1 = 3V_0 + \dot{\phi}V_{1,\phi}\,, \tag{4.2.76b}$$

$$\mathscr{X}_2 + \mathscr{Y}_0 = 2V_1 + \dot{\phi}V_{2,\phi}\,, \tag{4.2.76c}$$

$$\mathscr{X}_3 + \mathscr{Y}_1 = V_2 + \dot{\phi}V_{3,\phi}\,, \tag{4.2.76d}$$

where we have identified the arbitrary constant with the bare cosmological constant, $V_{-1} = \text{const.} = -\frac{1}{4}\Lambda_{\text{bare}}$. This is an important consistency check, since the vacuum energy renormalises this term. If the theory did not admit such an arbitrary constant, then we would have effectively fine-tuned the bare cosmological constant to zero against the vacuum energy, precisely the result of Weinberg's no-go theorem (as discussed in §3.2.1). Since the theory does allow for an arbitrary cosmological constant term, we can be confident in stating that the theory self-tunes.

Having ascertained relations between the functions $\mathscr{X}_j$ and $\mathscr{Y}_j$ in terms of arbitrary potentials $V_j(\phi)$, we can then use these to evaluate the Horndeski functions K and G_i $(i = 3, 4, 5)$ in the self-tuning limit by comparing Eqs. (4.2.76a) to (4.2.76d) with Eqs. (4.2.18a) to (4.2.18d).

4.2.4 Recovering the Fab-Four

Having derived a set of equations for the Horndeski functions K and G_i $(i = 3, 4, 5)$ in the self-tuning limit, we are now in the position to analyse particular cases of our disformally coupled self-tuning theory. The first is an important consistency check; since our self-tuning theory is a generalisation of the Fab-Four theory, it must reduce to the Fab-Four, in the special case where $\bar{A} = 1$, $B = 0$ and $N = 1$ (implied by eq. (4.2.15a)). We see that, in this particular case, the Jordan and Horndeski frames coincide, and so we have that $X = \bar{X} = \frac{1}{2}\dot{\phi}^2$, and $\bar{a} = a$. Thus, upon referring back to Eqs. (4.2.76a) to (4.2.76d) and Eqs. (4.2.18a) to (4.2.18d), we are left with the following set of differential equations

$$\dot{\phi}^2\tilde{G}_{5,\dot{\phi},\dot{\phi}} - 3\dot{\phi}\tilde{G}_{5,\dot{\phi}} + 3\tilde{G}_5 = \dot{\phi}V_{3,\phi} + V_2\,, \tag{4.2.77a}$$

$$6\dot{\phi}G_{4,\dot{\phi}} - 12G_4 - 3\dot{\phi}^2G_{5,\phi} + 3\dot{\phi}\tilde{G}_{5,\phi} = \dot{\phi}V_{2,\phi} + 2V_1\,, \tag{4.2.77b}$$

$$3\dot{\phi}G_3 - 3\tilde{G}_3 - 6\dot{\phi}G_{4,\phi} = \dot{\phi}V_{1,\phi} + 3V_0\,, \tag{4.2.77c}$$

$$K - \dot{\phi}\tilde{G}_{3,\phi} = -\Lambda_{\rm bare} + \dot{\phi}V_{0,\phi}\,, \tag{4.2.77d}$$

where we have recalled eq. (4.2.19), and noted that in this particular case $G_i = \tilde{G}_{i,\dot{\phi}}$ for $i = 3, 5$. Integrating each of these equations iteratively, we have that

$$\begin{aligned} K(\phi, X) =& -\Lambda_{\rm bare} + 2Xf_3'(\phi) + 4X^2 f_4''(\phi) - 2X^3 g_5'''(\phi) \\ & - \frac{1}{2}X^2\left[2\ln(\sqrt{2X}) - 1\right]V_3''''(\phi) \end{aligned} \tag{4.2.78a}$$

$$G_3(\phi, X) = f_3(\phi) - \frac{1}{4}X\left[2\ln(\sqrt{2X}) - 1\right]V_3'''(\phi) + 6Xf_4'(\phi) + 5X^2 g_5''(\phi)\,, \tag{4.2.78b}$$

$$\begin{aligned} \tilde{G}_3(\phi, X) =& -V_0(\phi) + \sqrt{2X}f_3(\phi) - \frac{1}{4}X\sqrt{2X}\left[2\ln(\sqrt{2X}) - 1\right]V_3'''(\phi) \\ & + 2X\sqrt{2X}f_4'(\phi) + X^2\sqrt{2X}g_5''(\phi)\,, \end{aligned} \tag{4.2.78c}$$

$$G_4(\phi, X) = 2Xf_4(\phi) - \frac{1}{2}X\ln(\sqrt{2X})V_3''(\phi) + 2X^2 g_5'(\phi) - \frac{1}{6}V_1(\phi)\,, \tag{4.2.78d}$$

$$G_5(\phi, X) = f_5(\phi) + 6Xg_5(\phi) - \frac{1}{2}\left[\ln(\sqrt{2X}) + 1\right]V_3'(\phi)\,, \tag{4.2.78e}$$

$$\tilde{G}_5(\phi, X) = \sqrt{2X}f_5(\phi) + 2X\sqrt{2X}g_5(\phi) - \frac{1}{2}\sqrt{2X}\ln(\sqrt{2X})V_3'(\phi) + \frac{1}{3}V_2(\phi)\,, \tag{4.2.78f}$$

where f_3, f_4, g_4, f_5 and g_5 are arbitrary functions of ϕ. We see, therefore, that the self-tuning limit of the Horndeski Lagrangian, eq. (4.2.17) on an FRW background, in the case where $\bar{A} = 1$, $N = 1$ and $B = 0$, has the form:

$$
\begin{aligned}
\mathcal{L} \;=\; & \bar{a}^3\left[K-\sqrt{2X}\tilde{G}_{3,\phi}+3\Big(2G_4-\sqrt{2X}\tilde{G}_{5,\phi}\Big)\left(\frac{s}{a}\right)^2\right] \\
& +\bar{a}^3\left[3\sqrt{2X}G_3-3\tilde{G}_3-6\sqrt{2X}G_{4,\phi}+3\Big(\sqrt{2X}G_5-\tilde{G}_{5,\phi}\Big)\left(\frac{s}{a}\right)^2\right]\bar{H} \\
& +\bar{a}^3\Big[12XG_{4,X}-6G_4-6XG_{5,\phi}\Big]\bar{H}^2+2\bar{a}^3X\sqrt{2X}G_{5,X}\bar{H}^3\,. \qquad (4.2.79)
\end{aligned}
$$

In order to deduce the covariant form of this Lagrangian we observe from eq. (4.2.78) that each of the functions f_3, f_4, g_4, f_5 and g_5 and the potential terms V_i ($i=3,4,5$) are completely de-coupled from one another, that is, we can express eq. (4.2.79) in terms of a sum of component Lagrangians, each purely in terms of one of the functions f_3, f_4, g_4, f_5, g_5 and V_i ($i=3,4,5$), and one containing only the bare cosmological constant, i.e. $\mathcal{L}_\Lambda=-\sqrt{-g}\Lambda_{\rm bare}$. This enables us to analyse the form of eq. (4.2.79) on a case-by-case basis, in which we "switch-on" each of these functions individually. In doing so, after integrating by parts, it is found that the contributions to eq. (4.2.78), pertaining to the functions V_0, V_2 and f_3, are vanishing.[9] The remaining functions, V_1, V_3, f_4, f_5 and g_5 provide non-trivial contributions, however, f_4 and f_5 lead to the same expression, implying that only four of the eight original functions yield independent terms in eq. (4.2.78). As component Lagrangians, these are given by,

$$
\mathcal{L}_{V_1} \;=\; -\,a^3V_1\left[\frac{\ddot{a}}{a}+H-\left(\frac{s}{a}\right)^2\right], \qquad (4.2.80a)
$$

$$
\mathcal{L}_{V_3} \;=\; \frac{3}{2}a^3V_3\frac{\ddot{a}}{a}\left[H^2-\left(\frac{s}{a}\right)^2\right], \qquad (4.2.80b)
$$

$$
\mathcal{L}_{f_4} \;=\; 12a^3Xf_4\left[H^2-\left(\frac{s}{a}\right)^2\right], \qquad (4.2.80c)
$$

$$
\mathcal{L}_{g_5} \;=\; 12a^3X\sqrt{2X}g_5\left[H^2-\left(\frac{s}{a}\right)^2\right]H\,. \qquad (4.2.80d)
$$

[9] It turns out that the reason for this stems from the fact that, in the fully covariant theory, these terms correspond to total derivatives, as was the case in the orignal Fab-Four construction [60].

Comparing these expressions with the following curvature invariants, evaluated on an FRW background,

$$R\big|_{FRW} = 6\left[\frac{\ddot{a}}{a} + H - \left(\frac{s}{a}\right)^2\right], \tag{4.2.81a}$$

$$\hat{G}\big|_{FRW} = 24\frac{\ddot{a}}{a}\left[H^2 - \left(\frac{s}{a}\right)^2\right], \tag{4.2.81b}$$

$$G^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi\Big|_{FRW} = 6X\left[H^2 - \left(\frac{s}{a}\right)^2\right], \tag{4.2.81c}$$

$$P^{\mu\nu\alpha\beta}\nabla_\mu\phi\nabla_\alpha\phi\nabla_\nu\nabla_\beta\phi\Big|_{FRW} = 6X\sqrt{2X}\left[H^2 - \left(\frac{s}{a}\right)^2\right]H\,. \tag{4.2.81d}$$

We arrive at the fully covariant expressions for the component Lagrangians [eq. (4.2.80)],

$$\mathcal{L}_{V_1} = -\sqrt{-g}\frac{V_1}{6}R\big|_{FRW}, \tag{4.2.82a}$$

$$\mathcal{L}_{V_3} = \frac{1}{16}\sqrt{-g}V_3\hat{G}\big|_{FRW}, \tag{4.2.82b}$$

$$\mathcal{L}_{f_4} = 2\sqrt{-g}f_4G^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi\Big|_{FRW}, \tag{4.2.82c}$$

$$\mathcal{L}_{g_5} = 2\sqrt{-g}g_5P^{\mu\nu\alpha\beta}\nabla_\mu\phi\nabla_\alpha\phi\nabla_\nu\nabla_\beta\phi\Big|_{FRW}. \tag{4.2.82d}$$

Referring back to the component Fab-Four Lagrangians [eq. (4.1.5)], and comparing these to the above expressions, it is evident, upon making the following identifications: $V_1 := -6V_g$, $V_3 := 16V_r$, $f_4 := \frac{1}{2}V_j$ and $g_5 := \frac{1}{2}V_p$, that we have recovered Fab-Four theory, as required.

Proceeding to the next case, in which $\bar{A} = \bar{A}(\phi)$, $B = 0$ and $N = \bar{A}$ (implied by eq. (4.2.15a)), we see that this corresponds to a conformal transformation of the metric. Upon following the same procedure as in the previous case, upon a redefinition of the arbitrary coefficient functions, one arrives back at the component Fab-Four Lagrangians [eq. (4.2.82)]. It is clear, therefore, that conformally related self-tuning Horndeski theories fall within the same class, i.e. they are all Fab-Four theories. This is to be expected, since it is inherited from the full Horndeski theory, itself a class of conformally related theories.[10] Moreover, it has recently been shown that if a Lagrangian is related to the Horndeski Lagrangian by a special disformal transformation, in which $\bar{A} = \bar{A}(\phi)$ and $\bar{B} = \bar{B}(\phi)$, then in can always be recast into Horndeski form [184]. Therefore, to generalise beyond Fab-Four theory, we must consider the most general case in which $\bar{A}$ and $\bar{B}$ (and implicitly N) are dependent on ϕ *and* $\bar{X} = \bar{X}(\phi, X)$.

4.2.5 Disformal couplings & beyond Fab-Four

Our starting point, in this most general case, is to refer back to Eqs. (4.2.76a) to (4.2.76d); observing that the right-hand sides of these equations are independent of $\ddot{\phi}$, it must be the case that the left-hand sides are also. Now, from the definitions of the functions $\mathscr{X}_j$ and $\mathscr{Y}_j$, (implied from Eqs. (4.2.18a) to (4.2.18d)), we see that they contain terms proportional to $\dot{\bar{A}}(\phi, \bar{X})$, which are dependent on ϕ, $\dot{\phi}$ *and* $\ddot{\phi}$. As such, for each of Eqs. (4.2.76a) to (4.2.76d), the sum of terms proportional to $\ddot{\phi}$ must vanish. Concentrating on eq. (4.2.76a), we see from eq. (4.2.18a), that the left-hand side has the form

$$\begin{aligned}\mathscr{X}_0 \;=\; & N\bar{A}^3K - \dot{\phi}\tilde{G}_{3,\phi} + \left[3\bar{A}^2\frac{\dot{\phi}}{N}G_3 - 6\bar{A}^2\frac{\dot{\phi}}{N}G_{4,\phi}\right]\dot{\bar{A}} \\ & + \left[6\frac{\bar{A}}{N}\left(\frac{\dot{\phi}}{N}\right)^2 G_{4,X} - 6\frac{\bar{A}}{N}G_4 - 3\frac{\bar{A}}{N}\left(\frac{\dot{\phi}}{N}\right)^2 G_{5,\phi}\right]\dot{\bar{A}}^2 + \frac{1}{N^2}\left(\frac{\dot{\phi}}{N}\right)^3 G_{5,X}\dot{\bar{A}}^3 \,,\end{aligned} \tag{4.2.83}$$

enabling us to isolate the terms proportional to $\dot{\bar{A}}$, which can be expressed in the schematic form

$$\mathscr{X}_0 \;\supset\; \alpha(\phi,\dot{\phi})\dot{\bar{A}}(\phi,\bar{X}) + \beta(\phi,\dot{\phi})\dot{\bar{A}}^2(\phi,\bar{X}) + \gamma(\phi,\dot{\phi})\dot{\bar{A}}^3(\phi,\bar{X}) \,, \tag{4.2.84}$$

and so, upon expanding $\dot{\bar{A}}$,

$$\dot{\bar{A}}(\phi,\bar{X}) \;=\; \dot{\phi}\bar{A}_{,\phi} + \dot{\bar{X}}\bar{A}_{,\bar{X}} \;=\; \lambda(\phi,\dot{\phi}) + \delta(\phi,\dot{\phi})\ddot{\phi} \,, \tag{4.2.85}$$

[10] Any theory related to Horndeski theory by a conformal transformation can be recast into "Horndeski form", i.e. upon appropriate redefinitions of the Horndeski functions $K \to \tilde{K}$ and $G_i \to \tilde{G}_i$ $(i = 3, 4, 5)$, its component Lagrangians will be of the form given in eq. (4.1.2).

we can extract the terms proportional to $\ddot{\phi}$, leading to the following constraint equation

$$\left(\alpha\delta + 2\beta\lambda\delta + 3\gamma\lambda^2\delta\right)\ddot{\phi} + \left(\beta\delta^2 + 3\gamma\lambda\delta^2\right)\ddot{\phi}^2 + \gamma\delta^3\ddot{\phi}^3 \;=\; 0\,. \tag{4.2.86}$$

Now, since we are considering the most general case, it must be that $\bar{A}_{,\phi} \neq 0$ and $\bar{A}_{,\bar{X}} \neq 0$, implying that $\lambda \neq 0$ and $\delta \neq 0$. By equating powers in $\ddot{\phi}$, it is evident from eq. (4.2.86) that $\gamma = \alpha = \beta = 0$, and hence, each of the terms proportional to $\dot{\bar{A}}$ vanish. This procedure can be applied to each of the remaining equations (4.2.76b), (4.2.76c) and (4.2.76d), and in each case, it is found that the coefficients of each power in $\dot{\bar{A}}$ vanish identically.

Let us now study the implications of this result. From eq. (4.2.83), we can immediately infer that

$$G_{5,X} \;=\; 0 \tag{4.2.87a}$$

$$G_4 \;=\; \left[2G_{4,X} - G_{5,\phi}\right]X\,, \tag{4.2.87b}$$

$$G_3 \;=\; 2G_{4,\phi}\,. \tag{4.2.87c}$$

These results then imply that

$$\begin{aligned}\mathscr{X}_1 \;&=\; 3\bar{A}^3\frac{\dot{\phi}}{N}G_3 - 3\tilde{G}_3 - 6\bar{A}^3\frac{\dot{\phi}}{N}G_{4,\phi} + \Bigg[12\frac{\bar{A}^2}{N}\left(\frac{\dot{\phi}}{N}\right)^2 G_{4,\bar{X}} \\ &\quad - 12\frac{\bar{A}^2}{N}G_4 - 6\frac{\bar{A}^2}{N}\left(\frac{\dot{\phi}}{N}\right)^2 G_{5,\phi}\Bigg]\dot{\bar{A}} + 3\frac{\bar{A}}{N^2}\left(\frac{\dot{\phi}}{N}\right)^3 G_{5,\bar{X}}\dot{\bar{A}}^2 \\ &=\; -3\tilde{G}_3 \;=\; 2V_1 + \dot{\phi}V_{2,\phi}\end{aligned}$$

$$\Rightarrow \qquad \tilde{G}_3(\phi, X) \;=\; -\frac{1}{3}\left(2V_1 + N\sqrt{2X}V_{2,\phi}\right)\,. \tag{4.2.88}$$

Moving onto eq. (4.2.76c), and applying eq. (4.2.87), we have

$$\begin{aligned}\mathscr{X}_2 + \mathscr{Y}_0 &= 6\frac{\bar{A}^3}{N}\left(\frac{\dot{\phi}}{N}\right)^2 G_{4,X} - 6\frac{\bar{A}^3}{N}G_4 - 3\frac{\bar{A}^3}{N}\left(\frac{\dot{\phi}}{N}\right)^2 G_{5,\phi} \\ &\quad - 6N\bar{A}G_4 + 3\dot{\phi}\tilde{G}_{5,\phi} + \left[6\frac{\bar{A}^2}{N^2}\left(\frac{\dot{\phi}}{N}\right)^3 G_{5,X} - 3\frac{\dot{\phi}}{N}G_5\right]\dot{\bar{A}} \\ &= -6\frac{\bar{A}^3}{N}G_4 - 3\frac{\bar{A}^3}{N}\left(\frac{\dot{\phi}}{N}\right)^2 G_{5,\phi} - 3\frac{\dot{\phi}}{N}G_5\dot{\bar{A}}\,, \end{aligned} \tag{4.2.89}$$

it immediately follows that $G_5 = 0$, and us such, we are left with

$$\begin{aligned}\mathscr{X}_2 + \mathscr{Y}_0 &= -6N\bar{A}G_4 + 3\dot{\phi}\tilde{G}_{5,\phi} = 2V_1 + \dot{\phi}V_{2,\phi}\,, \\ G_4(\phi, X) &= \frac{1}{6N\bar{A}}\left[N\sqrt{2X}\left(3\tilde{G}_{5,\phi} - V_{2,\phi}\right) - 2V_1\right], \end{aligned} \tag{4.2.90}$$

Finally, turning our attention to eq. (4.2.76d),

$$\begin{aligned}\mathscr{X}_3 + \mathscr{Y}_1 &= \frac{\bar{A}^3}{N^2}\left(\frac{\dot{\phi}}{N}\right)^3 G_{5,X} - 3\frac{\bar{A}}{N}\dot{\phi}G_5 + 3\tilde{G}_5 \\ &= 3\tilde{G}_5 = V_2 + \dot{\phi}V_{3,\phi} \\ \Rightarrow \qquad \tilde{G}_5 &= \frac{1}{3}\left(V_2 + N\sqrt{2X}V_{3,\phi}\right)\,. \end{aligned} \tag{4.2.91}$$

Observe that $G_5 = \frac{\bar{A}}{\sqrt{2X}}\tilde{G}_{5,X} = G_5 = 0$, which implies that $\tilde{G}_{5,X} = 0$. Given the expression for $\tilde{G}_5$ from eq. (4.2.91), it follows that

$$\frac{V_{3,\phi}}{3}\left(\sqrt{2X}N_{,X} + \frac{N}{\sqrt{2X}}\right) = 0 \qquad \Longrightarrow \qquad N(\phi, X) = \frac{\mathscr{J}(\phi)}{\sqrt{2X}}\,. \tag{4.2.92}$$

This is a contradiction, however, as ϕ and $\dot{\phi}$ are independent variables. Since it is the case that $\sqrt{2X} = \dot{\phi}$, it is not possible for N to have the form given by eq. (4.2.92). To resolve this contradiction, we are therefore forced to conclude that $\bar{A}_X = 0$. Consequently, the most general disformal coupling that admits self-tuning must have $\bar{A}$ as a function of ϕ only, which from eq. (4.2.7), implies that A is also at most a function of ϕ (note however, that this does not preclude $\bar{B}$ from remaining a function of both ϕ and $\bar{X}$). This is an important result, as the form of the Horndeski

Lagrangian does not change under conformal transformations, meaning that if we transform the gravitational metric, such that $g_{\mu\nu} \to \hat{g}_{\mu\nu} = A^2(\phi) g_{\mu\nu}$, we find

$$\begin{aligned}
\mathscr{L} &= \mathscr{L}_H(g,\phi,X) + \mathscr{L}_{\mathrm{M}}(\bar{g},\Psi) = \mathscr{L}_H(g,\phi,X) + \mathscr{L}_{\mathrm{M}}(A^2 g + A^2 B^2 \partial\phi\partial\phi,\Psi) \\
&= \tilde{\mathscr{L}}_H(\hat{g},\phi,X) + \mathscr{L}_{\mathrm{M}}(\hat{g} + \hat{B}^2\partial\phi\partial\phi,\Psi) \\
&= \tilde{\mathscr{L}}_H(\hat{g},\phi,X) + \mathscr{L}_{\mathrm{M}}(\bar{\hat{g}},\Psi)
\end{aligned} \tag{4.2.93}$$

where we have redefined B such that $\hat{B} = AB$. As such, in the self-tuning limit of the theory, we are at liberty to set $\bar{A} = A = 1$ without loss of generality. It follows from this, that $\dot{\bar{A}} = 0$, and as such, we are left with the following set of differential equations for the K and G_i

$$K - \sqrt{2X}\tilde{G}_{3,\phi} = -\Lambda_{\mathrm{bare}} + \sqrt{2X}V_{0,\phi}\,, \tag{4.2.94a}$$

$$3\sqrt{2X}G_3 - 3\tilde{G}_3 - 6\sqrt{2X}G_{4,\phi} = 3V_0 + N\sqrt{2X}V_{1,\phi}\,, \tag{4.2.94b}$$

$$\frac{12}{N}XG_{4,X} - \frac{6}{N}G_4 - \frac{6}{N}XG_{5,\phi} + 3N\sqrt{2X}\tilde{G}_{5,\phi} - 6NG_4 = 2V_1 + N\sqrt{2X}V_{2,\phi}\,, \tag{4.2.94c}$$

$$\frac{2}{N^2}X\sqrt{2X}G_{5,X} + 3\tilde{G}_5 - 3\sqrt{2X}G_5 = V_2 + N\sqrt{2X}V_{3,\phi}\,. \tag{4.2.94d}$$

We can use these expressions, along with several integrations by parts to drastically simplify the form of the cosmological Lagrangian [eq. (4.2.17)]. Indeed, it is found that the self tuning limit of the disformally coupled Horndeski Lagrangian, evaluated on an FRW background, can be expressed as

$$\begin{aligned}
\mathcal{L} = {}& \bar{a}^3\left[N\sqrt{2X}V_{1,\phi} - 2V_1\frac{s}{\bar{a}} + \left(3\tilde{G}_5 - 3\sqrt{2X}G_5 - V_2\right)\left(\frac{s}{\bar{a}}\right)^2\right]\left[\bar{H} - \frac{s}{\bar{a}}\right] \\
&+ \bar{a}^3\left[N\sqrt{2X}V_{2,\phi} + 6NG_4 - 3N\sqrt{2X}\tilde{G}_{5,\phi} + 2V_1\right]\left[\bar{H}^2 - \left(\frac{s}{\bar{a}}\right)^2\right] \\
&+ 2\bar{a}^3\frac{X\sqrt{2X}}{N^2}G_{5,X}\left[\bar{H}^3 - \left(\frac{s}{\bar{a}}\right)^3\right].
\end{aligned} \tag{4.2.95}$$

Note that eq. (4.2.95) has the nice feature that it is manifestly in self-tuning form.

Unfortunately, unlike the Fab-Four case, the differential equations for K and G_i, eq. (4.2.94), cannot be integrated in general form, due to the presence of an unknown function, $B(\phi, X)$ (appearing implicitly in eq. (4.2.94) through N); the system is *under-determined*. As such, one must first either specify the form of this arbitrary function, or alternatively, specify the form of at least one of the functions K and G_i. Ultimately, one aims to obtain a covariant form for this generalisation of the Fab-Four, however, it has proven to be a difficult task, and as of yet, has not been achieved.

Before closing our discussion on self-tuning solutions to the CCP, we shall present a particular solution to the set of differential equations (4.2.94). This serves to show firstly, that non-trivial (consistent) solutions do exist in the disformal case, and secondly, that the corresponding Lagrangian cannot be put into Fab-Four form, highlighting that it does constitute an extension beyond Fab-Four. To this end, we return to eq. (4.2.94); it is evident that we have a system of four differential equations with five unknown functions, N, K and G_i ($i = 3, 4, 5$). Focussing our attention on eq. (4.2.94a), we can trivially solve for K, leaving us with three remaining equations and four unknown functions. It is clear that, whichever way we look at it, the system is *under-determined*; this freedom is what enables us to choose how matter is to be disformally coupled to gravity. Given this, we first observe that we cannot choose N to have any form we like. In particular, it cannot be of the form $N = f(\phi)(2X)^{-1/2}$, as according to eq. (4.2.12) this would imply that ϕ and $\dot{\phi}$ are dependent on one another, which is inconsistent, since ϕ and $\dot{\phi}$ are independent variables. With this in mind, in the interest of clarity, we shall use eq. (4.2.94d), to implicity choose a form for N

$$\frac{2}{N^2} X\sqrt{2X} G_{5,X} \;=\; N\sqrt{2X} V_{3,\phi}\,. \tag{4.2.96}$$

With this choice, eq. (4.2.94d) can be readily solved to give

$$\begin{aligned} 2X\tilde{G}_{5,X} - \tilde{G}_5 \;=\; -\frac{V_2}{3} \qquad &\Rightarrow \qquad \tilde{G}_5(\phi, X) \;=\; f(\phi)\sqrt{2X} + \frac{V_2}{3} \\ &\Rightarrow \qquad G_5(\phi, X) \;=\; f(\phi)\,. \end{aligned} \tag{4.2.97}$$

However, note from eq. (4.2.96), that this implies $N\sqrt{2X}V_{3,\phi} = 0$, and since we require N to be non-vanishing, it follows that $V_3 = \text{const}$. We see then, that in fact eq. (4.2.96) places no constraint on the form of N. Now turning our attention to eq. (4.2.94c), in the interest of obtaining an analytic solution, we restrict to the case in which the arbitrary function $f(\phi) = f = \text{const.}$. Furthermore, since we have ascertained that eq. (4.2.96) places no restriction on the form of N, we shall now choose it to have the form

$$N(t) \;=\; g(\phi)X - 1\,. \tag{4.2.98}$$

Finally, we note that the function V_1 in eq. (4.2.94c) is also arbitrary, and so we shall set it to zero. In doing so, we find the following solution for G_4

$$12XG_{4,X} - 6g(\phi)XG_4 \;=\; 0 \qquad \Longrightarrow \qquad G_4(\phi,X) \;=\; h(\phi)\exp\left(\frac{g(\phi)X}{2}\right)\,. \tag{4.2.99}$$

It thus remains for us to determine G_3 and K, and to do so, we shall simplify eq. (4.2.94b) by setting $g(\phi) = g = \text{const.}$ and $h(\phi) = h = \text{const.}$, such that $G_{4,\phi} = 0$, and so

$$\begin{aligned} 6X\tilde{G}_{3,X} - 3\tilde{G}_3 \;=\; 3V_0 \qquad &\Rightarrow \qquad \tilde{G}_3(\phi,X) \;=\; I(\phi)\sqrt{2X} - V_0(\phi) \\ &\Rightarrow \qquad G_3(\phi,X) \;=\; I(\phi)\,. \end{aligned} \tag{4.2.100}$$

Given this, we can trivially determine K, from eq. (4.2.94a), to be

$$K(\phi,X) \;=\; 2I(\phi)X\,. \tag{4.2.101}$$

In summary, we find the following set of solutions

$$G_5(\phi,X) \;=\; f \;=\; \text{const.}\,, \tag{4.2.102a}$$

$$G_4(\phi,X) \;=\; h\exp\left(\frac{gX}{2}\right)\,, \tag{4.2.102b}$$

$$G_3(\phi,X) \;=\; I(\phi)\,, \tag{4.2.102c}$$

$$K(\phi,X) \;=\; 2I(\phi)X\,, \tag{4.2.102d}$$

and upon inserting them back into our disformally self-tuning Lagrangian [eq. (4.2.95)], we find that for this specific case, it has the following form

$$\mathcal{L} \;=\; 6\bar{a}^3Nh\exp\left(\frac{gX}{2}\right)\left[\bar{H}^2 - \left(\frac{s}{\bar{a}}\right)^2\right], \qquad N \;=\; gX - 1\,. \tag{4.2.103}$$

It is clear from this, that eq. (4.2.103) cannot be expressed in terms of the component Fab-Four Lagrangians [eq. (4.1.5)]. Note, however, that in this simple example, the self-tuning Lagrangian [eq. (4.2.103)] can be written in covariant form as

$$\mathcal{L} \;=\; \sqrt{-g}\, F(\bar{X})\bar{G}^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi \tag{4.2.104}$$

where $F(\bar{X}) = \frac{h}{X}\exp\left(\frac{g\bar{X}}{2N^2}\right)$, with N given by eq. (4.2.98) (in which X is implicitly a function of $\bar{X}$, cf. eq. (4.2.5)), and $\bar{G}^{\mu\nu}$ is the Einstein tensor evaluated in the Jordan frame (cf. eq. (4.2.81c)). Thus, in this particular case the Lagrangian is a subset of the beyond Horndeski extension of the Fab-Four found by Babichev *et al.* [183].

4.3 Discussion

In this chapter, we have have shown that it is possible to construct a theory that admits self-tuning solutions that absorb all contributions to the vacuum energy, and thus evading the CCP, at least at the classical level. This is possible due to the assumption that the vacuum configuration of the self-adjusting field ϕ breaks Poincaŕe invariance, thus evading Weinberg's no-go theorem [128], and enabling it to dynamically screen the vacuum energy from the spacetime curvature. We gave a brief overview of the original construction of such a self-tuning theory, the so-called Fab-Four. We then proceeded to consider whether this theory could be generalised to incorporate disformal couplings to the matter sector, whilst preserving the desired self-tuning dynamics of the scalar field.

Through a careful analysis, we were able to show that it is indeed possible to generalise the Fab-Four theory and obtain a self-tuning theory of gravity in which matter is disformally coupled to the self-tuning scalar field ϕ. In §4.2.4, we were able to show that this generalisation is consistent with the known results, reproducing the Fab-Four theory for both minimal coupling to the Horndeski metric and a minimal coupling to a Weyl-rescaled Horndeski metric. This is true so long as the scaling function depends on ϕ, but not its canonical kinetic term X. Interestingly, by passing the disformal theory through the self-tuning filter (cf. §4.1), we were forced to place strong constraints on the form of any disformal coupling to matter in the theory. Indeed, for self-tuning to be feasible, the conformal part of any disformal coupling to matter must necessarily be a function of ϕ alone (as opposed to being a function of both ϕ and X). This constraint significantly simplies the theory, since we can take advantage of the fact that a Lagrangian, conformally related to a Horndeski Lagrangian, remains in the Horndeski class, to effectively set the conformal factor $A(\phi)$ to unity. Furthermore, this simplification enabled us to express the theory in "self-tuning form" such that, upon solving the related differential equations for the Horndeski functions K and G_i ($i = 3, 4, 5$), the theory is guaranteed to be self-tuning.[11]

The caveat in this analysis is that, unlike the original Fab-Four, the set of coupled differential equations that must be solved in order to determine K and G_i, cannot be done so in general. Essentially, the system is underdetermined, arising from our freedom to choose the disformal coupling to matter, and so we can, at best, find solutions on a case-by-case basis, for each specfic choice of the disformal function B. Having said this, we were able to show that non-trivial solutions for K and G_i do exist, for a particular choice of B (implicity defined though N). Moreover, this simple analysis provided us with information on the inadmissable choices of the lapse function $N(t)$, and served to highlight that the resulting Lagrangian cannot be expressed in Fab-Four form.

Clearly there is still work to be done on this disformal generalisation. The main hurdle being that we are yet to discover a covariant form for the theory; owing to the presence of the arbitrary disformal function B, the path to covariantisation is less clear than in the original Fab-Four case. However, it may be possible to make a fairly general choice for the form of B as a starting point, and be able to identify associated curvature invariants for the solutions on an FRW background. More interestingly, as the disformal function B is dependent on both ϕ and X, the theory cannot be brought into Horndeski form under field re-definitions (this would be possible if B were a function of ϕ alone, cf. [184]), and as such, should one be able to find a covariant form for the theory, it would fall into a *beyond Horndeski* class. More recently, research has been conducted into the possibility of constructing theories that go beyond Horndeski, whilst still evading Ostrogradsky's instability [140] (cf. §3.2.2), see, for example [192–195], and the implications this may have for Vainshtein screening, e.g. [196, 197]. In particular, recently D. Langlois and K. Noui were able to derive a general class of degenerate higher-order scalar-tensor ($DHOST$) theories [198, 199]. Should one be able to construct a covariant form for our disformally self-tuning theory, one might expect that it would fall within this class, and it would be interesting to investigate this. A further avenue of future research, would be to

[11] This point is somewhat tautological, since the differential equations for K and G_i were derived by placing self tuning constraints on the theory. However, the fact that the theory can be recast into "self-tuning form" on an FRW background before determining K and G_i, explicitly shows that the theory is self-tuning, so long as there exist non-trivial solutions for K and G_i.

study the cosmology of our disformal theory, and whether there are phenomenologically viable solutions; one might naively hope, that as this appeared to be possible for the original Fab-Four, the same might be true for its disformal generalisation. However, it would require a detailed analysis to determine whether this is the case or not. Moreover, it is by no means certain that our theory is stable under radiative corrections, although, even if it turns out not to be, it may provide a useful stepping stone towards a self-tuning theory that is.

Aside from self-tuning, other long-distance modifications to gravity have been made, in an attempt to solve the CCP. As briely mentioned in §1.4, a particularly interesting idea is that of sequestering the SM vacuum energy, including all its loop corrections such that it never impacts on spacetime curvature. This approach, first developed by Kaloper and Padilla [58], proposes that one only modifies GR in the infinite wavelength limit in such a way that no new propagating degrees of freedom are introduced into the theory. Furthermore, by construction, it removes all contributions from the vacuum energy (to the energy-momentum tensor), including matter loops to all orders, rendering it a radiatively stable solution to the CCP. A particulary appealing attribute of this theory is that it reduces exactly to GR locally, not requiring any form of screening mechanism to pass solar system tests. For more in-depth details of this proposed solution, and recent analyses, we refer the reader to the following references [58, 59] and [200, 201].

Before bringing our discussion on self-tuning solutions to the CCP to a close, we would be remiss not to discuss the recent data obtained from the gravitational wave detection due to a binary neutron star collision. The data from the detection, named *GW170817*, has seemingly placed strong constraints on modified theories of gravity, in particular, scalar-tensor theories, having severely reduced their viable regions of parameter space. Indeed, several papers were published following the data release, completely ruling out the cubic Galileon (a subset of Horndeski theory), and essentially all variants of Horndeski theory (unless one accepts severe fine-tuning) apart from simple cases, as viable models for dark energy [202–206]. In particular, these analyses imply that Fab-Four theory is ruled out, since it relies on the quartic and quintic contributions in the Lagrangian, which have been ruled out, to provide Vainshtein screening [202].

The conclusions made rest on the constraints placed on the speed of gravitational waves, i.e. that $-3 \times 10^{-15} \leq \frac{c_g - c}{c} \leq 5 \times 10^{-16}$ (where c_g is the gravitational wave speed and c is the speed of light). Indeed, Horndeski theories general predict c_g to deviate from c, with the relative difference between the two values parameterised by a quantity $\alpha_T = \frac{c_g^2 - c^2}{c^2}$. It should be noted, however, that there are loopholes in the arguments made in these papers. For example, in [205], an expression for α_T is derived in terms of the Horndeski functions, $M_*^2 \alpha_T \equiv 2X \left[2G_{4,X} - 2G_{5,\phi} - \left(\ddot{\phi} - \dot{\phi} H\right) G_{5,X}\right]$ (where $M_*^2 \equiv 2\left(G_4 - 2XG_{4,X} + XG_{5,\phi}\right)$). The claim is that the only way we can arrive at $\alpha_T \approx 0$, is if $G_5 \approx$ const. and $G_{4,X} \approx 0$, otherwise there would have to be some delicate cancellation between $G_{4,X}$, $G_{5,\phi}$ and $G_{5,X}$, which would correspond to fine-tuning, and unstable to radiative corrections.

However, there are two options available, yet to be thoroughly explored, that could potentially lead to non-trivial solutions for ϕ that are not contrived. Indeed, one can use the EOM for the scalar field ϕ, to express $\ddot{\phi}$ in terms of $\dot{\phi}$, H and $\dot{H}$. We can then use the analogue Friedmann equations to express this in terms of the energy density arising from the matter sector. By inserting this into the expression into α_T, that we wish to vanish, the problem will then be recast into determining choices of the matter content of the universe for which H evolves in a particular fashion, such that it forces $\alpha_T \approx 0$. Of course, one might argue that such choices for the matter sector may need to be finally tuned, nonetheless, it is still worth investigating. Another option, is to seek solutions for the Horndeski functions K, G_3, G_4 and G_5, such that α_T vanishes independently of matter content of the universe. It could well be that such solutions do not exist, or are simply trivial, but again, one should not discard this possiblility without further exploration.

Of course, even if non-trivial solutions for the Horndeski functions exist, it remains to be seem whether they are able to describe a sensible evolution of the universe that matches observations. It should further be noted, that the constraints placed on beyond Horndeski, e.g. DHOST theories are less severe but still restrictive [202, 206]. Taking into account the possible loopholes discussed above, it is possible that these can be alleviated somewhat. If this were to be the case, since the disformal extension of the Fab-Four is a beyond Horndeski theory (most probably within a class of DHOST theories), the model would still remain a phenomenologically viable option. Having said this, recent work by Creminelli *et al.* on graviton decay into dark energy fluctuations has placed tighter constraints on the types of beyond Horndeski theories compatible with the gravitational wave data [207]. This will undoubtedly make it more difficult to keep the model viable, however, a more detailed analysis would need to be carried out to determine its status.

Chapter 5

(P)reheating the early universe

5.1 Preheating: a toy model

In §2 we briefly discussed the inflationary stage of the universe and the nature of the subsequent necessary reheating process, in which the energy density held in the inflaton condensate is deposited back into the universe, enabling the production of the SM particles, and allowing for the early universe to enter a radiation dominated epoch. As noted in §2.2.1, this is an important stage in the thermal evolution of the universe, having significant consequences for leptogenesis (see, e.g., ref. [105]) and for the generation of dark matter relic densities (see, e.g., ref. [106]). The process of reheating enables us to transition to the standard HBB model, whose description of BBN agrees well with experimental data [107].

In §2.2.2, we further commented that one must be careful to take into account the non-perturbative nature of the inflaton condensate at the end of inflation, and the non-trivial effects that this has on particle production, particularly during the early stages of reheating. Indeed, the coherent, collective behaviour exhibited by the inflaton condensate induces parametric resonances of the fields that it couples to, resulting in efficient, explosive particle production. Since parametric resonance is a phenomenon that occurs in the early stages of reheating, preceding the latter stages of perturbative decay, it is referred to as preheating.

We shall now give a more detailed review of preheating, following the original papers of Kofman, Linde and Starobinsky[1] [98, 113] and the reviews [100, 114, 213, 214], to which we refer the reader to for further details.

[1]Note that the phenomenon of parametric resonance was already known of, and well understood, having been previously applied to particle production by external electric fields, e.g. [208–210]. Indeed, early attempts to apply the theory of parametric resonance were made by Dologov and Kirilova [211], and by Traschen and Brandenberger [212], however, it was Kofman, Linde and Starobinsky who presented the first rigorous analysis of both narrow and broad resonance phenomena in reheating.

We start our discussion by referring back to the equations of motion for the classical inflaton field, φ, given by Eqs. (2.1.24) and (2.1.25). In the most simple case of chaotic inflation [90, 91], in which the inflaton potential is of the form $V(\varphi) = \frac{1}{2}m_\phi^2\varphi^2$, it is possible to parametrise eq. (2.1.25) in terms of the Hubble parameter H and an angular variable θ, as follows,

$$\dot{\varphi} \;=\; \sqrt{6}M_{\rm Pl}H\sin(\theta)\,, \tag{5.1.1a}$$

$$\varphi \;=\; \sqrt{6}\frac{M_{\rm Pl}}{m_\phi}H\cos(\theta)\,. \tag{5.1.1b}$$

Using eq. (2.1.24), we arrive at the equations of motion describing the dynamics of H and θ,

$$\dot{H} \;=\; -\,3H^2\sin^2(\theta)\,, \tag{5.1.2a}$$

$$\dot{\theta} \;=\; -\,m_\phi - \frac{3}{2}H\sin(2\theta)\,. \tag{5.1.2b}$$

Observe from eq. (5.1.2a), that H has a solution of the form,

$$H \;=\; \frac{2}{3t}\left(1 - \frac{\int^t {\rm d}t'\cos(2\theta)}{t}\right)^{-1}\,, \tag{5.1.3}$$

and thus decays over time. From this, it is then clear that the second term on the right-hand side of eq. (5.1.2b) describes oscillations with a decaying amplitude. Indeed, we find that $\dot{\theta} = -m_\phi\big[1 + \frac{1}{m_\phi t}\big(1 - \frac{1}{t}\int^t {\rm d}t'\cos(2\theta)\big)^{-1}\sin(2\theta)\big]$, and so for $m_\phi t \gg 1$, we can neglect this term to obtain an approximate solution for θ of the form $\theta(t) \simeq -m_\phi t$. The Hubble parameter can then be readily obtained from eq. (5.1.2a), such that

$$H(t) \simeq \frac{2}{3t}\left(1 - \frac{\sin(2m_\phi t)}{2m_\phi t}\right)^{-1}\,, \tag{5.1.4}$$

which is valid for $m_\phi t \gg 1$. Expanding this solution in powers of $(m_\phi t)^{-1}$, and substituting the result into eq. (5.1.1b), we thus obtain an approximate expression for the inflaton

$$\varphi(t) \simeq \varphi_0(t)\cos(m_\phi t)\big(1 + {\rm sinc}(2m_\phi t)\big)\,. \tag{5.1.5}$$

As inflation draws to a close, the "friction" term $3H\dot{\varphi}$ in eq. (2.1.24) becomes less and less important, with inflation terminating at $\varphi \sim M_{\rm Pl}/2$. Since eq. (5.1.5) is valid for $m_\phi t \gg 1$, it is applicable near the end of inflation, and it has been shown that the amplitude of φ drops off sufficiently enough that we can neglect the sinc-function contribution in eq. (2.1.24) [98]. Consequently, at the onset of reheating, the solution for the inflaton field φ asymptotically approaches the form

$$\varphi(t) \simeq \varphi_0(t)\cos(m_\phi t)\,, \qquad \varphi_0(t) \;\simeq\; \frac{2\sqrt{6}M_{\rm Pl}}{3m_\phi t}\,, \tag{5.1.6}$$

where $\varphi_0(t)$ is the (decaying) oscillation amplitude of the inflaton field.

Now, in order to capture the non-perturbative effects present during preheating, we shall again consider a Lagrangian of the form given in eq. (2.2.1), treating the inflaton classically and the scalar field χ (to which it is coupled) quantum mechanically. As such, we quantise the χ field in the presence of a classical time dependent background field $\varphi(t)$. In the Heisenberg picture, we can thus expand $\hat{\chi}$ in terms of its Fourier modes as follows

$$\hat{\chi}(t,\mathbf{x}) \;=\; \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\left[\hat{a}_{\mathbf{k}}\chi_{\mathbf{k}}(t)e^{+i\mathbf{k}\cdot\mathbf{x}} + \hat{a}^\dagger_{\mathbf{k}}\chi^*_{\mathbf{k}}(t)e^{-i\mathbf{k}\cdot\mathbf{x}}\right], \tag{5.1.7}$$

where $\hat{a}_{\mathbf{k}}$ and $\hat{a}^\dagger_{\mathbf{k}}$ are the annihilation and creation operators, respectively.[2] For a spatially flat ($k=0$) FRW background, the temporal part of each of the field modes satifies the following EOM

$$\ddot{\chi}_{\mathbf{k}} + 3H\dot{\chi}_{\mathbf{k}} + \left(\frac{\mathbf{k}^2}{a^2} + m_\chi^2 + g\varphi^2\right)\chi_k \;=\; 0\,, \tag{5.1.8}$$

where $\mathbf{k}$ is the comoving momentum. Note that this equation describes an oscillator with a variable frequency ω due to the presence of the time-dependent background field $\varphi(t)$, and the expansion of the universe (captured in the time-dependent scale factor $a(t)$). For the sake of clarity, let us neglect the expansion of the universe for now, and furthermore, we observe from eq. (5.1.6) that typically $\varphi_0(t)$ varies slowly with respect to the oscillation frequencies of both φ and χ. This corresponds to setting $a=1$ (and hence $H=0$) and $\varphi_0(t) \approx$ const. in eq. (5.1.8).[3] We are left, therefore, with an EOM for an oscillator with periodically changing fre-

[2]Note that we are employing the condensed notation $\hat{a}_{\mathbf{k}} \coloneqq \hat{a}(\mathbf{k})$, and $\chi_{\mathbf{k}}(t) \coloneqq \chi(\mathbf{k},t)$ (and likewise for $\hat{a}^\dagger_{\mathbf{k}}$ and $\chi^*_{\mathbf{k}}(t)$).

[3]In fact, if we neglect the additional decay of the inflaton amplitude φ_0 due to the transfer of energy to the χ field, then provided that the time period of preheating is small compared to the Hubble expansion time H^{-1}, this approximation $a \approx 1$ and $\varphi_0 \approx$ const. is a reasonable one.

quency $\omega^2_{\mathbf{k}}(t) = \mathbf{k}^2 + m^2_\chi + g\varphi_0^2 \cos^2(m_\phi t)$, which upon making a change of variables $mt = z - \pi/2$, can be recast into the form of a *Mathieu equation*

$$\chi''_{\mathbf{k}} + (A_{\mathbf{k}} - 2q\cos 2z)\,\chi_k \;=\; 0\,, \tag{5.1.9}$$

where the primes denote derivatives with respect to z, and the parameters $A_{\mathbf{k}}$ and q have been defined as follows,

$$A_{\mathbf{k}} \;=\; \frac{\mathbf{k}^2 + m^2_\chi}{m^2_\phi} + 2q\,, \tag{5.1.10a}$$

$$q \;=\; \frac{g\varphi_0^2}{4m^2_\phi}\,. \tag{5.1.10b}$$

The solutions to eq. (5.1.9) can be determined according to Floquet's theorem, and will be of the form [213],

$$\chi_{\mathbf{k}}(z) \;=\; e^{m_{\mathbf{k}}z}\mathscr{P}_1(z) + e^{-m_{\mathbf{k}}z}\mathscr{P}_2(z)\,, \tag{5.1.11}$$

where $\mathscr{P}_i$ $(i = 1, 2)$ are periodic functions of z, and $m_{\mathbf{k}}$ is a complex number known as the *Floquet exponent*, which depends on $A_{\mathbf{k}}$ and q. Note that the *real* part of the Floquet exponent $\mu_{\mathbf{k}} \coloneqq \mathrm{Re}\, m_{\mathbf{k}}$ is always *non-negative*, i.e. $\mu_{\mathbf{k}} \geq 0$. An important feature of the solutions to a Mathieu equation, is that they possess instabilities for certain ranges of $|\mathbf{k}|$. Specifically, for $\mu_{\mathbf{k}} > 0$, the corresponding modes $\chi_{\mathbf{k}}$ exhibit exponential growth, whereas, for $\mu_k = 0$, the modes are stable and no parametric resonance occurs. In this sense $\mu_{\mathbf{k}}$ parameterises the instability of the system, describing the exponential growth of the mode functions $\chi_{\mathbf{k}}$ in the unstable regions of parameter space $(q, A_{\mathbf{k}})$. As such, for clarity, we shall refer to $\mu_{\mathbf{k}}$ as the *instability parameter* of the system.

In general, it is found that the stabilities/instabilities present in such a system manifest a band structure in its corresponding parameter space, with the boundaries between stable and unstable regions being functions of $A_{\mathbf{k}}$ and q. This is evident in fig. 5.1, in which we can see the band structure of the stable and unstable regions for a range of values for q and $A_{\mathbf{k}}$. Note from eq. (5.1.10a) that only certain regions of instability are physically realisable, these correspond to values of momenta $|\mathbf{k}| \geq 0$, and as such, parametric resonance of the modes χ_k can only occur above the line $A_0 = (m_\chi/m_\phi)^2 + 2q$. From an in-depth analysis of the Mathieu equation

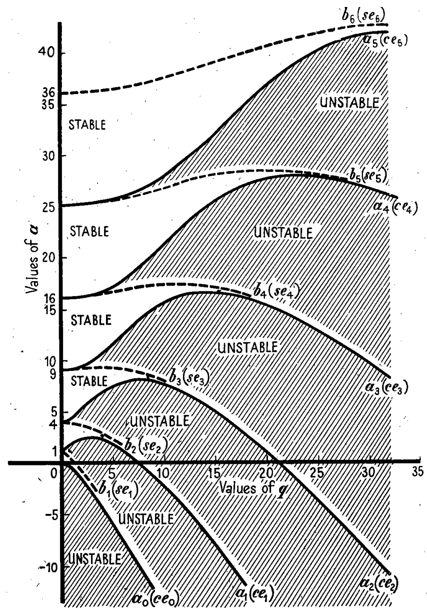


Figure 5.1.1: Stability chart for solutions to the Mathieu equation. The horizontal axis is the parameter q, and the vertical axis the value of A. The shaded regions correspond to regions of instability in parameter space in which parametric resonance occurs. [Source: McLachlan, 1951 [215]].

[215], it is found that the width of these resonance bands $\Delta A_{\mathbf{k}}^{(l)}$ (where $l \in \mathbb{Z}^+$), and furthermore the efficiency of preheating, are governed solely by the parameter q. Indeed, it is found that for small values of q, i.e. $q \ll 1$, resonance can only occur within narrow bands, a regime that is aptly referred to as *narrow resonance*. In the opposite extreme, i.e. $q \gg 1$, resonance can take place in broad bands, which include all long-wavelength modes $|\mathbf{k}| \to 0$; this regime is referred to as *broad resonance*.

Before moving on to discuss the details of narrow and broad resonance, we shall address the physical interpretation of the resonant effects on the mode functions $\chi_{\mathbf{k}}$ due to these instabilities. Indeed, the exponential growth of modes lying within a given resonance band can be interpreted physically as rapid particle production, leading to an exponential amplification of occupation numbers $n_{\mathbf{k}}(t) = \langle \hat{a}_{\mathbf{k}}^{\dagger} \hat{a}_{\mathbf{k}} \rangle$ in those modes. To see this, we observe that the energy per mode can be extracted from the expectation value of the effective Hamiltonian for the χ field,

$$\langle \hat{H}_{\mathbf{k}}(t) \rangle \;=\; \omega_{\mathbf{k}}(t) \left(N_{\mathbf{k}}(t) + \frac{1}{2} \right) \;=\; \frac{1}{2} \Big(|\dot{\chi}_{\mathbf{k}}|^2 + \omega_{\mathbf{k}}^2(t) |\chi_{\mathbf{k}}|^2 \Big) \,, \tag{5.1.12}$$

such that the number density $N_{\mathbf{k}}$ of χ particles per mode is given by,

$$N_{\mathbf{k}}(t) \;=\; \frac{1}{2\omega_{\mathbf{k}}(t)}\left(|\dot{\chi}_{\mathbf{k}}|^2 + \omega_{\mathbf{k}}^2(t)|\chi_{\mathbf{k}}|^2\right) - \frac{1}{2}\,. \tag{5.1.13}$$

Now, from eq. (5.1.11), we know that unstable modes grow as $\chi_{\mathbf{k}} \sim e^{\mu_{\mathbf{k}} z}$, and so we see from eq. (5.1.13), that this sources an exponentially increasing particle number density in these modes, $N_{\mathbf{k}} \sim e^{2\mu_{\mathbf{k}} z}$. This implies that the growth rate of the particle number densities in these modes is proportional to the current number density per mode, i.e. $\dot{N}_{\mathbf{k}} \sim 2\mu_{\mathbf{k}} m_\phi N_{\mathbf{k}} = \Gamma_{\mathrm{PR}} N_{\mathbf{k}}$, and thus capturing the effect of Bose enhancement, in agreement with the expected result in the perturbative calculation of this effect discussed in §2.2.2).

It is important to note that this rapid production of particles results in a highly non-thermal spectrum, as evidenced by the band structure in fig. 5.1. As such, the particles produced during the preheating phase are far from thermal equilibrium, and will therefore subsequently thermalise, eventually resulting in an equilibrium distribution, at which point the system will have a well-defined temperature.[4]

5.1.1 Narrow resonance

As briefly mentioned, the defining condition of narrow resonance is that $q \ll 1$, which from eq. (5.1.10b), implies that $g\varphi^2 \ll m_\phi^2$. In this case, resonance occurs only in narrow bands in momentum space centered around $A_{\mathbf{k}}^{(l)} \simeq l^2$ ($l \in \mathbb{Z}^+$), each with a width of order $\Delta A_{\mathbf{k}}^{(l)} = |A_{\mathbf{k}}^{(l)} - l^2| \sim q^l$, as dictated by the theory of Mathieu's equation [215]. The corresponding resonance then occurs for modes with $\mathbf{k}^2 \sim m_\phi^2\left(l^2 - 2q \pm q^l\right) - m_\chi^2$, and since $q \ll 1$, it is evident that, as l increases, these instability bands become progressively narrower. Consequently, it is clear that the widest and most important instability band is the first one, $A_{\mathbf{k}}^{(1)} \sim 1 \pm q$,, for which $\mathbf{k}^2 \sim m_\phi^2\left(1 - 2q \pm q\right) - m_\chi^2$. The instability parameter $\mu_{\mathbf{k}}^{(1)}$ for this first band is given by [215],

$$\mu_{\mathbf{k}}^{(1)} \simeq \sqrt{\left(\frac{q}{2}\right)^2 - \left(A_{\mathbf{k}}^{1/2} - 1\right)^2}\,. \tag{5.1.14}$$

[4]Note that in equilibrium, the distribution of particles across the available energy states is proportional to $(e^{\beta\varepsilon_i} \pm 1)^{-1}$ for fermions (+) and bosons (−), where $\beta = (k_B T)^{-1}$, and ε_i is the energy corresponding to the i^{th} state. We can therefore determine the inverse temperature T of the system from the gradient of the logarithmic plot of this distribution. However, far from thermal equilibrium, things are much more complex, in particular, the distribution of particles will not be parametrised by a single quantity β. Simply put, the system will not have a thermal distribution of particles and therefore will not have a well-defined notion of temperature.

At the edges of the resonance band $\mu_{\mathbf{k}}^{(1)} = 0$ as expected, since in stable regions the real part of the Floquet exponent, $m_{\mathbf{k}}$, must vanish. Furthermore, it has a maximal value $\mu_{\mathbf{k}}^{(1)} = \frac{q}{2}$ at the center of resonance, $k \sim m_\phi$, corresponding to a maximal growth in modes $\chi_{\mathbf{k}}$ with this momentum of order $\chi_{\mathbf{k}} \sim e^{qz/2} \simeq e^{qm_\phi t/2}$, with a corresponding increase in the number density of these modes $N_{\mathbf{k}} \sim \exp(\frac{g\varphi_0^2\pi}{4m_\phi^2}N)$, where $N = m_\phi t/2\pi$ counts the number of oscillations of the inflaton condensate since the start of reheating. One can provide a natural physical interpretation for the resonance in the first band peaking at $|\mathbf{k}| \sim m_\phi$. Indeed, since $g\varphi_0^2 \ll m_\phi^2$ (and assuming that $m_\chi \ll m_\phi$), the effective mass of the χ particles is much less than m_ϕ, and as such, two ϕ particles decay into two χ particles with momenta $|\mathbf{k}| \sim m_\phi$. We can generalise this particle picture by noting that, for the l^{th} instability band, the resonance is centered around $|\mathbf{k}| \sim lm_\phi$, and as such can be interpreted as the collective process $2l \times \phi \to \chi\chi$.

Note that the resonant production of particles is drastically different to the perturbative production discussed in §2.2.1. The decay rate Γ_φ of the latter is suppressed by a factor g^2/m_ϕ^2, and is thus very inefficient (since we take the coupling to be weak), whereas, in the former $\Gamma_{\text{PR}} \sim m_\phi q$, and so is much greater than Γ_φ for $\varphi_0^2 > g\sigma^2/2\pi$. Moreover, referring back to eq. (2.2.8), we see that in the perturbative approach the growth rate of the χ particle number density (per momentum mode) $N_{\mathbf{k}}$ is proportional to the number density of the inflatons N_ϕ, conversely, in the case of parametric resonance the growth rate is proportional to $N_{\mathbf{k}}$ itself, i.e. it is contingent on the number of χ particles already produced, resulting in an exponential growth. Note, however, that one can provide a link between the perturbative and non-perturbative analyses in this scenario. Indeed, one can interpret the Bose condensation effects in the leading-order perturbative process $\phi\phi \to \chi\chi$ as a parametric resonance of the χ modes due to the first ($l = 1$) instability band ($|\mathbf{k}| \sim m_\phi$). The higher-order resonance bands ($l > 1$) then correspond to the higher-order processes involving collective decays of $l \times \phi$ particles into pairs of χ particles, taking into account the Bose effects due to dense populations of the χ modes with $|\mathbf{k}| \sim lm_\phi$.

It is clear then, that by solving eq. (5.1.8) non-perturbatively, we capture the collective behaviour of the inflaton condensate φ, and thus opening up additional decay channels for φ to transfer energy to the χ field. Indeed, in the perturbative approach, it is assumed the χ particles produced by the oscillating inflaton field φ are on the (bare) mass-shell, i.e. $k_\chi^2 = -m_\chi^2$ (where k_χ^μ is the four-momentum of the produced

χ particles), however, if parametric resonance is taken into account, then the χ field has a time-dependent effective mass, and so particles can be produced off the (bare) mass-shell, i.e. $k_\chi^2 = -m_\chi^2 - g\varphi^2(t)$. Consequently, energy can be transferred from the inflaton condensate to the χ field much more efficiently, resulting in an exponential growth of field modes $\chi_{\mathbf{k}}$ (and accordingly the number density $n_{\mathbf{k}}$). Intuitively this makes sense, since the effective mass of the χ field fluctuates with the oscillations of the inflaton condensate, such that at certain points in its cycle the χ field can become very light relative to the inflaton mass, at which point becomes kinematically possible for a large number of χ particles to be created.

We note, however, that as the name suggests the resonance is *narrow*, and thus only occurs for modes $\chi_{\mathbf{k}}$ within a small region of momentum space. As a result, for most of momentum space, the perturbative techniques discussed in §2.2.1 are applicable. Indeed, if the resulting decay rate is smaller than $2\mu_{\mathbf{k}} m_\phi \sim m_\phi q$, then parametric resonance and perturbative decay of the inflaton field may coexist. Resonance of the χ field does not persist indefinitely, as φ gradually loses energy to the χ field, its amplitude decreases, eventually becoming smaller than $(g/8\pi)\sigma$, at which point perturbative decays take over and the amplitude of φ decreases exponentially (within a time $\sim \Gamma^{-1}_{\phi\to\chi\chi}$), terminating any remaining resonance. Indeed, for resonance to occur, it must proceed at a faster rate than perturbative decays, i.e. $\Gamma_{\text{PR}} \sim m_\phi q = \frac{g\varphi_0^2}{4m_\phi} \gtrsim \Gamma_{\phi\to\chi\chi}$, and once this condition is violated any resonant effects disappear. At this point, perturbative decays take over, in which one can treat the inflaton quanta as decaying independently into χ particles, and the analysis carried out in §2.2.1 is applicable. These will continue until the inflaton condensate has completely decayed, upon which the system will proceed to thermalise and ultimately rejoin with the standard HBB model.

5.1.2 Broad resonance

In models such as chaotic inflation, it is possible for the initial oscillations of the inflaton condensate to have a large amplitude, $\varphi_0 \sim 0.1\, M_{\text{Pl}}$. In this case q can be very large, i.e. $q = \frac{g\varphi_0^2}{4m_\phi^2} \gg 1$, and so the resonance bands of momenta, $|\mathbf{k}|$, can be very broad. In particular, all infrared modes ($|\mathbf{k}| \to 0$) reside within these bands, resulting in extremely efficient reheating. This is the regime of *broad resonance*, and since $q \gg 1$, the techniques and results used to determine $\mu_{\mathbf{k}}$ in the narrow resonance regime cannot be applied. One can still calculate $\mu_{\mathbf{k}}$ for $q \gg 1$, however, the procedure is more complicated. That being said, for completeness, we quote here the general solution for $\mu_{\mathbf{k}}$ (for fixed $A_{\mathbf{k}}$ and q),

$$\mu_{\mathbf{k}} \;=\; \frac{1}{\pi}\ln\left|\sqrt{F_{\mathbf{k}}^2}+\sqrt{F_{\mathbf{k}}^2-1}\right| , \tag{5.1.15}$$

where $F_{\mathbf{k}} = 1 + \left[\frac{\mathrm{d}}{\mathrm{d}z}\tilde{\chi}_{\mathbf{k}}^{(1)}(z=\frac{\pi}{2})\right]\tilde{\chi}_{\mathbf{k}}^{(2)}(z=\frac{\pi}{2})$, in which $\tilde{\chi}_{\mathbf{k}}^{(1)}$ and $\tilde{\chi}_{\mathbf{k}}^{(2)}$ are solutions to eq. (5.1.9), satisfying the intitial conditions $\tilde{\chi}_{\mathbf{k}}^{(1)} = 1$, $\frac{\mathrm{d}\tilde{\chi}_{\mathbf{k}}^{(1)}}{\mathrm{d}z} = 0$, and $\tilde{\chi}_{\mathbf{k}}^{(2)} = 0$, $\frac{\mathrm{d}\tilde{\chi}_{\mathbf{k}}^{(2)}}{\mathrm{d}z} = 1$ at $z = 0$, respectively [216]. In particular, note that parametric resonance occurs whenever $F_{\mathbf{k}} > 1$.

In the interest of clarity, however, we shall instead follow a different route to extract the salient features of broad resonance (and refer the reader to, e.g., [216–218] for further details on the procedure for determining $\mu_{\mathbf{k}}$). Indeed, let us return to eq. (5.1.8), and note that, in our current approximation ($a(t) = 1$ and $\varphi_0(t) \approx$ const.) this reduces to,

$$\ddot{\chi}_{\mathbf{k}}(t) + \omega_{\mathbf{k}}^2(t)\chi_{\mathbf{k}}(t) \;=\; 0 , \tag{5.1.16}$$

where $\omega_{\mathbf{k}}^2(t) = |\mathbf{k}|^2 + m_\chi^2 + g\varphi_0^2\cos^2(m_\phi t)$. This differential equation has the following approximate solution

$$\chi_{\mathbf{k}}(t) \approx \frac{\chi_{\mathbf{k}}(t_0)}{\sqrt{\omega_{\mathbf{k}}(t)}}\exp\left(\pm i\int \mathrm{d}t'\omega_{\mathbf{k}}(t')\right) , \tag{5.1.17}$$

which holds whenever the condition

$$\frac{|\dot{\omega}_{\mathbf{k}}(t)|}{\omega_{\mathbf{k}}^2(t)} < 1 , \tag{5.1.18}$$

is satisfied. This is the so-called WKB, or adiabatic approximation. Whenever eq. (5.1.18) is fulfilled then the system evolves adiabatically[5], and in particular, the number density $N_{\mathbf{k}}$ is (approximately) an adiabatic invariant, which can be seen by referring

[5]Note that, during reheating, the inflaton condensate induces a time-dependent oscillation frequency $\omega_{\mathbf{k}}(t)$ for any field coupled to the condensate (in this case χ). Qualitatively, an *adiabatic* evolution of the system then corresponds to a slow change in this oscillation frequency. Indeed, when φ is *not* near $\varphi(t) = 0$, $\omega_{\mathbf{k}}(t)$ varies sufficiently slowly, inducing a gradual evolution of the systems Hamiltonian. As such, the state of the system remains in the same instantaneous eigenstate of the Hamiltonian as it was in before the system started to evolve. However, when φ passes through $\varphi(t) = 0$, this causes an abrupt change in $\omega_{\mathbf{k}}(t)$, in turn, inducing a rapid change in the Hamiltonian of the system. Consequently, the system is no longer in the same instantaneous eigenstate of the Hamiltonian as it was before the *non-adiabatic* evolution occurred.

back to eq. (5.1.13)

$$N_{\mathbf{k}} \;=\; \frac{1}{2\omega_{\mathbf{k}}(t)}\Big(|\dot{\chi}_{\mathbf{k}}(t)|^2+\omega_{\mathbf{k}}^2(t)|\chi_{\mathbf{k}}(t)|^2\Big)-\frac{1}{2} \;\approx\; \omega_{\mathbf{k}}(t)|\chi_{\mathbf{k}}(t)|^2-\frac{1}{2} \;\approx\; \text{const.} \quad (5.1.19)$$

In fact, during broad resonance, the typical oscillation frequency of the χ field $\omega(t) = \sqrt{|\mathbf{k}|^2 + m_\chi^2 + 4m_\phi^2 q\cos^2(m_\phi t)}$ is much greater than that of the inflaton condensate, and as such the χ field oscillates many times ($\mathcal{O}(q^{1/2})$) during each period of oscillation of φ. For this reason, throughout most of this interval, the effective mass of the χ field varies adiabatically. We see then, that $N_{\mathbf{k}}$ can vary significantly when there is a *non-adiabatic* change in $\omega_{\mathbf{k}}$, i.e. when the condition given by eq. (5.1.18) is violated. As such, resonant particle production occurs during intervals of the inflaton oscillations where $\frac{|\dot{\omega}_{\mathbf{k}}|}{\omega_{\mathbf{k}}^2} \gtrsim 1$, i.e.,

$$\frac{g|\varphi||\dot{\varphi}|}{\big(|\mathbf{k}|^2 + m_\chi^2 + g\varphi^2\big)^{3/2}} \gtrsim 1\,. \quad (5.1.20)$$

It is clear then, that eq. (5.1.18) is strongly violated in cases where $|\dot{\varphi}| \gtrsim \sqrt{g}\,|\varphi|^2$, and only for particular values of momenta (as we shall discuss below). Moreover, $|\dot{\varphi}| \gtrsim \sqrt{g}\,|\varphi|^2$ at points in the inflaton's interval of oscillation near $\varphi = 0$, at which the effective mass of the χ field attains its minimal value $m_{\text{eff}} \approx m_\chi \ll m_\phi$. As such, during broad resonance, particles are produced in rapid bursts (and as we will see) across a broad range of momenta during short intervals of each oscillation of the inflaton condensate. The result is an exponentially fast growth of the number density. Note that this differs from the case of narrow resonance, in which particle production occurs smoothly over time. Indeed, the adiabatic condition [eq. (5.1.18)] is never violated during narrow resonance, however, it is always non-zero and therefore particle production can still take place, but proceeding at a much slower rate than resonant production. In this case the dominant contribution arises from the first resonance band, and since $q \ll 1$, we have that $|\mathbf{k}| \sim m_\phi \sim \omega_{\mathbf{k}}$ in its centre (for $m_\chi \ll m_\phi$). The corresponding mode $\chi_{\mathbf{k}}$ oscillates at approximately the same frequency as the inflaton condensate φ. Typically, the rate of parametric resonance in the narrow regime does not differ much for the rate of growth of these modes [98]. The reason why resonance occurs in the narrow regime is due to Bose enhancement of inflaton decays into certain χ modes (those within resonance bands of momenta identified in the Floquet analysis, cf. §5.1.1). As preheating progresses in this regime, these modes become densely populated, leading to a smooth exponential increase in their corresponding number densities over many oscillations of the inflaton condensate.

As we have established, it is clear from eq. (5.1.20) that resonant particle production occurs when φ passes through zero. In order to determine the range of momenta over which field modes can be parametrically excited, we note that for small $|\varphi|$ one may approximate $\dot{\varphi} \approx -m_\phi \varphi_0$. From eq. (5.1.20) it then follows that[6]

$$0 \leq |\mathbf{k}|^2 \lesssim (g m_\phi \varphi \varphi_0)^{2/3} - g\varphi^2 \,, \tag{5.1.21}$$

from which it is evident that parametric resonance can occur across a continuous and broad range of momenta, allowing for extremely efficient reheating. Clearly, the condition given by eq. (5.1.21) places a constraint on the amplitude of the inflaton condensate, in particular, we observe that $|\varphi|$ must be smaller than $\left(\frac{m_\phi \varphi_0}{\sqrt{g}}\right)^{1/2}$. If we then treat the momenta $|\mathbf{k}|^2$ that satisfy eq. (5.1.21) as a function of $\varphi(t)$, then we find that the maximal range of momenta for which particle production can occur corresponds to where φ takes the value,

$$\varphi(t) \;=\; \varphi_* \approx \frac{1}{2}\left(\frac{m_\phi \varphi_0}{\sqrt{g}}\right)^{1/2} \approx \frac{1}{3}\varphi_0 q^{-1/4} \,. \tag{5.1.22}$$

Accordingly, one can estimate the maximal momentum for particles produced to be $|\mathbf{k}|_{\max} \approx \left(\frac{\sqrt{g} m_\phi \varphi_0}{2}\right)^{1/2}$. Given that $|\varphi| \lesssim 2\varphi_*$ throughout the main part of each interval of oscillation, the maximum value of momentum attainable will decrease, but should remain of the same order of magnitude as $|\mathbf{k}|_{\max}$. As such, one can further estimate that the typical range of momenta $k = |\mathbf{k}|$ for particles produced during broad resonance should be

$$0 \leq k \lesssim \frac{k_*}{2} \;=\; \frac{1}{2}\left(\sqrt{g} m_\phi \varphi_0\right)^{1/2} \;=\; \frac{m_\phi q^{1/4}}{\sqrt{2}} \,, \tag{5.1.23}$$

where $k_* = \left(\sqrt{g} m_\phi \varphi_0\right)^{1/2}$ is a measure for the maximum momentum scale attainable for produced particles during preheating. Since $q \gg 1$, we see from eq. (5.1.23) that it is possible for $k \gg m_\phi$, indicating the collective behaviour of the interaction been the inflaton condensate and the χ field. Indeed, this highlights the fact that parametric resonance is a highly non-perturbative phenomenon in the broad resonance regime, in which the coherent nature of inflaton condensate manifests as the collective interaction of many inflaton quanta in the production of χ particles.

[6] Here we neglect the bare mass of the χ field, as we assume $m_\chi \ll m_\phi$

To conclude this part of the analysis, we note that as φ approaches the point $\varphi = 0$, it will spend an amount of time Δt_* in the region where eq. (5.1.18) is violated and particle production can occur. Within this region, we have that $|\varphi| \lesssim \varphi_*$, and so its width can be estimated as $|\dot{\varphi}|\Delta t_* \sim 2\varphi_*$, and therefore particle production occurs within a time interval,

$$\Delta t_* \sim \frac{2\varphi_*}{|\dot{\varphi}|} \approx \left(\sqrt{g} m_\phi \varphi_0\right)^{-1/2} = k_*^{-1} \,. \tag{5.1.24}$$

As briefly mentioned earlier, we see then that particle production occurs only within short time intervals near the points where $\varphi = 0$. In particular, note that during that time interval $k_* \sim \sqrt{g}\varphi_* \approx m_\chi$, such that $\omega_* = \sqrt{k_*^2 + m_\chi^2 + g\varphi_*^2} \sim k_*$. We see therefore, that particle production occurs within a time of order the oscillation period of the χ field, i.e. $\Delta t_* \sim \omega_*^{-1}$. Due to the strong violation of adiabaticity, the number density $N_{\mathbf{k}}$ jumps rapidly and it cannot be interpreted as a physical quantity within these short time intervals, however, it fairly quickly settles down to an approximately constant value and is (approximately) adiabatically invariant in between such jumps.

As the production process progresses, the amplitude of the inflaton will decrease due to energy being transferred to the χ field, furthermore, the created particles will generate a contribution to the effective mass of φ, $m_{\phi,\text{eff}}^2 = m_\phi^2 + g\langle\hat{\chi}^2\rangle$ (where $\langle\hat{\chi}^2\rangle$ is the renormalised expectation value of $\hat{\chi}^2$). As a result of these effects, eventually $q = \frac{g\varphi_0^2}{4m_{\phi,\text{eff}}^2} \lesssim 1$ when $\varphi_0^2 \sim \langle\hat{\chi}^2\rangle$, at which point broad resonance will transition to the narrow regime. This has been estimated to occur after a time $t_{\text{BR}} \sim m_\phi^{-1} \ln\left(\frac{m_\phi}{g^{5/2}M_{\text{Pl}}}\right)$ [113], which for $\sqrt{g} = 5 \times 10^{-4}$ and $m_\phi = \sqrt{8\pi}10^{-6}M_{\text{Pl}}$, corresponds to $N_{\text{BR}} \sim \mathcal{O}(10)$ oscillations of the condensate. Preheating then proceeds in the regime of narrow resonance, during which the energy of the inflaton field φ decreases further, such that it becomes much less than that of the produced χ particles. The decay of the inflaton field will eventually cease, when the field φ is small enough that the process becomes inefficient. Indeed, the inflaton can only completely decay if the single particle decay processes, $\varphi \to \chi\chi$ and/or $\varphi \to \psi\psi$, remain possible. In fact, it has been shown that when the amplitude of the condensate $\varphi_0 < g^{-1/2}m_\phi$, then the decay of φ will terminate, corresponding to a time $t_{\text{end}} \sim m_\phi^{-1}\left(\sqrt{g}M_{\text{Pl}}/m_\phi\right)^{1/3}$. If, however $m_\phi < g^{7/2}M_{\text{Pl}}$, then the decay of φ occurs at a later stage, once reheating has entered the perturbative regime [113].

5.2 Preheating: a more realistic scenario

Our discussion of preheating in the narrow resonance regime so far, whilst enlightening on the key aspects of parametric resonance, is nonetheless a toy model. A more realistic treatment would require us to include the effects arising from the expansion of the universe, the backreactions of the produced χ particles on the inflaton condensate and their secondary interactions (rescatterings and decays). In the case of narrow resonance, these effects, of which we shall discuss briefly in turn, all conspire to suppress the efficiency of parametric resonance and evolve the system towards the perturbative regime. We shall closely follow the analyses of Refs. [98], [114] and [100] in this section, and we refer the reader to them for further details.

Referring back to eq. (2.1.24), we see that the expansion of the universe increases the rate at which the inflaton decays, due to the presence of the "Hubble friction" term $3H\dot{\varphi}$. Comparing $\Gamma_{\rm PR}$ with the effective decay rate of the inflaton field, $3H + \Gamma_\varphi$ (cf. eq. (2.2.5)), we see that parametric resonance can only take place if $m_\phi q \gtrsim 3H + \Gamma_\varphi$. Furthermore, the momenta of the previously created χ particles will be redshifted due to the expansion, thus shifting them out of the resonance band, which can be estimated to occur after a time interval of order $\Delta t \sim qH^{-1}$. This places an additional condition on $\Gamma_{\rm PR}$, indeed, efficient decay of inflatons can only occur if $qm_\phi \gtrsim \Delta t^{-1}$. We see therefore, that in order for parametric resonance to play a role in reheating, the following two conditions must be satisfied:

$$m_\phi q \;\gtrsim 3H + \Gamma_\varphi \,, \tag{5.2.1a}$$

$$q^2 m_\phi \;\gtrsim H \,. \tag{5.2.1b}$$

Observe that, in the case of narrow resonance ($q \ll 1$), eq. (5.2.1b) is a much stronger condition than eq. (5.2.1a).

The backreactions of the the produced χ particles on the inflaton condensate further inhibit the efficiency of parametric resonance. As more and more χ particles are created, they shift the effective mass of φ, causing the structure of the resonance bands to change, and causing previously produced particles to be removed from the resonance. Finally, the secondary interactions of rescattering and decays of the produced χ particles will generally cause them to change momenta and, as such, shift them out of the resonance band, thus further reducing the efficiency of preheating. In the case of broad resonance ($q \gg 1$), the situation is a little different. In fact, the process of parametric resonance in the broad regime is significantly more complex when one re-introduces the expansion of spacetime into the picture. This more complex resonant behaviour is referred to as *stochastic resonance*. The name arises from the fact that, for $\dot{a} \neq 0$, $\omega_{\mathbf{k}}^2(t) = \frac{\mathbf{k}^2}{a^2} + m_\chi^2 + g\varphi_0^2 \cos^2(m_\phi t)$; the additional time-dependence provided by the scale factor a causes the χ field to oscillate non-periodically. As such, the phases of the a given resonant field mode $\chi_{\mathbf{k}}$ at successive crossings of the inflaton field through zero, $\varphi(t_*) = 0$, are completely uncorrelated. This leads to stochastic changes in the number density $N_{\mathbf{k}}$, in which it can increase or decrease at any given moment, only increasing exponentially on average. As a result, one is unable to interpret this process in terms of a classical particle picture; it is purely quantum mechanical.

Perhaps counter-intuitively, the expansion of spacetime actually competes against the de-stabilising effects of backreactions and rescatterings during preheating in the broad resonance regime. To see why this is the case, we note that particle production will occur for all values of momentum $|\mathbf{k}|$ within a sphere of radius $\sim k_*$. As a result of spacetime expansion, the physical momenta of the created χ particles are red-shifted ($|\mathbf{k}| \propto a^{-1}$), however, the radius of the resonance sphere, decreases at a slower rate; indeed, we see from eq. (5.1.23), that it falls off as $\varphi_0^{1/2} \propto t^{-1/2} \propto a^{-3/4}$. Consequently, created particles are shifted away from the boundary of the sphere, moving towards its center, where they are able to participate in further particle production and thus enhancing the probability of decay into these modes. Recall that, in the case of narrow resonance, the efficiency of the resonance was affected by the subsequent re-scatterings of the produced χ particles, and their backreaction on the inflaton condensate. Indeed, they reduced efficiency by removing particles from the resonance bands.

The same effect also occurs in the broad resonance regime, causing created particles near the boundary of the resonance sphere to be pushed away from it. However, as we have just noted, the expansion of spacetime causes particles to naturally move away from this region, and so the efficiency of the resonant particle production is determined by the competing effects of Hubble expansion, backreaction and rescatterings. That being said, the backreaction is already expected to be small throughout the most part of broad resonance. Indeed, the presence of χ particles changes the effective mass of the inflaton condensate. This effect is negligible so long as $\Delta m_\phi^2 < m_\phi^2$, where in the mean-field approximation $\Delta m_\phi^2 = g\langle\hat{\chi}^2\rangle$. One can show that within this approximation $\Delta m_\phi^2 \approx \left(1+\epsilon\cos\left(\frac{2\sqrt{g}\varphi_0}{m_\phi}\cos(m_\phi t)\right)\right)\frac{\sqrt{g}N_\chi}{|\varphi(t)|} \sim \frac{\sqrt{g}N_\chi}{|\varphi(t)|}$ (where $\epsilon < 1$). During the broad resonance regime (where $\sqrt{g}\varphi_0 \gg m_\phi$) the term is highly oscillatory and therefore has only a negligible effect on the evolution of the condensate $\varphi(t)$ (due to the sign of $\epsilon\cos\left(\frac{2\sqrt{g}\varphi_0}{m_\phi}\cos(m_\phi t)\right)$ changing rapidly during each oscillation of $\varphi(t)$).[7] In order to fully neglect backreaction effects on the dynamics of $\varphi(t)$, it must also be the case that the energy density of the χ field is sub-dominant to that of the inflaton condensate. Since the produced χ particles are predominantly relativistic, we shall estimate the energy density of the χ field in terms of its kinetic energy (density), as such $\rho_\chi \sim \langle(\partial_\mu\hat{\chi})^2\rangle \sim k_*^2\langle\hat{\chi}^2\rangle$.[8] Moreover, $k_*^2\langle\hat{\chi}^2\rangle \sim 2m_\phi N_\chi q^{1/2}$, and so comparing this with the potential energy density of the inflaton condensate $V_\phi \sim \frac{2m_\phi^4}{g}q$, we see that ρ_χ is smaller than V_ϕ so long as $q > 1$.[9] Therefore, backreaction will only start to have a non-negligible impact on the dynamics of $\varphi(t)$, and hence on the resonance, once preheating enters the narrow resonance regime ($q \ll 1$), and one can reasonably ignore its effects before this point. Finally, we should note that in many models the scattering of χ particles off the inflaton condensate can suppress, and even shut-off resonance before backreaction effects start to become relevant, as has been confirmed by numerical simulations [219–222].

With all of these considerations taken into account, it is clear that preheating is far from a straightforward affair, its efficiency (and viability) proving to be sensitive to an interplay of various factors. Importantly, we have seen that there are multiple effects that conspire to shut-off the resonance, particularly the backreaction and rescatterings of the produced particles. Moreover, resonance can be blocked by the generation of thermal effective masses of the produced particles [223]. In an expanding universe the point in time t_{PR} at which parametric resonance will typically cease when eq. (5.2.1) is violated. This can be estimated to occur when $q \sim \mathcal{O}(0.1)$, i.e.

[7] Kofman, Linde and Starobinsky showed that this rapid oscillatory behaviour does not lead to the parametric resonance of the inflaton condensate, but that a non-resonant particle production is possible (see ref. [98] for further details).

[8] To estimate the kinetic energy (density), one can consider the ultra-relativistic limit of the (expectation value of the) corresponding Hamiltonian: $\langle H\rangle \approx \frac{1}{2}\left(\langle(\partial_t\hat{\chi})^2\rangle + \langle(\nabla\hat{\chi})^2\rangle\right)$. By considering the Fourier transform, one can then make the formal replacements $\partial_t \to -\omega_{\mathbf{k}} \sim -k_*$ and $\nabla \to \mathbf{k} \sim \mathbf{k}_*$, such that $\langle H\rangle \sim k_*^2\langle\hat{\chi}^2\rangle$.

[9] Here, we have noted that $\Delta m_\phi^2 \sim \frac{\sqrt{g}N_\chi}{|\varphi(t)|}$, and for the effects of backreaction to be neglected, we require that $\Delta m_\phi^2 < m_\phi^2$. This implies the condition that $N_\chi \lesssim \frac{m_\phi^2|\varphi(t)|}{\sqrt{g}}$.

when $\sqrt{g}\varphi_0 \sim m_\phi$. At this point $\varphi_0 \approx \frac{2\sqrt{6}M_{\rm Pl}}{3m_\phi t_{\rm PR}}$, corresponding to

$$t_{\rm PR} \sim \frac{2\sqrt{6}}{3}\frac{\sqrt{g}M_{\rm Pl}}{m_\phi^2} , \tag{5.2.2}$$

meaning that, for $\sqrt{g} = 5 \times 10^{-4}$ and $m_\phi = \sqrt{8\pi}10^{-6}M_{\rm Pl}$, parametric resonance will persist for $N_{\rm PR} \sim \frac{\sqrt{6}}{3\pi}\frac{\sqrt{g}M_{\rm Pl}}{m_\phi} \sim 26$ oscillations of φ.

Once preheating terminates, perturbative decays take over and the elementary analysis discussed in §2.2.2 can be used. The production of χ particles will persist until the inflaton has completely decayed, and once this has occurred, the system of particles will eventually settle down to a state of equilibrium, with a corresponding reheat temperature, at which point the standard HBB model becomes applicable. In fact, one expects that the produced particles will start to thermalise even during preheating, as soon as the number density becomes appreciable in size. Thermalisation of the produced χ particles is a rather complex phenomenon, however, it is an important part of the reheating process and must be analysed carefully if one is to obtain a sensible prediction for the reheat temperature. We shall discuss this in detail in the following chapter, where we shall consider the processes of preheating and thermalisation simultaneously through use of the density matrix formalism.

Chapter 6

Thermalisation during preheating

At the end of the previous chapter we introduced the notion of thermalisation during reheating, noting that it is an important and essential part of the process, in order for inflationary theory to reconnect with the standard HBB model. Indeed, the elementary analysis and prediction of the reheating temperature carried out in §2.2.1 is severely lacking, in particular, due to its assumption that thermalisation happens almost instantaneously. In a realistic scenario, this generically would not be the case[1], in fact, thermalisation can only proceed so long as the rate of Hubble expansion is less than the interaction rate between the created particles.

In the early stages of reheating, non-perturbative processes can dominate, leading to explosive particle production. This occurs on time-scales much shorter than those needed for the produced particles to thermalise. Within the chaotic inflation scenario, preheating begins in the broad resonance regime, wherein the particle number grows exponentially across wide bands of momentum [98, 100]. As was discussed in §5.2, due to the decay of the amplitude of the inflaton condensate, and the expansion of the universe, the dynamics eventually transition to the narrow resonance regime, where the growth of particle number is restricted to ever narrower bands of momentum.

At this stage, the backreaction from the created particles and the Hubble expansion continue to conspire to reduce the efficiency of the parametric resonance [97]. Specifically, the backreaction alters the structure of the resonance bands, the cosmological expansion redshifts the momenta of the produced particles and both effects cause the created particles to be shifted out of the resonance bands. The resonance can also be blocked by the onset of effective thermal masses [223]. In any case, as the parametric resonance becomes increasingly narrow and inefficient, the dynamics inevitably transition to the perturbative regime [98, 113], and it is during this final stage of reheating that scatterings redistribute the occupancy of the momentum

[1] If reheating is predominantly perturbative, and particle production occurs slowly enough, then thermalisation is almost instantaneous [100].

modes, leading to the eventual kinetic equilibriation of the primordial plasma.

While the time-scales for preheating are much shorter than those needed for thermalisation, the processes driving that thermalisation are still relevant during the preheating phase. In particular, semi-classical lattice simulations have shown that sufficiently large self-interactions of the produced particles can suppress or prevent the resonant particle production [222]. The early thermal history of the universe depends strongly on how the primordial plasma attained kinetic equilibrium [224]. This has motivated the extensive study of thermalisation both in perturbative reheating [225–231] and after the phase of preheating [224, 232–239], and the relevant relaxation rates can be calculated by means of thermal quantum field theory [229, 240, 241].

With an aim of better understanding the thermalisation process during preheating, in this chapter we will recast the problem of preheating in the density matrix formalism [242] and study the impact of scatterings on this phase by means of a system of quantum Boltzmann equations. These Boltzmann equations are able to go beyond the usual mode-function analysis of preheating, based on the Mathieu equation, by accounting simultaneously for both the resonant particle production and the collisional processes.

As we shall show, the resonant particle production proceeds via the population of pair correlations of the form $M_{\mathbf{k}} \sim \langle \hat{a}_{-\mathbf{k}}(t)\hat{a}_{\mathbf{k}}(t)\rangle$ and $M^*_{\mathbf{k}} \sim \langle \hat{a}^\dagger_{\mathbf{k}}(t)\hat{a}^\dagger_{-\mathbf{k}}(t)\rangle$, requiring one to solve the coupled system of Boltzmann equations for the number density $N_{\mathbf{k}} \sim \langle \hat{a}^\dagger_{\mathbf{k}}(t)\hat{a}_{\mathbf{k}}(t)\rangle$ and these "particle-anti-particle" correlations. The presence of such pair correlations is expected in the absence of time translational invariance, as was identified in the context of preheating in Refs. [243–245] by means of the Schwinger-Keldysh closed-time-path and Kadanoff-Baym [246] formalisms of non-equilibrium quantum field theory [247, 248] (see also Refs. [249, 250]). Therein, the number density must be carefully defined [251], and the pair correlations can then be accounted for in the non-equilibrium Green's functions by working with coherent quasi-particle approximations [252–254] or directly in terms of the operator algebra by means of the so-called interaction-picture approach [255, 256]. Particle-anti-particle pair correlations have also been studied in the density matrix formalism in the context of neutrino kinetics, where they may play a role in core-collapse supernovae [257–261].

As will be demonstrated in this chapter, the pair correlations in fact play the pivotal role in mediating the particle production, and without them no particle production occurs. This leads to a powerful generalization of the previous observation by Morikawa and Sasaki [262] that small perturbations to such a system would destroy coherences between particle and anti-particle states. Indeed, it follows that any processes that cause such pair correlations to decohere will suppress, or shut off, the resonant particle production.

6.1 A density matrix approach to preheating

Following the introductory discussion to this chapter, we now proceed to construct a toy model of preheating. In doing so, we will account for the leading-order scattering processes of particles produced during this phase, thereby enabling one to study their impact on the preheating process. For this, we need to determine the behaviour of a statistical ensemble of relativistic particles as they evolve towards a state of thermal equilibrium. This requires a set of evolution equations describing a relativistic quantum many-body system.[2] The density matrix formalism, in which we describe the state of the system in terms of a density operator $\hat{\rho}(t)$, furnishes us with the means to derive such a set of equations. Ultimately, this will enable us to derive a self-consistent set of quantum Boltzmann equations that describe the evolution of the (scalar) particle number densities throughout preheating. With this in mind, in the following subsections we will elaborate on the details of a density matrix description of preheating.

6.1.1 Canonical quantisation on a time-dependent background

Let us consider a toy-model preheating theory, in which we assume single-field chaotic inflation has occurred, and that the inflaton $\phi(x) \coloneqq \phi(t, \mathbf{x})$ is coupled to some scalar field[3] $\chi(x) \coloneqq \chi(t, \mathbf{x})$ via a potential of the form $V(\varphi, \chi) = \frac{g}{4}\varphi^2\chi^2$. Throughout this analysis, we shall assume that the rate of post-inflationary expansion is small relative to the rate of particle production and thermalisation, such that the universe is approximately Minkowski at the end of inflation, i.e., the scale-factor $a(t) \approx \text{const.}$[4]

[2] Such a system will generically be in a mixed state throughout its evolution, necessitating a density matrix description.

[3] We restrict our attention to scalar fields to the avoid additional complications of dealing with spinor fields.

[4] One can choose co-moving coordinates such that $a(t) \approx 1$.

Although a toy model, it captures the salient features physical effects that we wish to study.

We therefore consider a Lagrangian of the form given by eq. (2.2.1), i.e.

$$\mathscr{L} \;=\; \mathscr{L}_0^{\phi} + \mathscr{L}_0^{\chi} + \mathscr{L}_{\text{int}} \;=\; -\frac{1}{2}\partial^{\mu}\phi\partial_{\mu}\phi - \frac{1}{2}m_{\phi}^2\phi^2 - \frac{1}{2}\partial^{\mu}\chi\partial_{\mu}\chi - \frac{1}{2}m_{\chi}^2\chi^2 - \frac{g}{4}\phi^2\chi^2\,, \tag{6.1.1}$$

where $\mathscr{L}_{\text{int}} = -\frac{g}{4}\phi^2\chi^2$ is the interaction Lagrangian, m_ϕ and m_χ are the bare masses of the inflaton $\phi = \phi(x)$ and the scalar field $\chi = \chi(x)$, and g is the coupling between ϕ and χ.[5] Given this, we can construct the corresponding Hamiltonian density from the Legendre transform of eq. (6.1.1) to give,

$$\mathscr{H} \;=\; \mathscr{H}_0^{\phi} + \mathscr{H}_0^{\chi} + \frac{g}{4}\phi^2\chi^2\,, \tag{6.1.2}$$

where $\mathscr{H}_0^{\phi} = \pi_\phi\dot{\phi} - \mathscr{L}_0^{\phi}$ and $\mathscr{H}_0^{\phi} = \pi_\chi\dot{\chi} - \mathscr{L}_0^{\chi}$ are the free-field Hamiltonian's of ϕ and χ, with $\pi_\phi = \frac{\partial\mathscr{L}}{\partial\dot{\phi}}$ and $\pi_\chi = \frac{\partial\mathscr{L}}{\partial\dot{\chi}}$ their respective conjugate momenta.

To proceed, we make a mean-field approximation, describing the inflaton condensate by the time-dependent background field $\varphi(t) \coloneqq \langle\hat{\phi}(t,\mathbf{x})\rangle$ (and assuming that $\langle\hat{\chi}(t,\mathbf{x})\rangle = 0$), where t is the cosmic time and the expectation value is with respect to a translationally invariant vacuum state. This homogeneous background $\varphi(t)$ corresponds to a condensate of zero-momentum inflaton quanta, coherently oscillating with frequency $\omega_k = m_\phi$. As discussed in §2.2.1, this is a reasonable approximation to make, since by the end of inflation the inflaton field ϕ has condensed, such that a large number of its quanta reside in the zero-momentum mode.[6] The evolution of the background value of the inflaton field $\varphi(t)$ is described by the following equations of motion:

$$\ddot{\varphi}(t) + 3H(t)\dot{\varphi}(t) + m_\phi^2\varphi(t) \;=\; 0 \tag{6.1.3a}$$

$$H^2(t) \;=\; \frac{1}{3M_{\text{Pl}}^2}\left[\frac{1}{2}\dot{\varphi}^2(t) + \frac{1}{2}m_\phi^2\varphi^2(t)\right]. \tag{6.1.3b}$$

[5] We note that quadratic inflation, where the inflaton potential is of the form $V(\phi) = \frac{1}{2}m_\phi^2\phi^2$ has been all but ruled out by observational data (essentially, the model predicts too large a tensor-to-scalar ratio [263]). However, here we are only interested in the behaviour of the inflaton at the end of inflation as it approaches the minimum of its potential. One expects that near the minimum of $V(\phi)$ it will be quadratic in form.

[6] As a result of the large occupancy in a single mode, it is reasonable to describe the inflaton in terms of classical scalar field.

These equations of course neglect the dissipative effects of particle production and its backreaction on the inflaton condensate, which increases the rate of decay. However, as remarked upon in §5.2, this backreaction is expected to be subdominant in the early stages of preheating that we study here. As such, one obtains an asymptotic solution to eq. (6.1.3) of the form (cf. §5.1)

$$\varphi(t) \simeq \varphi_0(t)\cos\left(m_\phi t\right)\,, \qquad \varphi_0(t) \simeq \frac{2\sqrt{6}M_{\mathrm{Pl}}}{3m_\phi t}\,, \tag{6.1.4}$$

which is valid during the post-inflationary reheating phase. Furthermore, recalling our assumption that the rate of post-inflationary expansion is small relative to the rate of particle production and thermalisation, i.e. $\dot{a} \approx 0$, we can take $\varphi_0(t) \approx \text{const.}$

Since we are primarily interested in the dynamics of the χ field, we use eq. (6.1.1) to construct an effective Lagrangian density for χ, in the presence of a time-dependent background (corresponding to the background value $\varphi(t)$ of the inflaton field)

$$\begin{aligned}\mathscr{L}_{\mathrm{eff}}(x) &= -\frac{1}{2}\partial^\mu\chi(x)\partial_\mu\chi(x) - \frac{1}{2}m_\chi^2\chi^2(x) - \frac{g}{4}\varphi^2(t)\chi^2(x)\\ &= -\frac{1}{2}\partial^\mu\chi(x)\partial_\mu\chi(x) - \frac{1}{2}m_{\mathrm{eff}}^2(t)\chi^2(x)\,.\end{aligned} \tag{6.1.5}$$

We see then that the time-dependence of the background manifests in the Lagrangian as a time-dependent effective mass for the χ field

$$m_{\mathrm{eff}}^2(t) = m_\chi^2 + \frac{g}{2}\varphi^2(t)\,. \tag{6.1.6}$$

Note that we have omitted the interactions between the χ and inflaton fluctuations, which give rise to perturbative decays, however, these play a subdominant role in the particle production.[7] We see that eq. (6.1.5) bears resemblance to the free theory of a scalar boson. By this, we mean that χ particles are produced non-perturbatively by virtue of the χ field interacting with the time-dependent background, but do not mutually interact in anyway; once they are produced, they can propagate freely. In reality, however, the theory implicitly contains interactions, since χ is coupled to a

[7] To see this, note that one can estimate both of the decay rates corresponding to the interactions $\mathscr{L}_{\mathrm{int}} \supset -\frac{g}{2}\varphi(t)\phi^2\chi^2$ and $\mathscr{L}_{\mathrm{int}} \supset -\frac{g}{4}\phi^2\chi^2$ to scale as $\Gamma_{\phi\to\chi\chi} \sim \frac{g^2\varphi_0^2}{32\pi m_\phi}$ and $\Gamma_{\phi\phi\to\chi\chi} \sim \frac{g^2\varphi_0^2}{256\pi m_\phi}$ (cf. §2.2.1). Moreover, the resonant production rate can be estimated as $\Gamma_* := 1/\Delta t_* \sim (\sqrt{g}m_\phi\varphi_0)^{1/2}$. Therefore, so long as $\left\{\frac{1}{32\pi}, \frac{1}{256\pi}\right\}\left(g^7\varphi_0^6/m_\phi^6\right)^{1/4} \ll 1$, then $\Gamma_* \gg \{\Gamma_{\phi\to\chi\chi}\,, \Gamma_{\phi\phi\to\chi\chi}\}$. In our analysis we take g to be $\mathcal{O}(10^{-7})$, furthermore, at the start of preheating $\varphi_0 \sim 0.1M_{\mathrm{Pl}}$. Therefore, if we take $m_\phi \sim 10^{-5}\varphi_0$, then it is clear that the resonant particle production dominates.

time-dependent background scalar field φ. As such, we shall refer to eq. (6.1.5) as the "collisionless" theory for χ, i.e. neglecting any self-interactions of the χ field.

The effective equation of motion for the χ field is then derived the usual way, by inserting eq. (6.1.5) into the Euler-Lagrange equation, and corresponds to an effective Klein-Gordon equation,

$$\left[\Box - m_{\text{eff}}^2(t)\right]\chi(x) \; = \; 0\,. \tag{6.1.7}$$

Furthermore, from eq. (6.1.5), we can construct the associated Hamiltonian density,

$$\mathscr{H}_0^{\text{eff}}(x) \; = \; \frac{1}{2}\pi_\chi^2(x) + \frac{1}{2}\big(\nabla\chi(x)\big)^2 + \frac{1}{2}m_{\text{eff}}^2(t)\chi^2(x)\,. \tag{6.1.8}$$

Before proceeding, we note that one can constrain the range of values the the coupling g (in eq. (6.1.6)) can take, such that: (a) the effective inflaton potential $V(\varphi) = \frac{1}{2}m_\phi^2\varphi^2$ is not destablised by quantum corrections, and (b) that inflation is solely driven by the inflaton. Indeed, one can obtain an upper limit for g by considering the interaction at the QFT level. The interaction potential $V(\varphi,\chi) = \frac{g}{4}\varphi^2\chi^2$, will generate radiative corrections to the inflaton effective potential $V(\varphi)$ of the form $\sim \frac{g^2\varphi^4}{64\pi^2}\ln\big(\sqrt{\frac{g}{2}}\varphi/\mu\big)$, where μ is an arbitrary subtraction scale (introduced via renormalisation, cf. §3.1.1). For large field inflation, the scale at which cosmological fluctuations are observed corresponds to inflaton values $\varphi \sim 4M_{\text{Pl}}$ [224]. As such, if we take the (bare) inflaton mass to be $m_\phi \sim 10^{-6}M_{\text{Pl}}$ (which for simple cases, such as this one, reproduces the correct amplitude for CMB anisotropies [224]), then the radiative corrections from the interaction potential do not alter the effective inflaton potential as long as $g \lesssim 10^{-5}$. Moreover, one can also provide a lower limit for g by noting that, during inflation, the χ field will have an effective mass (squared) given by eq. (6.1.6), which scales as $\frac{g}{2}\varphi^2 \sim \frac{g}{2}M_{\text{Pl}}^2$. Now, for scenarios in which (at least) two scalar fields φ and χ are present during inflation, then its latest stage will be driven by the lightest scalar [264]. For a consistent setting, where the end of inflation is driven by φ (i.e. inflation is purely due to φ), it must be that $\frac{g}{2}M_{\text{Pl}}^2 > m_\phi^2$ which implies that $g > 10^{-12}$.

Continuing with the construction of a toy model of preheating, we now canonically quantise the χ field in the presence of a time dependent background, $\varphi(t)$, by specifying that $\hat{\chi}(t,\mathbf{x})$ and its canonically conjugate momentum $\hat{\pi}_\chi(t,\mathbf{x}) = \frac{\partial\mathscr{L}_{\text{eff}}}{\partial(\partial_t\chi)}$ (where

$\partial_t \coloneqq \frac{\partial}{\partial t}$) satisfy the following equal-time commutation relations,

$$\left[\hat{\chi}(t,\mathbf{x}),\hat{\chi}(t,\mathbf{y})\right] \;=\; \left[\hat{\pi}_\chi(t,\mathbf{x}),\hat{\pi}_\chi(t,\mathbf{y})\right] \;=\; 0\,, \tag{6.1.9a}$$

$$\left[\hat{\chi}(t,\mathbf{x}),\hat{\pi}_\chi(t,\mathbf{y})\right] \;=\; i\delta^{(3)}\left(\mathbf{x}-\mathbf{y}\right)\,. \tag{6.1.9b}$$

To account for the thermalisation process, driven by the collisional processes between the produced χ particles, we introduce a self-interaction potential, the corresponding interactions of which are expected to be perturbative. As such, it is possible to treat the χ field within the framework of perturbative QFT. Thus we may work in the (modified) interaction picture, and expand $\hat{\chi}(x)$ in terms of its Fourier modes,

$$\hat{\chi}(t,\mathbf{x}) \;=\; \int_{\mathbf{k}}\left[\hat{a}(t,\mathbf{k})e^{+i\mathbf{k}\cdot\mathbf{x}}+\hat{a}^\dagger(t,\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{x}}\right], \tag{6.1.10}$$

where $\hat{a}^\dagger(t,\mathbf{k}) \coloneqq \hat{a}^\dagger_{\mathbf{k}}(t)$ and $\hat{a}(t,\mathbf{k}) \coloneqq \hat{a}_{\mathbf{k}}(t)$ create and destroy quanta of instantaneous frequency $\omega_{\mathbf{k}}(t)$ given by

$$\omega^2_{\mathbf{k}}(t) \;=\; \mathbf{k}^2+m^2_{\text{eff}}(t) \tag{6.1.11}$$

Proceeding with our analysis, we note that the creation and annihilation operators $\hat{a}^\dagger_{\mathbf{k}}(t)$ and $\hat{a}_{\mathbf{k}}(t)$ (respectively) can be expressed in terms of $\hat{\chi}(t,\mathbf{x})$ and $\hat{\pi}_\chi(t,\mathbf{x})$,

$$\hat{a}_{\mathbf{k}}(t) \;=\; \int_{-\infty}^{+\infty}\mathrm{d}^3\mathbf{x}\;e^{-i\mathbf{k}\cdot\mathbf{x}}\left[\sqrt{\frac{\omega_{\mathbf{k}}(t)}{2}}\hat{\chi}(t,\mathbf{x})+\frac{i}{\sqrt{2\omega_{\mathbf{k}}(t)}}\hat{\pi}_\chi(t,\mathbf{x})\right], \tag{6.1.12a}$$

$$\hat{a}^\dagger_{\mathbf{k}}(t) \;=\; \int_{-\infty}^{+\infty}\mathrm{d}^3\mathbf{x}\;e^{+i\mathbf{k}\cdot\mathbf{x}}\left[\sqrt{\frac{\omega_{\mathbf{k}}(t)}{2}}\hat{\chi}(t,\mathbf{x})-\frac{i}{\sqrt{2\omega_{\mathbf{k}}(t)}}\hat{\pi}_\chi(t,\mathbf{x})\right]. \tag{6.1.12b}$$

The creation and annhilation operators have mass dimension $-3/2$, and satisfy the canonical equal-time commutation relations

$$\left[\hat{a}_{\mathbf{p}}(t),\hat{a}_{\mathbf{k}}(t)\right] \;=\; \left[\hat{a}^\dagger_{\mathbf{p}}(t),\hat{a}^\dagger_{\mathbf{k}}(t)\right] \;=\; 0\,, \tag{6.1.13a}$$

$$\left[\hat{a}_{\mathbf{p}}(t),\hat{a}^\dagger_{\mathbf{k}}(t)\right] \;=\; (2\pi)^3\delta^{(3)}\left(\mathbf{p}-\mathbf{k}\right)\,. \tag{6.1.13b}$$

This convention is advantageous in the sense that the right-hand side of eq. (6.1.13b) is time-independent. Moreover, $\hat{\chi}(x)$ and $\hat{\pi}_\chi(x)$ are the canonical variables of the system, and therefore (by definition) are not explicitly dependent on time, i.e. $\partial_t\hat{\chi}=0$,

$\partial_t \hat{\pi}_\chi = 0$ (that is, they are held fixed with respect to partial time derivatives).[8] However, this is not true for the corresponding creation and annihilation operators: the time dependence of the effective (instantaneous) frequency $\omega_{\mathbf{k}}(t)$ generates an explicit time dependence for $\hat{a}_{\mathbf{k}}(t)$ and $\hat{a}^\dagger_{\mathbf{k}}(t)$. There full time-evolution is governed by the (interaction picture) Heisenberg equations,

$$\partial_t \hat{a}_{\mathbf{k}}(t) = \frac{\dot{\omega}_{\mathbf{k}(t)}}{2\omega_{\mathbf{k}}(t)} \hat{a}^\dagger_{-\mathbf{k}}(t)\,, \tag{6.1.14a}$$

$$\partial_t \hat{a}^\dagger_{\mathbf{k}}(t) = \frac{\dot{\omega}_{\mathbf{k}(t)}}{2\omega_{\mathbf{k}}(t)} \hat{a}_{-\mathbf{k}}(t)\,. \tag{6.1.14b}$$

Since we are working in the (modified) interaction picture it follows that the full time evolution of a given operator $\hat{\mathcal{O}}(t)$ (in general, with explicit time dependence) is governed by the Heisenberg equation of motion,

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathcal{O}}(t) \coloneqq \dot{\hat{\mathcal{O}}}(t) = i\left[\hat{H}_0(t), \hat{\mathcal{O}}(t)\right] + \partial_t \hat{\mathcal{O}}(t)\,, \tag{6.1.15}$$

where where $\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathcal{O}}(t) \coloneqq \dot{\hat{\mathcal{O}}}(t)$ denotes the *total* time derivative of an operator $\hat{\mathcal{O}}(t)$ (taking into account both its implicit and explicit time dependence), and

$$\hat{H}_0(t) = \int_{-\infty}^{+\infty} \mathrm{d}^3\mathbf{x}\, \hat{\mathscr{H}}_0^{\mathrm{eff}}(t, \mathbf{x}) \tag{6.1.16}$$

is the effective Hamiltonian operator of the "collisionless" theory (i.e. the Legendre transform of eq. (6.1.5)), where $\hat{\mathscr{H}}_0^{\mathrm{eff}}(t, \mathbf{x})$ is given by the quantised version of eq. (6.1.8).

Hence, the full time evolutions of $\hat{a}_{\mathbf{k}}(t)$ and $\hat{a}^\dagger_{\mathbf{k}}(t)$ are given by,

$$\dot{\hat{a}}_{\mathbf{k}}(t) = i\left[\hat{H}_0(t), \hat{a}_{\mathbf{k}}(t)\right] + \frac{\dot{\omega}_{\mathbf{k}}(t)}{2\omega_{\mathbf{k}}(t)} \hat{a}^\dagger_{-\mathbf{k}}(t)\,, \tag{6.1.17a}$$

$$\dot{\hat{a}}_{\mathbf{k}}(t) = i\left[\hat{H}_0(t), \hat{a}^\dagger_{\mathbf{k}}(t)\right] + \frac{\dot{\omega}_{\mathbf{k}(t)}}{2\omega_{\mathbf{k}}(t)} \hat{a}_{-\mathbf{k}}(t)\,. \tag{6.1.17b}$$

[8] Their time evolution is generated by the Hamiltonian of the system (note that this is true even in cases where the Hamiltonian has explicit time-dependence, as it does here). This is in analogy to the Hamiltonian formulation of classical mechanics, where the EOM of the canonical coordinates (p, q) are given by $\frac{\mathrm{d}q}{\mathrm{d}t} = \{q, H\}$, $\frac{\mathrm{d}p}{\mathrm{d}t} = \{p, H\}$, with $\{u, H\} = \frac{\partial u}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial u}{\partial p}\frac{\partial H}{\partial q}$ the Poisson bracket of some quantity u with respect to the Hamiltonian H. In general, functions u of the canonical variables will have explicit time dependence, i.e. $u = u(q, p, t)$, in which case their EOM are given by $\frac{\mathrm{d}u}{\mathrm{d}t} = \{u, H\} + \frac{\partial u}{\partial t}$. See, e.g., ref. [265] for further details.

From Eqs. (6.1.12a) and (6.1.12b), we can subsequently express the conjugate momentum $\hat{\pi}_\chi(x)$ in terms of $\hat{a}_{\mathbf{k}}(t)$ and $\hat{a}^\dagger_{\mathbf{k}}(t)$ as,

$$\hat{\pi}_\chi(t,\mathbf{x}) \;=\; i\int_{\mathbf{k}} \omega_{\mathbf{k}}(t)\left[\hat{a}^\dagger_{\mathbf{k}}(t)e^{-i\mathbf{k}\cdot\mathbf{x}} - \hat{a}_{\mathbf{k}}(t)e^{+i\mathbf{k}\cdot\mathbf{x}}\right]. \tag{6.1.18}$$

It can be shown that the expression for $\hat{\pi}_\chi(x)$ given above is consistent with the fact that $\hat{\pi}_\chi(x) = \dot{\hat{\chi}}(t,\mathbf{x}) = \frac{\mathrm{d}}{\mathrm{d}t}\hat{\chi}(t,\mathbf{x})$, and that it also adheres to the canonical commutation relations, Eqs. (6.1.9a) and (6.1.9b).[9] Furthermore, inserting our expressions for $\hat{\chi}$ and $\hat{\pi}_\chi$ (Eqs. (6.1.10) and (6.1.18) respectively) into (6.1.16), and normal-ordering the operators[10], we can express the effective "collisionless" (normal-ordered) Hamiltonian $\hat{H}_0(t)$ in terms of the creation and annihilation operators,

$$\hat{H}_0(t) \;=\; \int_{-\infty}^{+\infty}\frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\,\omega_{\mathbf{k}}(t)\hat{a}^\dagger_{\mathbf{k}}(t)\hat{a}_{\mathbf{k}}(t)\,. \tag{6.1.19}$$

Accordingly, it is found that,

$$\left[\hat{H}_0(t),\hat{a}_{\mathbf{k}}(t)\right] \;=\; -\omega_{\mathbf{k}}(t)\hat{a}_{\mathbf{k}}(t)\,, \tag{6.1.20a}$$

$$\left[\hat{H}_0(t),\hat{a}^\dagger_{\mathbf{k}}(t)\right] \;=\; +\omega_{\mathbf{k}}(t)\hat{a}^\dagger_{\mathbf{k}}(t)\,. \tag{6.1.20b}$$

Such that the full time evolutions of $\hat{a}_{\mathbf{k}}(t)$ and $\hat{a}^\dagger_{\mathbf{k}}(t)$ given in eq. (6.1.17) can be written explicity as,

$$\dot{\hat{a}}_{\mathbf{k}}(t) \;=\; -i\omega_{\mathbf{k}}(t)\hat{a}_{\mathbf{k}}(t) + \frac{\dot{\omega}_{\mathbf{k}}(t)}{2\omega_{\mathbf{k}}(t)}\hat{a}^\dagger_{-\mathbf{k}}(t)\,, \tag{6.1.21a}$$

$$\dot{\hat{a}}^\dagger_{\mathbf{k}}(t) \;=\; +i\omega_{\mathbf{k}}(t)\hat{a}^\dagger_{\mathbf{k}}(t) + \frac{\dot{\omega}_{\mathbf{k}(t)}}{2\omega_{\mathbf{k}}(t)}\hat{a}_{-\mathbf{k}}(t)\,. \tag{6.1.21b}$$

6.1.2 Introducing thermalisation into the picture

The process of thermalisation is driven by collisional processes between the χ particles, and therefore we need to incorporate a means whereby they can mutually interact. With this in mind, we assume that such collisional processes arise from a

[9] Note that the corresponding EOM for the canonical variables $\hat{\chi}(x)$ and $\hat{\pi}_\chi(x)$ are given by $\frac{\mathrm{d}}{\mathrm{d}t}\hat{\chi}(x) = i\big[\hat{H}_0(t),\hat{\chi}(x)\big]$ and $\frac{\mathrm{d}}{\mathrm{d}t}\hat{\pi}_\chi(x) = i\big[\hat{H}_0(t),\hat{\pi}_\chi(x)\big]$, respectively.

[10] We are safe to do this here as we are considering the flat (Minkowski) space limit, i.e. the scale factor $a(t) \approx \text{const}$.

self-interaction potential for the χ field, of the form

$$\mathscr{L}_{\rm int}(x) \;=\; -\frac{\lambda}{4!}\,\chi^4(x)\,, \tag{6.1.22}$$

where λ is a (dimensionless) coupling constant. Transitioning to the quantum regime, via canonical quantisation, this amounts to introducing an interaction Hamiltonian operator of the form,

$$\hat{H}_{\rm int}(t) \;=\; \int_{-\infty}^{+\infty} {\rm d}^3\mathbf{x}\;\frac{\lambda}{4!}\,\hat{\chi}^4(t,\mathbf{x})\,. \tag{6.1.23}$$

As such, the effective Hamiltonian becomes,

$$\begin{aligned}\hat{H}(t) &= \hat{H}_0(t) + \hat{H}_{\rm int}(t)\\ &= \int_{-\infty}^{+\infty} {\rm d}^3\mathbf{x}\left[\frac{1}{2}\hat{\pi}_\chi^2(t,\mathbf{x}) + \frac{1}{2}\big(\nabla\hat{\chi}(t,\mathbf{x})\big)^2 + \frac{1}{2}m_{\rm eff}^2(t)\hat{\chi}^2(t,\mathbf{x}) + \frac{\lambda}{4!}\,\hat{\chi}^4(t,\mathbf{x})\right].\end{aligned} \tag{6.1.24}$$

At this point, we note that the additional effective interaction $\mathscr{L}_{\rm int}^{\rm eff} \supset -\frac{g}{2}\varphi(t)\phi\chi^2$, generated from the coupling between ϕ and χ, gives rise to ϕ mediated two-to-two scatterings of χ particles.[11] A legitimate concern is that these processes may dominate over those arising from eq. (6.1.22). To determine whether this is the case or not, we need to calculate the relevant two-to-two scattering cross-sections for each of the interactions (which for clarity we denote as σ_λ and σ_g). Having said this, we can already obtain an order-of-magnitude estimate. Indeed, one expects that $\sigma_\lambda \sim \frac{\lambda^2}{s}$, whereas $\sigma_g \sim \frac{g^4\varphi^4}{s^3}$, where $\sqrt{s}$ is the centre-of-mass energy of the scattering process. In order for the scattering processes arising from eq. (6.1.22) to dominate over the ϕ-mediated processes, we require that $\sigma_\lambda \gg \sigma_g$, which translates to the condition $\lambda^2 \gg \frac{g^4\varphi^4}{s^2}$. In intervals of adiabacity, the centre-of-mass energy scales as $\sqrt{s} \sim 2g\varphi^2$, and as such this places the bound $\lambda \gg g/2$. In non-adiabatic intervals, the centre-of-mass energy scales as $\sqrt{s} \sim 2m_\phi^2$, and so we require that $\lambda \gg \frac{g^2\varphi^2}{4m_\phi^2}$. Note, however, in this regime $|\varphi| \ll 1$ and so this does not really place a constraint on λ (other than it should be non-zero). Having determined constraints on the value of λ, let us now carry out a rigorous calculation of the relevant cross-sections so as to provide a confirmation of our expectations. We start by considering the zero temperature limit of the cross-section for two-to-two scattering processes, which will provide us with a

[11] There is also the effective interaction $\mathscr{L}_{\rm int}^{\rm eff} \supset -\frac{g}{4}\phi^2\chi^2$, however, two-to-two scattering processes arising from this interaction are loop suppressed, and thus will be sub-dominant to the tree-level contributions from $\mathscr{L}_{\rm int}^{\rm eff} \supset -\frac{\lambda}{4!}\chi^4$ and $\mathscr{L}_{\rm int}^{\rm eff} \supset -\frac{g}{2}\varphi(t)\phi\chi^2$.

leading order estimate for how it scales at finite temperature. Having determined its generic form, we shall then apply the result to the relevent interactions.

In general, the differential cross-section for two-to-n scattering processes is determined via the following formula [121]:

$$\mathrm{d}\sigma \;=\; \frac{1}{(2E_1)(2E_2)|\mathbf{v}_1-\mathbf{v}_2|}|\mathcal{M}|^2\mathrm{d}\Pi^{(n)}\,, \tag{6.1.25}$$

where $E_i = \sqrt{m_i^2+\mathbf{p}_i^2}$ $(i=1,2)$ are the energies of the incoming particles (with m_i and $\mathbf{p}_i$ their corresponding masses and momenta), $\mathbf{v}_1-\mathbf{v}_2$ is their relative velocity, $\mathcal{M}$ is the corresponding S-matrix element for the interaction, and $\mathrm{d}\Pi^{(n)}$ is the n-body Lorentz invariant phase space (LIPS) measure, given by [121]

$$\mathrm{d}\Pi^{(n)} \;\coloneqq\; (2\pi)^4\delta^{(4)}\Big(p_1+p_2-\sum_j p_j\Big)\prod_j \frac{\mathrm{d}^3\mathbf{p}_j}{(2\pi)^3}\frac{1}{2E_j}\,, \tag{6.1.26}$$

where $p_i = (E_i,\,\mathbf{p}_i)$ are the 4-momenta of the particles, and the product and sum are taken over the final states j. Since they will prove useful later, here we also define three Lorentz scalars s, t and u, the so-called Mandelstam variables (where we consider processes of the form $p_1+p_2\to p_3+p_4$),

$$s \;\coloneqq\; -(p_1+p_2)^2 \;=\; -(p_3+p_4)^2\,, \tag{6.1.27a}$$

$$t \;\coloneqq\; -(p_1-p_3)^2 \;=\; -(p_2-p_4)^2\,, \tag{6.1.27b}$$

$$u \;\coloneqq\; -(p_1-p_4)^2 \;=\; -(p_2-p_3)^2\,, \tag{6.1.27c}$$

which satisfy $s+t+u=\sum_{i=1}^4 m_i^2$. In particular s is the (square of the) centre-of-mass energy.

Returning to eq. (6.1.26), for two-to-two scattering processes in the CM frame $p_1+p_2\to p_3+p_4$, where $\mathbf{p}_1 = -\mathbf{p}_2$, such that $E_1+E_2 = \sqrt{-(p_1+p_2)^2} = \sqrt{s}$, the LIPS measure has the form

$$\begin{aligned}\mathrm{d}\Pi^{(2)} &= \frac{\mathrm{d}^3\mathbf{p}_3}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{p}_4}{(2\pi)^3}\frac{1}{2E_3}\frac{1}{2E_4}(2\pi)^4\delta^{(4)}(p_1+p_2-p_3-p_4)\\ &= \frac{\mathrm{d}^3\mathbf{p}_3}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{p}_4}{(2\pi)^3}\frac{1}{2E_3}\frac{1}{2E_4}(2\pi)^4\delta(\sqrt{s}-E_3-E_4)\delta^{(3)}(\mathbf{p}_3+\mathbf{p}_4)\,.\end{aligned} \tag{6.1.28}$$

We can then use the δ-function over the 3-momenta to integrate over $\mathrm{d}^3\mathbf{p}_4$, such that

$$\mathrm{d}\Pi^{(2)} = \frac{\mathrm{d}^3\mathbf{p}_3}{(2\pi)^2 4E_3E_4}\delta(\sqrt{s}-E_3-E_4)\,, \tag{6.1.29}$$

where now $E_3 = \sqrt{|\mathbf{p}_3|^2+m_3^2}$ and $E_4 = \sqrt{|\mathbf{p}_3|^2+m_4^2}$. Let us now transform to a spherical polar coordinate system, such that $\mathrm{d}^3\mathbf{p}_3 = |\mathbf{p}_3|^2\mathrm{d}|\mathbf{p}_3|\mathrm{d}\Omega_{\mathrm{CM}}$, in which $\mathrm{d}\Omega_{\mathrm{CM}} = \sin(\theta)\mathrm{d}\theta\mathrm{d}\phi$ is the differential solid angle (in the CM frame), ϕ is the azimuthal angle (about the collision axis), and θ is the angle between $\mathbf{p}_1$ and $\mathbf{p}_3$ in the CM frame. Given this, we can then integrate over $\mathrm{d}|\mathbf{p}_3|$ by recalling the following result

$$\delta(f(x)) = \sum_i \frac{1}{|f'(x_i)|}\delta(x-x_i)\,, \tag{6.1.30}$$

where x_i are the roots of $f(x)$ and the $f' := \frac{\mathrm{d}f}{\mathrm{d}x}$. In this particular case, the δ-function vanishes at only one value of $|\mathbf{p}_3|$. Note further, that the derivative of its argument with respect to $|\mathbf{p}_3|$ is

$$\frac{\partial}{\partial|\mathbf{p}_3|}\left(\sqrt{s}-E_3-E_4\right) = -\frac{|\mathbf{p}_3|}{E_3}-\frac{|\mathbf{p}_3|}{E_4} = -\frac{|\mathbf{p}_3|\sqrt{s}}{E_3E_4}\,. \tag{6.1.31}$$

Thus,

$$\delta(\sqrt{s}-E_3-E_4) = \frac{E_3E_4}{|\mathbf{p}_3^*|\sqrt{s}}\,\delta(|\mathbf{p}_3|-|\mathbf{p}_3^*|)\,, \tag{6.1.32}$$

where $|\mathbf{p}_3^*|$ is the (postive) root of the function $f(|\mathbf{p}_3|) = \sqrt{s}-E_3-E_4$, corresponding to its CM expression [eq. (6.1.35)]. Applying this result to eq. (6.1.29), and integrating over $\mathrm{d}|\mathbf{p}_3|$, we arrive at the final expression

$$\mathrm{d}\Pi^{(2)} = \frac{|\mathbf{p}_3^*|}{16\pi^2\sqrt{s}}\,\mathrm{d}\Omega_{\mathrm{CM}}\,. \tag{6.1.33}$$

Finally, upon noting that $\mathbf{v}_i = \frac{1}{E_i}\mathbf{p}_i$, such that $|\mathbf{v}_1-\mathbf{v}_2| = \frac{1}{E_1E_2}|\mathbf{p}_1^*|\sqrt{s}$ (where $\mathbf{p}_1^*$, $\mathbf{p}_2^* = -\mathbf{p}_1^*$ are the 3-momenta of the incoming particles in the CM frame), one finds that eq. (6.1.25) takes the form

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega_{\mathrm{CM}}}\right|_{2\to 2} = \frac{|\mathcal{M}|^2}{64\pi^2 s}\frac{|\mathbf{p}_3^*|}{|\mathbf{p}_1^*|}\,, \tag{6.1.34}$$

where $|\mathbf{p}_1^*| = |\mathbf{p}_2^*|$ and $|\mathbf{p}_3^*| = |\mathbf{p}_4^*|$. Their values in the CM frame are given by

$$|\mathbf{p}_1^*|^2 \;=\; \frac{1}{4s}\lambda(s, m_1, m_2)\,, \tag{6.1.35a}$$

$$|\mathbf{p}_3^*|^2 \;=\; \frac{1}{4s}\lambda(s, m_3, m_4)\,, \tag{6.1.35b}$$

in which

$$\lambda(s, m_i, m_j) \;=\; s^2 - 2(m_i^2 + m_j^2)s + (m_i^2 - m_j^2)^2\,, \tag{6.1.36}$$

with $i = m_1, m_3$ and $j = m_2, m_4$, in which $\lambda(s, m_i, m_j)$ is the Källén function, defined as

$$\lambda(a, b, c) \;=\; a^2 + b^2 + c^2 - 2ab - 2ac - 2bc \;=\; a^2 - 2a(b + c) + (b - c)^2\,. \tag{6.1.37}$$

Note, in particular, that for elastic scattering (where $m_1 = m_3$ and $m_2 = m_4$) in the CM frame, it follows that $|\mathbf{p}_1^*| = |\mathbf{p}_2^*| = |\mathbf{p}_3^*| = |\mathbf{p}_4^*|$, and $E_1 = E_3$, $E_2 = E_4$. In the special case where the masses of the incoming and outgoing particles are all equal, this further reduces to $E_1 = E_2 = E_3 = E_4 = \frac{1}{2}\sqrt{s}$.

For spinless particles, the differential cross-section $\mathrm{d}\sigma_{2\to 2}$ for two-to-two scattering is rotationally symmetric about the collision axis (the azimuthal angle ϕ) and so depends on only two independent variables: the CM energy $\sqrt{s}$ and the scattering angle θ. Moreover, the $t-$channel Mandelstam variable is related to θ as

$$t \;=\; m_1^2 + m_3^2 - 2E_1E_3 + 2|\mathbf{p}_1||\mathbf{p}_3|\cos(\theta) \tag{6.1.38}$$

and therefore, $\mathrm{d}\sigma_{2\to 2}$ is equivalently a function of $\sqrt{s}$ and t. For fixed s, we have that $\mathrm{d}t = 2|\mathbf{p}_1||\mathbf{p}_3|\mathrm{d}\cos(\theta) = 2|\mathbf{p}_1||\mathbf{p}_3|\frac{\mathrm{d}\Omega_{\mathrm{CM}}}{2\pi}$ (where we have used that $\mathrm{d}\Omega_{\mathrm{CM}} = 2\pi\mathrm{d}\cos(\theta)$). Given this, one can straightforwardly express eq. (6.1.34) in a frame-independent manner,

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}t}\right|_{2\to 2} \;=\; \frac{\mathrm{d}\Omega_{\mathrm{CM}}}{\mathrm{d}t}\left.\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega_{\mathrm{CM}}}\right|_{2\to 2} \;=\; \frac{|\mathcal{M}|^2}{64\pi s|\mathbf{p}_1^*|^2}\,, \tag{6.1.39}$$

where $|\mathbf{p}_1^*|$ is a function of s given by eq. (6.1.35a). To obtain the total cross-section, one must integrate (6.1.25) over all outgoing momenta for fixed CM energy $\sqrt{s}$, and divide by the appropriate symmetry factor S: if there are n_j identical outgoing particles of type j, then $S = \prod_j n_j!$, and so

$$\sigma = \frac{1}{S} \int \mathrm{d}\sigma \, . \tag{6.1.40}$$

In the case of two-to-two scattering we can use eq. (6.1.39), and integate with respect to t,

$$\sigma_{2\to 2} = \frac{1}{S} \int_{t_{\min}}^{t_{\max}} \left. \frac{\mathrm{d}\sigma}{\mathrm{d}t} \right|_{2\to 2} \mathrm{d}t \, , \tag{6.1.41}$$

where $S = 1$ ($S = 2$) if the outgoing particles are (in)distinguishable. The limits $t_{\min}$ and $t_{\max}$ can be determined from eq. (6.1.38) evaluated in the CM frame with $\cos(\theta) = -1$ and $\cos(\theta) = +1$, respectively. In the case where the incoming particles are identical (i.e. $m_1 = m_2$), and outgoing particles are also identical (i.e. $m_3 = m_4$), it follows from eq. (6.1.35) that $|\mathbf{p}_1^*| = \frac{1}{2}(s - 4m_1^2)^{1/2}$, and $|\mathbf{p}_3^*| = \frac{1}{2}(s - 4m_1^2)^{1/2}$. Moreover, from eq. (6.1.38), we have that $t_{\min} = m_1^2 + m_1^2 - 2E_1E_2 - 2|\mathbf{p}_1^*||\mathbf{p}_3^*|$ and $t_{\max} = m_1^2 + m_1^2 - 2E_1E_2 + 2|\mathbf{p}_1^*||\mathbf{p}_3^*|$. Given this, eq. (6.1.41) can be expressed as

$$\sigma_{2\to 2} = \frac{1}{2} \int_{t_{\min}}^{t_{\max}} \frac{|\mathcal{M}|^2}{64\pi^2 s |\mathbf{p}_1^*|^2} \mathrm{d}t = \int_{t_{\min}}^{t_{\max}} \frac{|\mathcal{M}|^2}{32\pi s (s - 4m_1^2)} \mathrm{d}t \, . \tag{6.1.42}$$

If $|\mathcal{M}|^2$ is independent of t, then one can readily integrate eq. (6.1.42) to give

$$\sigma_{2\to 2} = \frac{|\mathcal{M}|^2}{32\pi s} \sqrt{\frac{s - 4m_3^2}{s - 4m_1^2}} \, . \tag{6.1.43}$$

Note further, that in the special case where the incoming and outgoing particles are all identical (i.e. $m_1 = m_2 = m_3 = m_4 := m$), then $|\mathbf{p}_1^*| = |\mathbf{p}_3^*| = \frac{1}{2}(s - 4m^2)^{1/2}$, and $t_{\min} = -(s - 4m^2)$ and $t_{\max} = 0$, such that eq. (6.1.42) becomes

$$\sigma_{2\to 2} = \int_{-(s-4m^2)}^{0} \frac{|\mathcal{M}|^2}{32\pi s (s - 4m^2)} \mathrm{d}t \, . \tag{6.1.44}$$

We are now in a position to determine the scattering cross-sections for the quartic and ϕ-mediated interactions arising from eq. (6.1.22), and $\mathscr{L}_{\text{int}}^{\text{eff}} \supset -\frac{g}{2}\varphi(t)\phi\chi^2$, respectively. Let us first consider the quartic interaction. Since the scattering processes are perturbative, the dominant contribution to the cross-section arises from the tree-level diagram (cf. fig. 6.1.1), which corresponds to the S-matrix element $\mathcal{M} = -i\lambda$. Accordingly, one can use eq. (6.1.43), with $m_1 = m_3$, to get

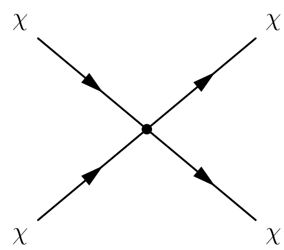


Figure 6.1.1: Tree-level Feynman diagram for two-to-two scattering of χ particles arising from the quartic self-interaction $\frac{\lambda}{4!}\,\hat{\chi}^4(x)$.

$$\sigma_\lambda \;=\; \frac{\lambda^2}{32\pi s}\;. \tag{6.1.45}$$

Turning our attention to the ϕ-mediated interaction, the situation is a little more complicated. In this case, there are three contributing processes at tree-level, arising from the $s-$, $t-$ and $u-$channels (cf. fig. 6.1.2). The corresponding S-matrix element

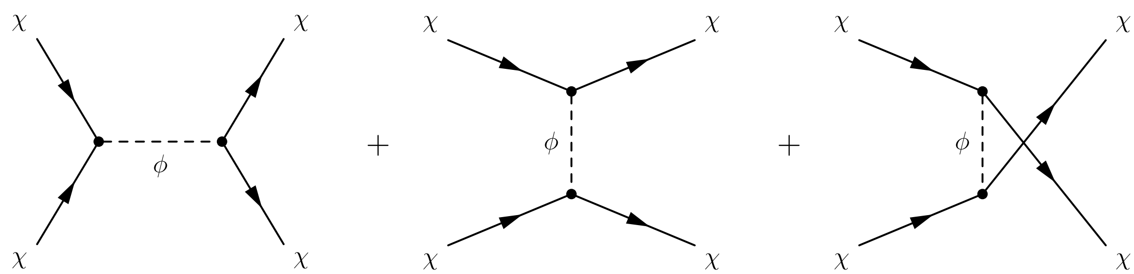


Figure 6.1.2: Tree-level $s-$, $t-$ and $u-$channel Feynman diagrams contributing to the scattering matrix element for ϕ-mediated two-to-two scatterings of χ particles, arising from the effective interaction Lagrangian $\mathscr{L}_{\rm int}^{\rm eff} \supset -\frac{g}{2}\varphi(t)\phi\chi^2$.

in this case is $\mathcal{M} = -ig^2\varphi^2\Big[\frac{1}{m_\phi^2-s} + \frac{1}{m_\phi^2-t} + \frac{1}{m_\phi^2-u}\Big]$. Upon noting that $u = 4m_\chi^2 - s - t$, eq. (6.1.44) can be integrated, yielding the result

$$\begin{aligned}\sigma_g \;=\; \frac{g^4\varphi^4}{32\pi s(s-4m_{\rm eff}^2)}\Bigg[&\frac{s-4m_{\rm eff}^2}{(s-m_\phi^2)^2} + \frac{2}{m_\phi^2} - \frac{2}{s+m_\phi^2-4m_{\rm eff}^2}\\ &- \frac{4(3m_\phi^2-4m_{\rm eff}^2)}{(s-m_\phi^2)(s+2m_\phi^2-4m_{\rm eff}^2)}\ln\!\left(\frac{s+m_\phi^2-4m_\chi^2}{m_\phi^2}\right)\Bigg]\,,\end{aligned} \tag{6.1.46}$$

where $m_{\rm eff}$ is the effective mass of the χ particles, given by eq. (6.1.6). Given Eqs. (6.1.45) and (6.1.46), we are now in a position to determine which of the two interactions dominates the scattering rate for $\chi\chi \to \chi\chi$. To do so, it suffices to consider the ratio of their corresponding cross-sections. With this in mind, let us consider the two regimes of each oscillation interval of φ, namely the intervals of adiabaticity and non-adiabaticity, with the coupling constants taken to be $g \sim 10^{-7}$ and $\lambda \sim 10^{-1}$.[12] Recall from §5.1.2, that the momenta of the produced particles will typically be $|\mathbf{k}| \lesssim \left(\sqrt{g}m_{\phi}\varphi_0\right)^{1/2}$, and in the adiabatic regime $m_{\text{eff}} \sim \sqrt{\frac{g}{2}}\varphi_0$. Therefore, the COM energy will be dominated by contributions from the effective mass of the χ field ($\sqrt{\frac{g}{2}}\varphi_0 \gg m_{\phi},\, m_{\chi}$), thus scaling as $\sqrt{s} \sim \sqrt{2g}\varphi_0$. In this case, we find that the t- and u-channels dominate the ϕ-mediated processes, however their combined cross-section is still suppressed by $\mathcal{O}(10^{-3})$ relative to that of the self-interaction. We now consider the non-adiabtatic regime. In this case, the time-dependent coupling $g\varphi(t) \to 0$, and as such the total cross-section vanishes (even for the smallest centre-of-mass energies $\sqrt{s} \to 2m_{\chi}$). We see, therefore, that the results of this detailed analysis are in agreement with what we expected from our order-of-magnitude estimate of the values that λ can assume in order for the interactions arising from eq. (6.1.22) to dominate the scattering rate. One is thus safe to proceed under the assumption that the dominant collisional processes are those stemming from eq. (6.1.22).

With these initial considerations in mind, we now continue towards deriving a set of equations describing the evolution of the χ particle number density during preheating. For this, we first define the χ particle number density operator $\hat{N}_{\mathbf{k}}(t)$ (per momentum mode) as a dimensionless quantity,

$$\hat{N}_{\mathbf{k}}(t) \;\coloneqq\; \hat{N}(t, \mathbf{k}) \;=\; \frac{\hat{a}^{\dagger}_{\mathbf{k}}(t)\hat{a}_{\mathbf{k}}(t)}{\text{Vol}}\,, \tag{6.1.47}$$

where $\text{Vol} = (2\pi)^3\delta^{(3)}(0)$ is the corresponding spatial volume (note that $\delta^{(3)}(0)$ is the δ-function in momentum-space[13]). To find the total number density operator $\hat{N}(t)$ we must then integrate over the momentum space volume (taking into account the density-of-states per unit volume of momentum space):

[12] This order or magnitude for g is typical in studies or preheating (cf. [97, 98, 224, 237]), and is within the required bounds as discussed in §6.1.1.

[13] To see this, note that

$$(2\pi)^3\delta^{(3)}(0) \;=\; \int_{-\infty}^{+\infty} \mathrm{d}^3\mathbf{x}\; e^{i(\mathbf{p}-\mathbf{k})\cdot\mathbf{x}}\Big|_{\mathbf{p}\,=\,\mathbf{k}} \;=\; \int_{-\infty}^{+\infty} \mathrm{d}^3\mathbf{x} \;=\; \text{Vol}$$

$$\hat{N}(t) \;=\; \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\, \hat{N}_{\mathbf{k}}(t)\,. \tag{6.1.48}$$

Furthermore, as a result of the non-adiabatic evolution of the system, the pair correlations $\langle \hat{a}_{-\mathbf{k}}(t)\hat{a}_{\mathbf{k}}(t)\rangle_t$ and $\langle \hat{a}^\dagger_{\mathbf{k}}(t)\hat{a}^\dagger_{-\mathbf{k}}(t)\rangle_t$, which can be interpreted as quantifying correlations between particle and antiparticle states, will be non-trivial. This is due to the coupling between $\hat{\chi}$ and the time-dependent background field φ generating intervals of non-adiabacity in the evolution of the system. The result is that the evolution of the creation and annihilation operators of $\hat{\chi}$ are coupled during the preheating phase, thus causing them to become mixed as they evolve in time[14]. Once the inflaton has completely decayed, such that $\hat{\chi}$ has essentially decoupled from it, the system evolves adiabatically, i.e. $\dot{\omega}_{\mathbf{k}}(t) \to 0$. The creation and annihilation operators then evolve independently of one another, and as such, there are no correlations between particle and antiparticle states during purely adiabatic evolution.

Accordingly, we shall introduce corresponding *pair* operators $\hat{M}_{\mathbf{k}}(t)$ and $\hat{M}^\dagger_{\mathbf{k}}(t)$, defined respectively as

$$\hat{M}_{\mathbf{k}}(t) \;:=\; \frac{\hat{a}_{-\mathbf{k}}(t)\hat{a}_{\mathbf{k}}(t)}{\mathrm{Vol}}\,, \tag{6.1.49a}$$

$$\hat{M}^\dagger_{\mathbf{k}}(t) \;:=\; \frac{\hat{a}^\dagger_{\mathbf{k}}(t)\hat{a}^\dagger_{-\mathbf{k}}(t)}{\mathrm{Vol}}\,. \tag{6.1.49b}$$

As we are working in the (modified) interaction picture, we note that all operators evolve with respect to the effective Hamiltonian of the collisionless theory, $\hat{H}_0(t)$ according to eq. (6.1.15), with the exception of the the density operator $\hat{\rho}(t)$ which (since it describes the state of the system) evolves with respect to the interaction Hamiltonian $\hat{H}_{\mathrm{int}}(t)$ [eq. (6.1.23)]. With this in mind, we take the total time derivative of $\hat{N}_{\mathbf{k}}(t)$, noting that $\big[\hat{H}_0(t), \hat{N}_{\mathbf{k}}(t)\big] = 0$, such that,

$$\dot{\hat{N}}_{\mathbf{k}}(t) \;=\; \partial_t \hat{N}_{\mathbf{k}}(t) \;=\; \frac{\dot{\omega}_{\mathbf{k}}(t)}{\omega_{\mathbf{k}}(t)} \mathrm{Re}\, \hat{M}_{\mathbf{k}}(t)\,. \tag{6.1.50}$$

Moreover, the pair correlations (Eqs. (6.1.49a) and (6.1.49b)) evolve according to

$$\dot{\hat{M}}_{\mathbf{k}}(t) \;=\; -2i\omega_{\mathbf{k}}(t)\hat{M}_{\mathbf{k}}(t) + \frac{\dot{\omega}_{\mathbf{k}}(t)}{2\omega_{\mathbf{k}}(t)}\big(2\hat{N}_{\mathbf{k}}(t) + 1\big)\,. \tag{6.1.51}$$

[14] cf. eq. (6.1.21) which clearly shows the coupled evolution of $\hat{a}_{\mathbf{k}}(t)$ and $\hat{a}^\dagger_{-\mathbf{k}}(t)$, thus generating time-dependent correlations between them.

where $\frac{\mathrm{d}}{\mathrm{d}t}\hat{M}^\dagger_{\mathbf{k}}$ can be obtained from the complex conjugate of eq. (6.1.51). In particular, note from eq. (6.1.50), that the evolution of the number density operator, $\hat{N}_{\mathbf{k}}(t)$, is solely dependent on the pair operators, $\hat{M}_{\mathbf{k}}(t)$ and $\hat{M}^\dagger_{\mathbf{k}}(t)$. If there were no non-adiabaticity present in the evolution of the system, then $\hat{M}_{\mathbf{k}}(t)$ and $\hat{M}^\dagger_{\mathbf{k}}(t)$ would vanish, and $\hat{N}_{\mathbf{k}}(t)$ would be an (approximate) adiabatic invariant. This observation will have important implications on the production of particles due to parametric resonance.

We have determined how the operators $\hat{N}_{\mathbf{k}}(t)$, $\hat{M}_{\mathbf{k}}(t)$ and $\hat{M}^\dagger_{\mathbf{k}}(t)$ evolve, however, our aim is to study the evolution of the corresponding observable, i.e. the expecation value of the number density operator $N(t) = \big\langle \hat{N}(t) \big\rangle_t$ (where $\hat{N}(t)$ is given by eq. (6.1.48)). Within the framework of the density matrix formalism, the expectation value of a given operator $\hat{\mathcal{O}}(t)$ is given by its ensemble average

$$\big\langle \hat{\mathcal{O}}(t) \big\rangle_t \;\coloneqq\; \frac{\mathrm{Tr}\big[\hat{\rho}(t)\hat{\mathcal{O}}(t)\big]}{\mathrm{Tr}\big[\hat{\rho}(t)\big]} \;=\; \mathrm{Tr}\big[\hat{\rho}(t)\hat{\mathcal{O}}(t)\big]\,, \tag{6.1.52}$$

where Tr is the trace operator[15]. The subscript in the expecation value $\big\langle \hat{\mathcal{O}}(t) \big\rangle_t$ denotes that we are taking the expectation value of an operator $\hat{\mathcal{O}}(t)$ with respect to the density matrix $\hat{\rho}$, evaluated at time t. Thus, in order to establish how the number density $N(t)$ evolves during preheating, we need to determine the evolution of the expectation values of the relevant operators during preheating. This further requires one to specify how the density operator $\hat{\rho}(t)$ evolves. To proceed, we note that the evolution of $\hat{\rho}(t)$ is governed by the quantum Liouville equation

$$\dot{\hat{\rho}}(t) \;=\; \partial_t\hat{\rho}(t) - i\big[\hat{H}_{\mathrm{int}}(t), \hat{\rho}(t)\big] \;=\; -i\big[\hat{H}_{\mathrm{int}}(t), \hat{\rho}(t)\big]\,, \tag{6.1.53}$$

where, to reduce complexity, we have assumed that $\hat{\rho}(t)$ has no explicit time dependence. Integrating eq. (6.1.53) we obtain the formal solution

$$\hat{\rho}(t) \;=\; \hat{\rho}(t_0) - i\int_{t_0}^{t}\mathrm{d}t'\,\big[\hat{H}_{\mathrm{int}}(t'), \hat{\rho}(t')\big]\,, \tag{6.1.54}$$

where $t_0 \leq t' \leq t$, in which t_0 is the initial time. Applying the method of successive substitution, we have that

[15]Note here, that we take the density matrix $\hat{\rho}(t)$ to be normalised to unity, i.e. $\mathrm{Tr}\big[\hat{\rho}(t)\big] \;=\; 1$.

$$\hat{\rho}(t) \;=\; \hat{\rho}(t_0) - i\int_{t_0}^{t} \mathrm{d}t' \left[\hat{H}_{\rm int}(t'), \hat{\rho}(t_0)\right] - \int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t'} \mathrm{d}t'' \left[\hat{H}_{\rm int}(t'), \left[\hat{H}_{\rm int}(t''), \hat{\rho}(t'')\right]\right], \tag{6.1.55}$$

where $t_0 \leq t'' \leq t' \leq t$. Finally, differentiating with respect to t, we can recast the time-evolution of $\hat{\rho}(t)$ in the form

$$\dot{\hat{\rho}}(t) \;=\; -i\big[\hat{H}_{\rm int}(t), \hat{\rho}(t_0)\big] - \int_{t_0}^{t} \mathrm{d}t' \left[\hat{H}_{\rm int}(t), \left[\hat{H}_{\rm int}(t'), \hat{\rho}(t')\right]\right], \tag{6.1.56}$$

where $t_0 \leq t' \leq t$.

Given this result, we then differentiate the number density (per momentum mode) $N_{\mathbf{k}}(t) = \mathrm{Tr}(\hat{\rho}(t)\hat{N}_{\mathbf{k}}(t))$ with respect to t, replacing $\dot{\hat{\rho}}(t)$ with the right-hand side of eq.(6.1.56) wherever it occurs,

$$\begin{aligned}\dot{N}_{\mathbf{k}}(t) \;&= \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Tr}\big[\hat{\rho}(t)\hat{N}_{\mathbf{k}}(t)\big] \;=\; \mathrm{Tr}\big[\hat{\rho}(t)\dot{\hat{N}}_{\mathbf{k}}(t)\big] + \mathrm{Tr}\big[\dot{\hat{\rho}}(t)\hat{N}_{\mathbf{k}}(t)\big] \\ &= \frac{\dot{\omega}_{\mathbf{k}}(t)}{2\omega_{\mathbf{k}}(t)}\Big(\mathrm{Tr}\big[\hat{\rho}(t)\hat{M}_{\mathbf{k}}(t)\big] + \mathrm{Tr}\big[\hat{\rho}(t)\hat{M}^{\dagger}_{\mathbf{k}}(t)\big]\Big) - i\,\mathrm{Tr}\big\{\big[\hat{H}_{\rm int}(t), \hat{\rho}(t_0)\big]\hat{N}_{\mathbf{k}}(t)\big\} \\ &\quad - \int_{t_0}^{t} \mathrm{d}t'\; \mathrm{Tr}\big\{\big[\hat{H}_{\rm int}(t), \big[\hat{H}_{\rm int}(t'), \hat{\rho}(t')\big]\big]\hat{N}_{\mathbf{k}}(t)\big\}\,.\end{aligned} \tag{6.1.57}$$

Using the cyclicity property of the trace and eq. (6.1.52) we can then recast this as,

$$\dot{N}_{\mathbf{k}}(t) \;=\; \frac{\dot{\omega}_{\mathbf{k}}(t)}{\omega_{\mathbf{k}}(t)}\mathrm{Re}\,M_{\mathbf{k}}(t) - i\big\langle\big[\hat{N}_{\mathbf{k}}(t), \hat{H}_{\rm int}(t)\big]\big\rangle_{t_0} - \int_{t_0}^{t} \mathrm{d}t'\;\big\langle\big[\big[\hat{N}_{\mathbf{k}}(t), \hat{H}_{\rm int}(t)\big], \hat{H}_{\rm int}(t')\big]\big\rangle_{t'}\,. \tag{6.1.58}$$

Proceeding similarly for $M_{\mathbf{k}}(t) \coloneqq \mathrm{Tr}\big[\hat{\rho}(t)\hat{M}_{\mathbf{k}}(t)\big]$ and $M^{*}_{\mathbf{k}}(t) \coloneqq \mathrm{Tr}\big[\hat{\rho}(t)\hat{M}^{\dagger}_{\mathbf{k}}(t)\big]$, we obtain the following evolution equation for the pair correlation $M_{\mathbf{k}}(t)$

$$\begin{aligned}\dot{M}_{\mathbf{k}}(t) \;=&\; -2i\omega_{\mathbf{k}}(t)M_{\mathbf{k}}(t) + \frac{\dot{\omega}_{\mathbf{k}}(t)}{2\omega_{\mathbf{k}}(t)}\big(2N_{\mathbf{k}}(t)+1\big) - i\big\langle\big[\hat{M}_{\mathbf{k}}(t), \hat{H}_{\rm int}(t)\big]\big\rangle_{t_0} \\ &- \int_{t_0}^{t} \mathrm{d}t'\;\big\langle\big[\big[\hat{M}_{\mathbf{k}}(t), \hat{H}_{\rm int}(t)\big], \hat{H}_{\rm int}(t')\big]\big\rangle_{t'}\,.\end{aligned} \tag{6.1.59}$$

The corresponding evolution equation for $M^{*}_{\mathbf{k}}(t)$ can be obtained from the complex conjugate of eq. (6.1.59).

Equations (6.1.58) and (6.1.59) compose a coupled set of self-consistent Boltzmann equations that describe the complete evolution of the system, including non-Markovian memory effects. The latter make the solution of this system technically challenging. However, under the assumption that the non-Markovian effects are subdominant, we can make the system tractable by means of a Wigner-Weisskopf (or Markovian) approximation [266]. This relies on two assumptions:

1. molecular chaos: momentum correlations are lost between collisions;

2. a separation of time-scales: the evolution of the system is slow compared with the microscopic QFT processes.

The first assumption introduces a notion of time irreversibility into the evolution equations. That is, prior to a given collision, assuming previous correlations between the momenta of the particles involved have been lost (on time-scales relevant to the statistical evolution of the system), results in an asymmetry in trajectory of the system. Given a set of initial conditions, it is possible to solve the evolution equations and determine its trajectory in the positive time direction. However, it is not possible to evolve the system back to its initial state[16]). This is what enables the system to evolve irreversibly towards a state of thermal equilibrium, i.e. a state of maximal entropy. A consequence of this is that we can express the correlation functions (within the collision integrals) in terms of (products of) single particle distribution functions, viz. the number density and pair correlations.

In order to make use of the second assumption, and closely following ref. [256], we consider the integral

$$\begin{aligned} \mathcal{I} \;&=\; \int_{t_0}^{t} \mathrm{d}t'\, \mathrm{Tr}\big\{ \big[\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t) \big], \hat{H}_{\mathrm{int}}(t') \big] \hat{\rho}(t) \big\} \\ &=\; \int_{t_0}^{t} \mathrm{d}t'\, \mathrm{Tr}\big\{ \big[\hat{F}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t') \big] \hat{\rho}(t') \big\}\,, \end{aligned} \tag{6.1.60}$$

where $\hat{\mathcal{O}}_{\mathbf{k}}(t)$ generically denote the operators of interest, and we have defined the following time-dependent operator

$$\hat{F}_{\mathbf{k}}(t) \;\coloneqq\; \big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t) \big]\,. \tag{6.1.61}$$

[16] One can of course find a solution to the time-reversed evolution equations, but since they are no longer symmetric under $t \to -t$, this will not correspond to a trajectory that leads back to the same state as the initial state of the system, since information about the past dynamics of the system has been lost.

Inserting the Fourier transforms of both $\hat{F}_{\mathbf{k}}(t)$ and $\hat{H}_{\text{int}}(t')$, we have,

$$\mathcal{I} \;=\; \int_{t_0}^{t} \mathrm{d}t' \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega}{2\pi} \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega'}{2\pi}\, e^{i\omega t} e^{i\omega' t'}\, \mathrm{Tr}\big\{\big[\hat{\tilde{F}}_{\mathbf{k}}(\omega), \hat{\tilde{H}}_{\text{int}}(\omega')\big]\hat{\rho}(t')\big\}\,, \qquad (6.1.62)$$

and after making a change of variables $\omega \to \omega - \omega'$, we can re-express this as

$$\mathcal{I} \;=\; \int_{t_0}^{t} \mathrm{d}t' \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega}{2\pi} \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega'}{2\pi}\, e^{i\omega t}\, e^{i\omega'(t'-t)}\, \mathrm{Tr}\left\{\big[\hat{\tilde{F}}_{\mathbf{k}}(\omega-\omega'), \hat{\tilde{H}}_{\text{int}}(\omega')\big]\hat{\rho}(t')\right\}\,. \qquad (6.1.63)$$

Since we are ultimately interested in the evolution of the system over macroscopic time-scales, we shall now assume that there exists a separation of scales between the QFT processes and the coarse-grained statistical evolution of the system, i.e. that the statistical averages of (products of) operators evolve slowly compared to the evolution of the operators $\hat{\mathcal{O}}_{\mathbf{k}}$.[17] Given this, we observe from eq. (6.1.62), that the exponential $e^{i\omega'(t'-t)}$ rapidly oscillates at times far away from $t \sim t'$. Therefore, as long as the inverse Fourier transform of $\hat{\tilde{F}}(\omega-\omega')$ remains dynamical, i.e. it is not sharply peaked around $\omega \sim \omega'$, the integral over ω' is dominated by contributions at $t \sim t'$. We can therefore replace $\hat{\rho}(t')$ by $\hat{\rho}(t)$ (or, equivalently $\langle\cdots\rangle_{t'} \to \langle\cdots\rangle_t$) in the integrand of eq. (6.1.62),

$$\mathcal{I} \simeq \int_{t_0-t}^{0} \mathrm{d}t' \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega}{2\pi} \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega'}{2\pi}\, e^{i\omega t} e^{i\omega' t'}\, \mathrm{Tr}\big\{\big[\hat{\tilde{F}}(\omega-\omega'), \hat{\tilde{H}}_{\text{int}}(\omega')\big]\hat{\rho}(t)\big\}\,, \qquad (6.1.64)$$

where we have made a judicious change of variables $t' \to t' - t$. Moreover, the separation of time-scales implies that we can extend the lower limit of the integral to the infinite past, i.e. $t_0 \to -\infty$, since its tails contribute negligibly (due the incoherent nature of neighbouring values of the integrand). Note that we do not actually change the boundary time t_0; this remains fixed. Rather, we can extend the lower bound of the integral without altering its value significantly. As such, we can write

$$\mathcal{I} \simeq \int_{-\infty}^{0} \mathrm{d}t' \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega}{2\pi} \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega'}{2\pi}\, e^{i\omega t} e^{i\omega' t'}\, \mathrm{Tr}\big\{\big[\hat{\tilde{F}}(\omega-\omega'), \hat{\tilde{H}}_{\text{int}}(\omega')\big]\hat{\rho}(t)\big\}\,. \qquad (6.1.65)$$

[17] One expects this to be true for perturbatively small collisional processes, since these occur at a slow rate (i.e. the time between each collision is macroscopically large), resulting in a slow macroscopic evolution of the system.

Now, let us compute the integral over t' via an analytic continuation of ω' to $\omega' - i\varepsilon$ (with $\varepsilon > 0$):

$$\int_{-\infty}^{0} \mathrm{d}t'\, e^{i\omega' t'} \;=\; \lim_{\varepsilon\to 0}\int_{-\infty}^{0} \mathrm{d}\tilde{t}\, e^{i(\omega'-i\varepsilon)\tilde{t}} \;=\; \lim_{\varepsilon\to 0}\frac{1}{i\omega'+\varepsilon}e^{i\omega'\tilde{t}+\varepsilon\tilde{t}}\bigg|_{-\infty}^{0}. \tag{6.1.66}$$

Note that $\left|e^{i(\omega'-i\varepsilon)t'}\right| = e^{\varepsilon t'} \to 0$ as $t' \to -\infty$, and therefore we arrive at the following formal solution,

$$\int_{-\infty}^{0} \mathrm{d}t'\, e^{i\omega' t'} \;=\; \lim_{\varepsilon\to 0}\left[\frac{\varepsilon}{\omega'^2+\varepsilon^2} - i\frac{\omega'}{\omega'^2+\varepsilon^2}\right] \;=\; \pi\delta(\omega') - i\mathcal{P}\frac{1}{\omega'}\,, \tag{6.1.67}$$

where we have used that the δ-function can be represented as $\delta(\omega') = \lim_{\varepsilon\to 0}\frac{1}{\pi}\frac{\varepsilon}{\omega'^2+\varepsilon^2}$, and that $\mathcal{P}\frac{1}{\omega'} = \lim_{\varepsilon\to 0}\frac{\omega'}{\omega'^2+\varepsilon^2}$, where $\mathcal{P}$ denotes the Cauchy principal value. Since $\delta(\omega')$ can also be represented via a Fourier integral $\delta(\omega') = \int_{-\infty}^{+\infty}\frac{\mathrm{d}t'}{2\pi}\,e^{i\omega' t'}$, we can recast the right-hand side of eq. (6.1.67) as

$$\int_{-\infty}^{0} \mathrm{d}t'\, e^{i\omega' t'} \;=\; \frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{d}t'\, e^{i\omega' t'} - i\mathcal{P}\frac{1}{\omega'}\,. \tag{6.1.68}$$

Given this result, we find that, in the Wigner-Weisskopf approximation, eq. (6.1.60) can be expressed as,

$$\begin{aligned}\mathcal{I} \simeq &\,\frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{d}t'\,\mathrm{Tr}\big\{\big[\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big], \hat{H}_{\mathrm{int}}(t')\big]\hat{\rho}(t)\big\} \\ &- i\mathcal{P}\int_{-\infty}^{+\infty}\frac{d\omega'}{2\pi}\frac{e^{i\omega' t'}}{\omega'}\,\mathrm{Tr}\big\{\big[\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big], \hat{\tilde{H}}_{\mathrm{int}}(\omega')\big]\hat{\rho}(t)\big\}\,. \end{aligned}\tag{6.1.69}$$

The relevant collision contributions are subsumed in the first term in eq. (6.1.69). The second term gives rise to dispersive self-energy corrections[18] [256], which we shall hereafter neglect. As such, eq. (6.1.69) becomes

$$\begin{aligned}\mathcal{I} &\simeq \frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{d}t'\;\mathrm{Tr}\big\{\big[\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big], \hat{H}_{\mathrm{int}}(t')\big]\hat{\rho}(t)\big\} \\ &= \frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{d}t'\;\big\langle\big[\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big], \hat{H}_{\mathrm{int}}(\tilde{t})\big]\big\rangle_t\,, \end{aligned}\tag{6.1.70}$$

[18] These correspond (upon renormalising) to shifts in energy levels, essentially through the dressing of the mass of the χ field.

where we have made use of eq. (6.1.52).

Implementing the above approximations to Eqs. (6.1.58) and (6.1.59), we arrive at the following coupled set of Markovian master equations for the number density $N_{\mathbf{k}}(t)$, and the pair correlations $M_{\mathbf{k}}(t)$ and $M^*_{\mathbf{k}}(t)$, valid at order λ^2:

$$\begin{aligned}\dot{N}_{\mathbf{k}}(t) \;\simeq\; & \frac{\dot{\omega}_{\mathbf{k}}(t)}{\omega_{\mathbf{k}}(t)}\mathrm{Re}\, M_{\mathbf{k}}(t) - i\big\langle\big[\hat{N}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big]\big\rangle_{t_0} \\ & - \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{d}t' \,\big\langle\big[\big[\hat{N}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big], \hat{H}_{\mathrm{int}}(t')\big]\big\rangle_t\,, \end{aligned} \tag{6.1.71a}$$

$$\begin{aligned}\dot{M}_{\mathbf{k}}(t) \;\simeq\; & -2i\omega_{\mathbf{k}}(t)M_{\mathbf{k}}(t) + \frac{\dot{\omega}_{\mathbf{k}}(t)}{2\omega_{\mathbf{k}}(t)}\big(2N_{\mathbf{k}}(t)+1\big) - i\big\langle\big[\hat{M}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big]\big\rangle_{t_0} \\ & - \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{d}t' \,\big\langle\big[\big[\hat{M}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big], \hat{H}_{\mathrm{int}}(t')\big]\big\rangle_t\,, \end{aligned} \tag{6.1.71b}$$

$$\begin{aligned}\dot{M}^*_{\mathbf{k}}(t) \;\simeq\; & 2i\omega_{\mathbf{k}}(t)M^*_{\mathbf{k}}(t) + \frac{\dot{\omega}_{\mathbf{k}}(t)}{2\omega_{\mathbf{k}}(t)}\big(2N_{\mathbf{k}}(t)+1\big) - i\big\langle\big[\hat{M}^\dagger_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big]\big\rangle_{t_0} \\ & - \frac{1}{2}\int_{-\infty}^{\infty} \mathrm{d}t' \,\big\langle\big[\big[\hat{M}^\dagger_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big], \hat{H}_{\mathrm{int}}(t')\big]\big\rangle_t\,, \end{aligned} \tag{6.1.71c}$$

Note that, while we are applying these Markovian master equations in the context of preheating, they can be applied in a more general setting to describe the evolution of any interacting scalar field with a time-dependent mass term. In particular, we observe from eq. (6.1.71a), that the particle production terms for the number density are proportional to $M_{\mathbf{k}}(t)$ and $M^*_{\mathbf{k}}(t)$. We see, therefore, that non-adiabatic particle production depends crucially on the existence and non-vanishing values of pair correlations, such that any processes that destroy these coherences will suppress (and eventually shut off) the resonant production (cf. ref. [262]). Furthermore, since $M_{\mathbf{k}}(t)$ and $M^*_{\mathbf{k}}(t)$ source $N_{\mathbf{k}}(t)$, one should expect their values to be of the same order of magnitude, so long as $|\dot{\omega}_{\mathbf{k}}/\omega^2_{\mathbf{k}}| \sim \mathcal{O}(1)$, which is true during intervals of non-adiabaticity.

In a more realistic scenario, the inflaton condensate will decay. Consequently, the particle production terms (the first term in the right-hand side of each of the master equations) will decrease as time progresses. Once the condensate has completely diminished, the production terms in the master equations will vanish and the system will continue to thermalise. At this point, the evolution of the system will become adiabatic, resulting in the decoherence of the pair correlations, $M_{\mathbf{k}}(t)$ and $M^*_{\mathbf{k}}(t)$.

Once the remaining collision terms have died off, the system will have equilibriated, and one expects that $N_{\mathbf{k}}(t)$ will have the form of a Bose-Einstein distribution.

6.2 A self-consistent set of Boltzmann transport equations

6.2.1 Derivation of the approximate mode functions

In the previous section, we derived a set of Boltzmann transport equations describing the evolutions of $N_{\mathbf{k}}(t)$, $M_{\mathbf{k}}(t)$ and $M^*_{\mathbf{k}}(t)$, to leading-order in the Markovian approximation. We now wish to express the master equations entirely in terms of $N_{\mathbf{k}}(t)$, $M_{\mathbf{k}}(t)$ and $M^*_{\mathbf{k}}(t)$, so as to obtain a self-consistent set of evolution equations. In order to do so, we first need to solve the Heisenberg equations for $\hat{a}_{\mathbf{k}}(t)$ and $\hat{a}^\dagger_{\mathbf{k}}(t)$ (see eq. (6.1.21)), so that we can evolve all of the operators appearing in the collision terms to equal times, namely the time t of the density operator. To this end, we assume a Bogoliubov ansatz [267], relating the creation and operators at some time t to their counterparts at some other time $\tilde{t}$

$$\hat{a}_{\mathbf{p}}(t) \;=\; \alpha_{\mathbf{p}}(t,\tilde{t})\hat{a}_{\mathbf{p}}(\tilde{t}) + \beta^*_{\mathbf{p}}(t,\tilde{t})\hat{a}^\dagger_{-\mathbf{p}}(\tilde{t})\,, \tag{6.2.1a}$$

$$\hat{a}^\dagger_{\mathbf{p}}(t) \;=\; \alpha^*_{\mathbf{p}}(t,\tilde{t})\hat{a}^\dagger_{\mathbf{p}}(\tilde{t}) + \beta_{\mathbf{p}}(t,\tilde{t})\hat{a}_{-\mathbf{p}}(\tilde{t})\,. \tag{6.2.1b}$$

Such transformations preserve the canonical algebra (cf. eq. (6.1.13)), which implies that the Bogoliubov coefficients $\alpha_{\mathbf{k}}(t,\tilde{t})$ and $\beta_{\mathbf{k}}(t,\tilde{t})$ satisfy the constraint

$$\left|\alpha_{\mathbf{k}}(t,\tilde{t})\right|^2 - \left|\beta_{\mathbf{k}}(t,\tilde{t})\right|^2 \;=\; 1\,. \tag{6.2.2}$$

Furthermore, by setting $t = \tilde{t}$, we see that $\alpha_{\mathbf{k}}(t,\tilde{t})$ and $\beta_{\mathbf{k}}(t,\tilde{t})$ satisfy the following boundary conditions:

$$\alpha^{(*)}_{\mathbf{k}}(\tilde{t},\tilde{t}) \;=\; 1\,, \quad \beta^{(*)}_{\mathbf{k}}(\tilde{t},\tilde{t}) \;=\; 0\,. \tag{6.2.3}$$

Additionally, their isotropy is guaranteed since $\hat{\chi}$ is a real scalar field, i.e.,

$$\alpha_{\mathbf{k}}(t,\tilde{t}) \;=\; \alpha_{-\mathbf{k}}(t,\tilde{t})\,, \quad \beta_{\mathbf{k}}(t,\tilde{t}) \;=\; \beta_{-\mathbf{k}}(t,\tilde{t})\,. \tag{6.2.4}$$

The Bogoliubov coefficients $\alpha_{\mathbf{k}}(t,\tilde{t})$ and $\beta_{\mathbf{k}}(t,\tilde{t})$ satisfy the differential equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\alpha_{\mathbf{k}}(t,\tilde{t}) \;=\; -\,i\,\omega_{\mathbf{k}}(t)\alpha_{\mathbf{k}}(t,\tilde{t}) \;+\; \frac{\dot{\omega}_{\mathbf{k}}(t)}{2\,\omega_{\mathbf{k}}(t)}\beta_{\mathbf{k}}(t,\tilde{t})\;, \tag{6.2.5a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\beta_{\mathbf{k}}(t,\tilde{t}) \;=\; +\,i\,\omega_{\mathbf{k}}(t)\beta_{\mathbf{k}}(t,\tilde{t}) \;+\; \frac{\dot{\omega}_{\mathbf{k}}(t)}{2\,\omega_{\mathbf{k}}(t)}\alpha_{\mathbf{k}}(t,\tilde{t})\;, \tag{6.2.5b}$$

which follow directly from Eqs. (6.1.21). Now, observe that eq. (6.2.1) can be neatly recast into matrix form

$$\begin{pmatrix} \hat{a}_{\mathbf{p}}(t) \\ \hat{a}^{\dagger}_{-\mathbf{p}}(t) \end{pmatrix} = \begin{pmatrix} \alpha_{\mathbf{p}}(t,\tilde{t}) & \beta^{*}_{\mathbf{p}}(t,\tilde{t}) \\ \beta_{\mathbf{p}}(t,\tilde{t}) & \alpha^{*}_{\mathbf{p}}(t,\tilde{t}) \end{pmatrix} \begin{pmatrix} \hat{a}_{\mathbf{p}}(\tilde{t}) \\ \hat{a}^{\dagger}_{-\mathbf{p}}(\tilde{t}) \end{pmatrix}\,. \tag{6.2.6}$$

Expressing the transformation in this form has the particular advantage that one can readily obtain its inverse

$$\begin{aligned} \begin{pmatrix} \hat{a}_{\mathbf{p}}(\tilde{t}) \\ \hat{a}^{\dagger}_{-\mathbf{p}}(\tilde{t}) \end{pmatrix} &= \begin{pmatrix} \alpha_{\mathbf{p}}(t,\tilde{t}) & \beta^{*}_{\mathbf{p}}(t,\tilde{t}) \\ \beta_{\mathbf{p}}(t,\tilde{t}) & \alpha^{*}_{\mathbf{p}}(t,\tilde{t}) \end{pmatrix}^{-1} \begin{pmatrix} \hat{a}_{\mathbf{p}}(t) \\ \hat{a}^{\dagger}_{-\mathbf{p}}(t) \end{pmatrix} = \begin{pmatrix} \alpha^{*}_{\mathbf{p}}(t,\tilde{t}) & -\beta^{*}_{\mathbf{p}}(t,\tilde{t}) \\ -\beta_{\mathbf{p}}(t,\tilde{t}) & \alpha_{\mathbf{p}}(t,\tilde{t}) \end{pmatrix} \begin{pmatrix} \hat{a}_{\mathbf{p}}(t) \\ \hat{a}^{\dagger}_{-\mathbf{p}}(t) \end{pmatrix} \\ &\coloneqq \begin{pmatrix} \alpha_{\mathbf{p}}(\tilde{t},t) & \beta^{*}_{\mathbf{p}}(\tilde{t},t) \\ \beta_{\mathbf{p}}(\tilde{t},t) & \alpha^{*}_{\mathbf{p}}(\tilde{t},t) \end{pmatrix} \begin{pmatrix} \hat{a}_{\mathbf{p}}(t) \\ \hat{a}^{\dagger}_{-\mathbf{p}}(t) \end{pmatrix}\,. \end{aligned} \tag{6.2.7}$$

It is clear then, that $\alpha_{\mathbf{k}}(\tilde{t},t)$ and $\beta_{\mathbf{k}}(\tilde{t},t)$ are related to $\alpha_{\mathbf{k}}(t,\tilde{t})$ and $\beta_{\mathbf{k}}(t,\tilde{t})$ as follows:

$$\alpha_{\mathbf{k}}(\tilde{t},t) \;= \alpha^{*}_{\mathbf{k}}(t,\tilde{t})\,, \tag{6.2.8a}$$

$$\beta_{\mathbf{k}}(\tilde{t},t) \;=\; -\,\beta_{\mathbf{k}}(t,\tilde{t})\,. \tag{6.2.8b}$$

Moreover, one can iterate eq. (6.2.6) to derive the following composition properties:

$$\alpha_{\mathbf{k}}(t,\tilde{t}') \;=\; \alpha_{\mathbf{k}}(t,\tilde{t})\alpha_{\mathbf{k}}(\tilde{t},\tilde{t}') + \beta^{*}_{\mathbf{k}}(t,\tilde{t})\beta_{\mathbf{k}}(\tilde{t},\tilde{t}')\,, \tag{6.2.9a}$$

$$\beta_{\mathbf{k}}(t,\tilde{t}') \;=\; \beta_{\mathbf{k}}(t,\tilde{t})\alpha_{\mathbf{k}}(\tilde{t},\tilde{t}') + \alpha^{*}_{\mathbf{k}}(t,\tilde{t})\beta_{\mathbf{k}}(\tilde{t},\tilde{t}')\,. \tag{6.2.9b}$$

This then provides us with a good consistency check; by setting $t = \tilde{t}'$, we see that eq. (6.2.9) implies eq. (6.2.8). Equipped with the Bogoliubov transformation [eq. (6.2.1)] and its inverse [eq. (6.2.7)], we can now express the creation and annihilation operators at time $t' \leq t$ in terms of their later time forms $\hat{a}_{\mathbf{k}}(t)$ and $\hat{a}^{\dagger}_{\mathbf{k}}(t)$ and vice-verse. Indeed, upon applying these results to eq. (6.1.10), the field at time

$t' \leq t$ is given by

$$\hat{\chi}(t', \mathbf{x}) \;=\; \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \left[\tilde{\chi}_{\mathbf{k}}(t', t)\hat{a}_{\mathbf{k}}(t)e^{+i\mathbf{k}\cdot\mathbf{x}} \;+\; \tilde{\chi}^*_{\mathbf{k}}(t', t)\hat{a}^\dagger_{\mathbf{k}}(t)e^{-i\mathbf{k}\cdot\mathbf{x}}\right], \tag{6.2.10}$$

where the mode function $\tilde{\chi}_{\mathbf{k}}(t', t)$ is given by

$$\tilde{\chi}_{\mathbf{k}}(t', t) \;\coloneqq\; \frac{1}{\sqrt{2\omega_{\mathbf{k}}(t')}} \left[\alpha_{\mathbf{k}}(t', t) \;+\; \beta_{\mathbf{k}}(t', t)\right]. \tag{6.2.11}$$

which satisfies the following EOM,

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\chi_{\mathbf{k}}(t', t) + \omega^2_{\mathbf{k}}(t)\chi_{\mathbf{k}}(t', t) \;=\; 0\,. \tag{6.2.12}$$

This can further be recast in the form of a Mathieu equation, upon a change of variables, $m_\phi t \to z = m_\phi t + \frac{\pi}{2}$,

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}\chi_{\mathbf{k}}(z', z) + \left[A_k - 2q\cos(2z)\right]\chi_{\mathbf{k}}(z', z) \;=\; 0\,, \tag{6.2.13}$$

where $A_k = \frac{\mathbf{k}^2 + m_\chi^2}{m_\phi^2} + 2q$ and $q = \frac{g\varphi_0^2}{8m_\phi^2}$. As discussed in §5.1, q parameterises whether we are in narrow ($q \ll 1$) or broad ($q \gg 1$) resonance. We therefore see that one can recover the mode function analysis within the framework of the density matrix formalism. One would of course expect this to be the case, and therefore it provides a good consistency check.

In their present form, Eqs. (6.2.5a) and (6.2.5b) do not lend themselves to an analytical calculation. However, we intend to make use of the solutions to them only in the collision terms. Moreover, recall the Wigner-Weisskopf approximation of the collision terms, described earlier in subsection 6.1.2, wherein we argued that the collision integral is dominated by times $t' \sim t$. This allowed us to replace $\hat{\rho}(t')$ by $\hat{\rho}(t)$ and extend the upper and lower limits of the integration over t' to positive and negative infinity, respectively. Once all of the creation and annihilation operators have been evolved to the time t, we are therefore interested in the Bogoliubov coefficients only for times t' near t, while treating the integration over this neighbourhood in t' as effectively infinite as far as the dynamics of the fast modes is concerned. Accordingly, for $t \sim t'$, we can set $\beta_{\mathbf{k}}(t, t') \simeq 0$, such that (cf. eq. (6.2.5))

$$\dot{\alpha}_{\mathbf{k}}(t, t') \;\simeq\; -\,i\omega_{\mathbf{k}}(t)\alpha_{\mathbf{k}}(t, t')\,, \tag{6.2.14}$$

effectively imposing adiabatic evolution in the neighbourhood of the time t. The solution to eq. (6.2.14), satisfying the boundary condition in eq. (6.2.3), is

$$\alpha_{\mathbf{k}}(t,t') \;\simeq\; \exp\Big[-i\int_{t'}^{t} \mathrm{d}\tilde{t}\;\omega_{\mathbf{k}}(\tilde{t})\Big]\,, \tag{6.2.15}$$

and the inverse function $\alpha_{\mathbf{k}}(t',t)$ can be obtained from the property $\alpha_{\mathbf{k}}(t',t)=\alpha^{*}_{\mathbf{k}}(t,t')$ (cf. eq. (6.2.8)).

Given eq. (6.2.15), we are then left with the task of evaluating the integral in the exponent. Within the adiabatic approximation, this can be done in closed form; specifically

$$\int_{t'}^{t} \mathrm{d}\tilde{t}\,\omega_{\mathbf{k}}(\tilde{t}) \;=\; \frac{\omega_{\mathbf{k}}|_{\max}}{m_\phi}\Big[\mathrm{E}\left(m_\phi t, z_{\mathbf{k}}\right) \;-\; \mathrm{E}\left(m_\phi t', z_{\mathbf{k}}\right)\Big]\,, \tag{6.2.16}$$

where $\mathrm{E}\left(m_\phi t, z_{\mathbf{k}}\right)$ is the incomplete elliptic integral of the second kind, and the argument $z_{\mathbf{k}}$ is given by $z_{\mathbf{k}}=\frac{g\varphi_0^2}{2(\omega_{\mathbf{k}}|_{\max})^2}$, in which,

$$\omega_{\mathbf{k}}(t)|_{\max} \;:=\; \sqrt{\mathbf{k}^2+m_\chi^2+\frac{g\varphi_0^2}{2}}\,, \tag{6.2.17}$$

is the maximum value of $\omega_{\mathbf{k}}(t)$. While it is not appropriate to expand $\omega_{\mathbf{k}}(t)$ perturbatively in the coupling g, since $g\varphi_0^2 \gg m_\chi^2$, a numerical analysis of eq. (6.2.16) suggests that it is possible to approximate the solution well by making a series expansion of $\mathrm{E}\left(m_\phi t, z_{\mathbf{k}}\right)$ with respect to $z_{\mathbf{k}}$ (for $z_{\mathbf{k}}<1$). To linear order, we obtain

$$\mathrm{E}\left(m_\phi t, z_{\mathbf{k}}\right)-\mathrm{E}\left(m_\phi t', z_{\mathbf{k}}\right) \;\simeq\; \left(1-\frac{z_{\mathbf{k}}}{4}\right)m_\phi(t-t')+\frac{z_{\mathbf{k}}}{4}m_\phi\big[t\,\mathrm{sinc}(2m_\phi t)-t'\,\mathrm{sinc}(2m_\phi t')\big]\,. \tag{6.2.18}$$

For $z_{\mathbf{k}}\ll 1$, the relative error is $\sim 0.01\%$; for $z_{\mathbf{k}}\sim\mathcal{O}(1)$, it is at most $\sim 15\%$ (see fig. 6.2.1)[19]. Away from $t,t'=0$, the first term dominates, and we can neglect the terms involving sinc functions, whose relative contributions decrease linearly with time for $t>\frac{1}{m_\phi}$. Doing so, we are left with the result

$$\int_{t'}^{t} \mathrm{d}\tilde{t}\;\omega_{\mathbf{k}}(\tilde{t}) \;\simeq\; \bar{\omega}_{\mathbf{k}}(t-t')\,. \tag{6.2.19}$$

[19]We can tolerate an error of $\sim 15\%$ as it is a global error in the overall contribution from the collision integrals, and not a relative error between its constituent terms.

We see therefore, that eq. (6.2.16) can be interpreted as the approximate time-average of $\omega_{\mathbf{k}}(t)$ over an interval $\Delta t = t - t'$, i.e.

$$\bar{\omega}_{\mathbf{k}} \;\simeq\; \frac{1}{\Delta t}\int_{t'}^{t} \mathrm{d}\tilde{t}\; \omega_{\mathbf{k}}(\tilde{t}) \;. \tag{6.2.20}$$

To be consistent with this approximation, we should make the replacement $\omega_{\mathbf{k}}(t) \to \bar{\omega}_{\mathbf{k}}$ in the mode functions (cf. eq. (6.2.11)) and whenever it appears in the collision integrals. In this way, eq. (6.2.11) for the mode function reduces to

$$\tilde{\chi}_{\mathbf{k}}(t', t) \;\simeq\; \frac{1}{\sqrt{2\bar{\omega}_{\mathbf{k}}}}\, e^{+i\bar{\omega}_{\mathbf{k}}(t-t')} \;. \tag{6.2.21}$$

Notice that, in the limit $\varphi_0 \to 0$, corresponding to the decay of the amplitude of the inflaton oscillations, $\bar{\omega}_{\mathbf{k}} \to \sqrt{\mathbf{k}^2 + m_\chi^2}$, and we correctly recover the usual evolution of the free field.

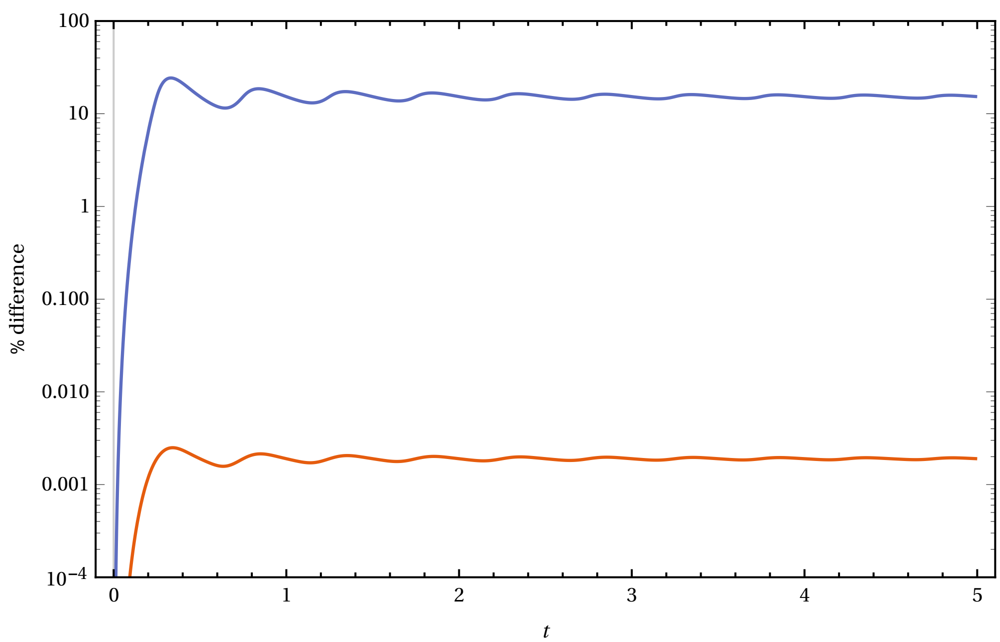


Figure 6.2.1: Plot of the percentage difference between $\int_{t_0}^{t} \mathrm{d}t'\, \omega_{\mathbf{k}}(t')$ and its Taylor approximation (cf. eq. (6.2.18)) as a function of t (in units of $m_\phi/2\pi$) over several periods of oscillation of the inflaton condensate for $z_{\mathbf{k}} \sim \mathcal{O}(1)$ (blue line) and $z_{\mathbf{k}} \ll 1$ (orange).

Observe that the approximation we have discussed above imposes adiabatic evolution in the collision integral between the times t and t'. Since the periods of non-adiabatic particle production are much shorter than the intervening periods of adiabatic evolution, this is expected to be a valid procedure, so long as the time-scale for the collisions is much larger than that of each burst of particle production. We now wish to show that such a separation of scales can be present during the preheating phase. To do so, we first recall from §5.1.2, that during the broad resonance phase, one can estimate the particle production rate to be $\Gamma_* := \frac{1}{\Delta t_*} \sim \left(\sqrt{\frac{g}{2}} m_\phi \varphi_0\right)^{1/2}$, and the produced particles typically have momenta lying in the interval $0 \leq |\mathbf{k}| \lesssim m_\phi (\frac{q}{4})^{1/4} \approx 4 m_\phi < \sqrt{\frac{g}{2}}\varphi_0$ for $q = \mathcal{O}(10^3)$ [98]. During the very early stages of preheating, we can therefore estimate the total number density of χ particles to be $N \propto m_\phi^3$. Now, turning our attention to the collision processes, we note that the collision rate is given by $\Gamma_{\chi\chi\to\chi\chi} = |\mathbf{v}_{\rm rel}| N \sigma_{\chi\chi\to\chi\chi}$, where $|\mathbf{v}_{\rm rel}|$ is the relative velocity between the two colliding particles, and $\sigma_{\chi\chi\to\chi\chi}$ is the cross-section for the two-to-two scattering processes. To remain conservative with our estimate for the collision rate, herein we shall assume that the relative velocity, which scales as $|\mathbf{v}_{\rm rel}| \sim 2\frac{|\mathbf{k}|}{m_{\rm eff}(t)}$, is of order one.

Let us now consider $\sigma_{\chi\chi\to\chi\chi}$ in vacuo (cf. eq. (6.1.45))[20], which in this case is given by eq. (6.1.45). Accordingly, when $m_{\rm eff}^2 \simeq \frac{g\varphi_0^2}{2}$ is maximal, $|\mathbf{k}| \lesssim m_{\rm eff}$, and the two-to-two scattering cross-section scales as $\sigma_{\chi\chi\to\chi\chi} \sim \frac{\lambda^2}{16\pi\, g\varphi_0^2}$. When $m_{\rm eff}^2 \simeq \frac{g\varphi_0^2}{2}$ is maximal, $|\mathbf{k}| \lesssim m_{\rm eff}$, and the two-to-two scattering cross-section scales as $\sigma_{\chi\chi\to\chi\chi} \sim \frac{\lambda^2}{64\pi\, g\varphi_0^2}$. On the other hand, when φ nears the turning point of an oscillation and for modes with $|\mathbf{k}| \lesssim m_\chi$, the cross-section scales as $\sigma_{\chi\chi\to\chi\chi} \sim \frac{\lambda^2}{128\pi m_\chi^2}$. For modes with $|\mathbf{k}| \lesssim m_\phi$, the cross-section scales as $\sigma_{\chi\chi\to\chi\chi} \sim \frac{\lambda^2}{128\pi m_\phi^2}$. Hence, in order to achieve the separation of scales required above, i.e. $\Gamma_{\chi\chi\to\chi\chi} \ll \Gamma_*$, we need $\frac{\lambda^2 m_\phi^{5/2}}{128\pi}\left(\sqrt{\frac{g}{2}}\varphi_0\right)^{-1/2}\left\{\frac{1}{g\varphi_0^2/2}, \frac{1}{m_\chi^2}, \frac{1}{m_\phi^2}\right\} \ll 1$, and all three cases can be satisfied for $\lambda = \mathcal{O}(0.1)$.

At this point, we note that the particle number can become extremely large during preheating, in which case the system can become effectively strongly coupled [233], such that many-to-many processes dominate over the two-to-two scatterings that we consider here. However, in this analysis we are interested only in the first few oscillations of the inflaton, wherein the particle number remains relatively small and the collision integral amounts to perturbatively small corrections to the Boltzmann equations. Nevertheless, one would anticipate that such strong coupling would only increase the impact of the collisions on the dynamics of the resonance.

[20] This should provide us with a leading order estimate for how it scales at finite temperature.

6.2.2 Collision terms in the Boltzmann equations

Having determined approximate forms for the mode functions $\tilde{\chi}_{\mathbf{k}}(t', t)$, we can now proceed to evaluate the remaining expectation values in the Markovian master equations [eq. (6.1.71)]. In order to do so, we make two additional assumptions about the system: (i) that the state is (approximately) Gaussian so that all higher-order correlations can be expressed in terms of one- and two-point functions by Wick's theorem and (ii) that the system is spatially homogeneous, such that momentum-space two-point correlation functions can be written in the form

$$\big\langle \hat{\mathcal{O}}_{\mathbf{p}}(t)\hat{\mathcal{O}}_{\mathbf{k}}(t)\big\rangle = \frac{(2\pi)^3}{\mathrm{Vol}}\delta^{(3)}(\mathbf{p}+\mathbf{k})\big\langle \hat{\mathcal{O}}_{-\mathbf{k}}(t)\hat{\mathcal{O}}_{\mathbf{k}}(t)\big\rangle\,, \tag{6.2.22a}$$

$$\big\langle \hat{\mathcal{O}}^{\dagger}_{\mathbf{p}}(t)\hat{\mathcal{O}}_{\mathbf{k}}(t)\big\rangle = \frac{(2\pi)^3}{\mathrm{Vol}}\delta^{(3)}(\mathbf{p}-\mathbf{k})\big\langle \hat{\mathcal{O}}_{\mathbf{k}}(t)\hat{\mathcal{O}}_{\mathbf{k}}(t)\big\rangle\,. \tag{6.2.22b}$$

We note from eq. (6.1.71), that the various correlation functions (at $\mathcal{O}(\lambda)$ and $\mathcal{O}(\lambda^2)$) can be expressed generically as

$$\big\langle\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big]\big\rangle_{t_0} = \frac{\lambda}{4!}\int_{\mathbf{x}}\big\langle\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{\chi}^4(t,\mathbf{x})\big]\big\rangle_{t_0} \tag{6.2.23}$$

and

$$\big\langle\big[\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big], \hat{H}_{\mathrm{int}}(t')\big]\big\rangle_{t} = \left(\frac{\lambda}{4!}\right)^2\int_{\mathbf{x},\mathbf{y}}\big\langle\big[\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{\chi}^4(x)\big], \hat{\chi}^4(y)\big]\big\rangle_{t}\,, \tag{6.2.24}$$

where $\hat{\mathcal{O}}_{\mathbf{k}} \in \{\hat{N}_{\mathbf{k}}, \hat{M}_{\mathbf{k}}\}$, and $x^\mu = (t, \mathbf{x})$ and $y^\mu = (t', \mathbf{y})$. With a view to deriving explicit expressions for Eqs. (6.2.23) and (6.2.24), let us expand their corresponding commutators:

$$\begin{aligned}
\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big] &= \frac{\lambda}{4!}\int_{\mathbf{x}}\big[\hat{\mathcal{O}}_{\mathbf{k}}(t),\, \hat{\chi}^4(t,\mathbf{x})\big] \\
&= \frac{\lambda}{4!}\int_{\mathbf{x}}\Big\{\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{\chi}(t,\mathbf{x})\big]\hat{\chi}^3(t,\mathbf{x}) + \hat{\chi}(t,\mathbf{x})\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{\chi}(t,\mathbf{x})\big]\hat{\chi}^2(t,\mathbf{x}) \\
&\qquad + \hat{\chi}^2(t,\mathbf{x})\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{\chi}(t,\mathbf{x})\big]\hat{\chi}(t,\mathbf{x}) + \hat{\chi}^3(t,\mathbf{x})\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{\chi}(t,\mathbf{x})\big]\Big\} \\
&= \frac{\lambda}{4!}\int_{\mathbf{x}}\Big\{4\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{\chi}(t,\mathbf{x})\big]\hat{\chi}^3(t,\mathbf{x}) - 6\Delta_{\hat{\mathcal{O}}_{\mathbf{k}}}\hat{\chi}^2(t,\mathbf{x})\Big\}\,,
\end{aligned} \tag{6.2.25}$$

and,

$$\begin{aligned}
\big[\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{H}_{\rm int}(t)\big], \hat{H}_{\rm int}(t')\big] \;&=\; \left(\frac{\lambda}{4!}\right)^2 \int_{\mathbf{x},\,\mathbf{y}} \big[\big[\hat{\mathcal{O}}_{\mathbf{k}}(t),\, \hat{\chi}^4(t,\mathbf{x})\big],\, \hat{\chi}^4(t',\mathbf{y})\big] \\
= \left(\frac{\lambda}{4!}\right)^2 &\int_{\mathbf{x},\,\mathbf{y}} \Big\{4\big[\hat{\mathcal{O}}_{\mathbf{k}},\, \hat{\chi}\big]\big[\hat{\chi}^3,\, \hat{\chi}'^4\big] + 4\big[\big[\hat{\mathcal{O}}_{\mathbf{k}},\, \hat{\chi}\big],\, \hat{\chi}'^4\big]\hat{\chi}^3 - 6\Delta_{\hat{\mathcal{O}}_{\mathbf{k}}}\big[\hat{\chi}^2,\, \hat{\chi}'^4\big]\Big\} \\
= \left(\frac{\lambda}{4!}\right)^2 &\int_{\mathbf{x},\,\mathbf{y}} \Big\{16\Delta'_{\hat{\mathcal{O}}_{\mathbf{k}}}\big[\hat{\chi}^3\hat{\chi}'^3 - 9\Delta\hat{\chi}^2\hat{\chi}'^2 + 18\Delta^2\hat{\chi}\hat{\chi}' - 6\Delta^3\big] + 96\Delta^3\big[\hat{\mathcal{O}}_{\mathbf{k}},\, \hat{\chi}\big]\hat{\chi}' \\
&\quad - 144\Delta^2\big[\hat{\mathcal{O}}_{\mathbf{k}},\, \hat{\chi}\big]\hat{\chi}\hat{\chi}'^2 + 48\Delta\big[\hat{\mathcal{O}}_{\mathbf{k}},\, \hat{\chi}\big]\hat{\chi}^2\hat{\chi}'^3 - 24\Delta_{\hat{\mathcal{O}}_{\mathbf{k}}}\Delta\big[2\hat{\chi}\hat{\chi}'^3 - 3\Delta\hat{\chi}'^2\big]\Big\}\,,
\end{aligned} \tag{6.2.26}$$

where all operators with the latest time $t \geq t'$ have been commuted to the left (to simplify their Wick contraction). Herein, we have suppressed the spacetime arguments of the various operators and used the shorthand notations $\hat{\chi} \equiv \hat{\chi}(t,\mathbf{x})$ and $\hat{\chi}' \equiv \hat{\chi}(t',\mathbf{y})$. In addition, we have defined the various Pauli-Jordan-like functions

$$\Delta \;\equiv\; \Delta(x,y) \;\coloneqq\; \big[\hat{\chi}(x),\, \hat{\chi}(y)\big] \;=\; \int_{\mathbf{p}} \Big(\tilde{\chi}^*_{\mathbf{p}}(t',t)e^{+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} - \tilde{\chi}_{\mathbf{p}}(t',t)e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}\Big)\,, \tag{6.2.27a}$$

$$\Delta'_{\hat{\mathcal{O}}_{\mathbf{k}}} \;\equiv\; \Delta_{\hat{\mathcal{O}}_{\mathbf{k}}}(x,y) \;\coloneqq\; \big[\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{\chi}(t,\mathbf{x})\big], \hat{\chi}(t',\mathbf{y})\big]\,, \tag{6.2.27b}$$

$$\Delta_{\hat{\mathcal{O}}_{\mathbf{k}}} \;\equiv\; \Delta_{\hat{\mathcal{O}}_{\mathbf{k}}}(x,x)\,. \tag{6.2.27c}$$

We now proceed to take the trace of Eqs. (6.2.25) and (6.2.26) with density operator, and Wick contract the resulting expressions. In the case of eq. (6.2.25), we have

$$\big\langle\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{H}_{\rm int}(t)\big]\big\rangle_{t_0} \;=\frac{\lambda}{4}\int_{\mathbf{x}} \Big\{\big\langle\big[\hat{\mathcal{O}}_{\mathbf{k}}, \hat{\chi}\big]\hat{\chi}\big\rangle_{t_0} - \big\langle\hat{\chi}\big[\hat{\mathcal{O}}_{\mathbf{k}}, \hat{\chi}\big]\big\rangle_{t_0}\Big\}\big\langle\hat{\chi}^2\big\rangle_{t_0}\,. \tag{6.2.28}$$

In fact, the contributions to eq. (6.1.71) from eq. (6.2.28) correspond to radiative corrections to the mass of the χ field. Since we are interested in the thermalisation process, which is precipitated by collisional processes, we shall henceforth neglect the contributions from eq. (6.2.28).

Let us now Wick contract eq. (6.2.26):

$$
\begin{aligned}
&\big\langle\big[\big[\hat{\mathcal{O}}_{\mathbf{k}}, \hat{H}_{\rm int}\big], \hat{H}'_{\rm int}\big]\big\rangle_t \\
&= -\frac{1}{2}\int_{\mathbf{x},\mathbf{y}} \Bigg\{ \Big[\Pi^{<}(x,y)G^{>}_{\mathcal{O}_{\mathbf{k}}}(y,x) \;-\; \Pi^{>}(x,y)G^{<}_{\mathcal{O}_{\mathbf{k}}}(y,x)\Big] \\
&\qquad - \frac{\lambda^4}{2}\Big[G^{<}_{\mathcal{O}_{\mathbf{k}}}(y,x)G^{>}(x,y) \;-\; G^{>}_{\mathcal{O}_{\mathbf{k}}}(y,x)G^{<}(x,y)\Big]\big\langle\hat{\chi}^2(x)\big\rangle_t\big\langle\hat{\chi}^2(y)\big\rangle_t \\
&\qquad - \frac{\lambda^2}{8}\Big[G^{<}_{\mathcal{O}_{\mathbf{k}}}(x,x)\big(G^{>}(x,y)\big)^2 \;-\; G^{<}_{\mathcal{O}_{\mathbf{k}}}(x,x)\big(G^{<}(x,y)\big)^2 \\
&\qquad\qquad + G^{>}_{\mathcal{O}_{\mathbf{k}}}(x,x)\big(G^{<}(x,y)\big)^2 \;-\; G^{>}_{\mathcal{O}_{\mathbf{k}}}(x,x)\big(G^{>}(x,y)\big)^2\Big]\big\langle\hat{\chi}^2(y)\big\rangle_t \Bigg\}\,,
\end{aligned}
\tag{6.2.29}
$$

where $G^{\gtrless}(x,y)$ are the positive- ($>$) and negative-frequency ($<$) Wightman propagators, to which the functions $G^{\gtrless}_{\mathcal{O}_{\mathbf{k}}}(x,y)$ are related, and the $\Pi^{\gtrless}(x,y)$ correspond to the two cut self-energies. These are defined respectively as

$$G^{>}(x,y) \;\coloneqq\; \langle\chi(x)\chi(y)\rangle_t \;=\; \big(G^{<}(x,y)\big)^*\,, \tag{6.2.30a}$$

$$\Pi^{>}(x,y) \;\coloneqq\; \frac{\lambda^2}{3!}\big(G^{>}(x,y)\big)^3 \;=\; \big(\Pi^{<}(x,y)\big)^*\,, \tag{6.2.30b}$$

$$G^{>}_{\mathcal{O}_{\mathbf{k}}}(x,y) \;\coloneqq\; \big\langle\hat{\chi}(x)\big[\hat{\mathcal{O}}_{\mathbf{k}}(t),\hat{\chi}(y)\big]\big\rangle_t\,, \tag{6.2.30c}$$

$$G^{<}_{\mathcal{O}_{\mathbf{k}}}(x,y) \;\coloneqq\; \big\langle\big[\hat{\mathcal{O}}_{\mathbf{k}}(t),\hat{\chi}(y)\big]\hat{\chi}(x)\big\rangle_t\,. \tag{6.2.30d}$$

Note from eq. (6.2.29), that the presence of $\Pi^{>}(x,y)$ and $\Pi^{<}(x,y)$ implies that all of the terms in the integrand correspond to on-shell processes, and thus one should view them diagramatically as the cuts of their corresponding bubble diagrams, such that the internal cut lines become external legs. Furthermore, the insertion of the number density in each term, arising from the correlation functions containing the commutator between $\hat{\chi}$ and $\hat{\mathcal{O}}_{\mathbf{k}}$, indicates that one should "open-up" the corresponding bubble diagrams at the point where the number density is inserted (as it carries an external on-shell momentum $\mathbf{k}$). In doing so, one should identify the endpoints of the two corresponding external legs.

Given this, in the Markovian limit, the terms $\Pi^{>}(x,y)G^{<}_{\mathcal{O}_{\mathbf{k}}}(y,x)$ and $\Pi^{<}(x,y)G^{>}_{\mathcal{O}_{\mathbf{k}}}(y,x)$ in eq. (6.2.29) give rise to the relevant two-to-two scattering processes. That they are the correct terms follows from the fact that diagramatically, they correspond to

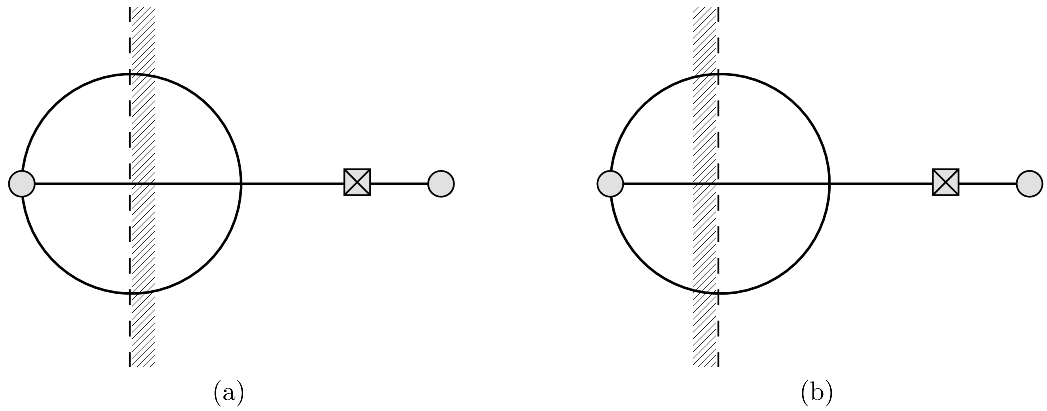


Figure 6.2.2: The forward (a) and backward (b) cuts of the scalar two-loop "sunset" diagram. The small shaded circles indicate coincident points, and the crossed boxes indicate insertions of the operator $\hat{\mathcal{O}}_{\mathbf{k}}$. The net energy flow is from the unshaded to shaded regions.

the absorptive forward and backward cuts (corresponding to the direction of energy flow, respectively) of the two-particle irreducible (2PI) sunset diagram (cf. fig. 6.2.2). Thermalization of the produced particles is precipitated by collision processes. These correspond to two-to-two scatterings, of which the kinematics of such a diagram permits. Thus, we can interpret Figs. 6.2.3a and 6.2.3b as gain and loss terms, in which particles are added and subtracted, respectively, from a given momentum mode. Moreover, the remaining terms in eq. (6.2.29) (2nd, 3rd and 4th lines of eq. (6.2.29)) correspond to $\mathcal{O}(\lambda^2)$ shifts in the mass of the χ field. As was the case for the $\mathcal{O}(\lambda)$ contributions [eq. (6.2.28)], we shall omit these in the present analysis, since we are interested only in the terms that drive thermalisation. Therefore, only collisional contributions remain in eq. (6.2.29), and it thus reduces to the following collision integral

$$\begin{aligned} C_{\mathbf{k}}^{(\mathcal{O})}[N, M; t] \;&=\; -\,\frac{1}{2}\int_{\infty}^{+\infty} \mathrm{d}t' \,\big\langle \big[\big[\hat{\mathcal{O}}_{\mathbf{k}}(t), \hat{H}_{\mathrm{int}}(t)\big], \hat{H}_{\mathrm{int}}(t')\big]\big\rangle_t \\ &\simeq \frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{d}t' \int_{\mathbf{x},\mathbf{y}} \Big[\Pi^{<}(x,y) G^{>}_{\mathcal{O}_{\mathbf{k}}}(y,x) \;-\; \Pi^{>}(x,y) G^{<}_{\mathcal{O}_{\mathbf{k}}}(y,x)\Big]\,. \end{aligned} \tag{6.2.31}$$

While we calculate the collision terms directly from the operator algebra, the fact that we have been able to write the collision contributions in terms of the self-energies and Green functions provides us with a good consistency check. This being that it enables us to make contact with approaches based on non-equilibrium quantum field theory (see, e.g., Refs. [249, 250]). In this case, the collision terms can be associated with the absorptive cuts of the non-equilibrium self-energies, which can be calculated by means of the Kobes-Semenoff [268, 269] cutting rules that generalize those of Cutkosky [270], and 't Hooft and Veltman [271]. The gain and loss terms are then associated with the forward and backward cuts of the two-loop sunset diagram, as depicted in fig. 6.2.2, wherein all cut lines are placed on-shell with the net energy flow proceeding from the unshaded to the shaded regions.

Let us now consider the collision integrals for $N_{\mathbf{k}}(t)$ and $M_{\mathbf{k}}(t)$ (the collision integral for $M^*_{\mathbf{k}}(t)$ can be obtained through complex conjugation of M's). In the case of $N_{\mathbf{k}}(t)$, we note that it is a purely real quantity (as is $\dot{N}_{\mathbf{k}}(t)$), and therefore it must be that $C^{(N)}_{\mathbf{k}}[N,M;t] = \mathrm{Re}\, C^{(N)}_{\mathbf{k}}[N,M;t]$. Indeed, the integrand can be written as

$$
\begin{aligned}
&\Pi^<(x,y)G^>_{N_{\mathbf{k}}}(y,x) - \Pi^>(x,y)G^<_{N_{\mathbf{k}}}(y,x) \\
&= \frac{1}{\mathrm{Vol}\sqrt{2\omega_{\mathbf{k}}(t)}}\Big[\Big(\Pi^<(x,y)\big\langle\hat{\chi}(y)\hat{a}^\dagger_{\mathbf{k}}(t)\big\rangle_t - \Pi^>(x,y)\big\langle\hat{a}^\dagger_{\mathbf{k}}(t)\hat{\chi}(y)\big\rangle_t\Big)e^{-i\mathbf{k}\cdot\mathbf{x}} \\
&\qquad + \Big(\Pi^>(x,y)\big\langle\hat{a}_{\mathbf{k}}(t)\hat{\chi}(y)\big\rangle_t - \Pi^<(x,y)\big\langle\hat{\chi}(y)\hat{a}_{\mathbf{k}}(t)\big\rangle_t\Big)e^{+i\mathbf{k}\cdot\mathbf{x}}\Big] \\
&= 2\,\mathrm{Re}\Big[\frac{1}{\mathrm{Vol}\sqrt{2\omega_{\mathbf{k}}(t)}}\Big(\Pi^<(x,y)\big\langle\hat{\chi}(y)\hat{a}^\dagger_{\mathbf{k}}(t)\big\rangle_t - \Pi^>(x,y)\big\langle\hat{a}^\dagger_{\mathbf{k}}(t)\hat{\chi}(y)\big\rangle_t\Big)e^{-i\mathbf{k}\cdot\mathbf{x}}\Big]\,,
\end{aligned}
\tag{6.2.32}
$$

and thus is a real quantity, such that the integral is itself real. To proceed, we note that $\Pi^{\lessgtr}$, $\langle\hat{a}^\dagger_{\mathbf{k}}\hat{\chi}'\rangle_t$ and $\langle\hat{\chi}'\hat{a}^\dagger_{\mathbf{k}}\rangle_t$ can be expressed in terms of the number density $N_{\mathbf{k}}$ and pair correlations $M_{\mathbf{k}}$ and $M^*_{\mathbf{k}}$ as follows

$$
\Pi^>(x,y) \;=\; \frac{\lambda^2}{3!}\prod_{i=1}^{3}\int_{\boldsymbol{\kappa}_i}\mathrm{Tr}\big[\mathbb{X}_{\boldsymbol{\kappa}_i}(t',t)\mathbb{N}_{\boldsymbol{\kappa}_i}(t)\mathbb{A}_{(1)}\big]e^{+i\boldsymbol{\kappa}_i\cdot(\mathbf{x}-\mathbf{y})}\,, \tag{6.2.33a}
$$

$$
\big\langle\hat{a}^\dagger_{\mathbf{k}}(t)\hat{\chi}(y)\big\rangle_t \;=\; \mathrm{Tr}\big[\mathbb{X}_{\mathbf{k}}(t',t)\mathbb{N}_{\mathbf{k}}(t)\mathbb{A}_{(2)}\big]e^{+i\mathbf{k}\cdot\mathbf{y}} \;=\; \big\langle\hat{\chi}(y)\hat{a}_{\mathbf{k}}(t)\big\rangle^*_t\,, \tag{6.2.33b}
$$

$$
\big\langle\hat{\chi}(y)\hat{a}^\dagger_{\mathbf{k}}(t)\big\rangle_t \;=\; \big[\mathrm{Tr}\big[\mathbb{X}_{\mathbf{k}}(t',t)\mathbb{N}_{\mathbf{k}}(t)\mathbb{A}_{(3)}\big]\big]^*e^{+i\mathbf{k}\cdot\mathbf{y}} \;=\; \big\langle\hat{a}_{\mathbf{k}}(t)\hat{\chi}(y)\big\rangle^*_t\,, \tag{6.2.33c}
$$

where $\boldsymbol{\kappa}_i \in \{\mathbf{p},\, \mathbf{q},\, \mathbf{l}\}$, and we have introduced the compact matrix notation $\mathbb{A}_{(1)}$, $\mathbb{X}_{\boldsymbol{\kappa}}(t',t)$ and $\mathbb{N}_{\boldsymbol{\kappa}}(t)$ such that

$$\mathbb{A}_{(1)} \;\coloneqq\; \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\,, \quad \mathbb{A}_{(2)} \;\coloneqq\; \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\,, \quad \mathbb{A}_{(3)} \;\coloneqq\; \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \tag{6.2.34a}$$

$$\mathbb{X}_{\boldsymbol{\kappa}_i}(t',t) \;\coloneqq\; \begin{pmatrix} \tilde{\chi}_{\boldsymbol{\kappa}_i}(t',t) & 0 \\ 0 & \tilde{\chi}^*_{\boldsymbol{\kappa}_i}(t',t) \end{pmatrix}\,, \tag{6.2.34b}$$

$$\mathbb{N}_{\boldsymbol{\kappa}_i}(t) \;\coloneqq\; \begin{pmatrix} N_{\boldsymbol{\kappa}_i}(t) & M_{\boldsymbol{\kappa}_i}(t) \\ M^*_{\boldsymbol{\kappa}_i}(t) & 1+N_{\boldsymbol{\kappa}_i}(t) \end{pmatrix}\,, \tag{6.2.34c}$$

where $\chi_{\boldsymbol{\kappa}_i}(t',t)$ is the approximate solution for the mode function, as defined in eq. (6.2.11). As such, using the expressions for $\Pi^{\lessgtr}$, $\langle \hat{a}^\dagger_{\mathbf{k}}\chi'\rangle_t$ and $\langle \chi'\hat{a}^\dagger_{\mathbf{k}}\rangle_t$ [eq. (6.2.33)] in eq. (6.2.32), and upon some simplifications[21], the collision integral $C^{(N)}_{\mathbf{k}}[N,M;t]$ can be expressed as

$$C^{(N)}_{\mathbf{k}}[N,M;t] \;=\; \frac{\lambda^2}{3!}\mathrm{Re}\int_{-\infty}^{+\infty}\mathrm{d}t'\int_{\mathbf{p},\mathbf{q}}\frac{1}{\sqrt{2\bar{\omega}_{\mathbf{k}}}\sqrt{2\bar{\omega}_{\mathbf{k}+\mathbf{p}-\mathbf{q}}}}\tilde{f}^{(N)}_{\mathbf{k},\mathbf{p},\mathbf{q}}(t,t') \tag{6.2.35}$$

where the function $\tilde{f}^{(N)}_{\mathbf{k},\mathbf{p},\mathbf{q}}(t,t')$ is defined as,

$$\begin{aligned}\tilde{f}^{(N)}_{\mathbf{k},\mathbf{p},\mathbf{q}}(t,t') \;=&\; \big[\mathrm{Tr}\big[\mathbb{X}_{\mathbf{k}}(t',t)\mathbb{N}_{\mathbf{k}}(t)\mathbb{A}_{(3)}\big]\big]^*\prod_{j=1}^{3}\big[\mathrm{Tr}\big[\mathbb{X}_{\boldsymbol{\kappa}'_j}(t',t)\mathbb{N}_{\boldsymbol{\kappa}'_j}(t)\mathbb{A}_{(1)}\big]\big]^* \\ &- \mathrm{Tr}\big[\mathbb{X}_{\mathbf{k}}(t',t)\mathbb{N}_{\mathbf{k}}(t)\mathbb{A}_{(2)}\big]\prod_{j=1}^{3}\mathrm{Tr}\big[\mathbb{X}_{\boldsymbol{\kappa}'_j}(t',t)\mathbb{N}_{\boldsymbol{\kappa}'_j}(t)\mathbb{A}_{(1)}\big]\,.\end{aligned} \tag{6.2.36}$$

with $\boldsymbol{\kappa}'_i \in \{\mathbf{p},\, \mathbf{q},\, \mathbf{k}+\mathbf{p}-\mathbf{q}\}$. Note that in eq. (6.2.35) we have replaced the time-dependent frequency $\omega_{\boldsymbol{\kappa}}(t)$ with its (approximate) time-averaged form $\bar{\omega}_{\boldsymbol{\kappa}}$ (cf. eq. (6.2.20)) wherever it appears in the integrand. As discussed at the end of §6.2.1, we are required to do this so as to be consistent with the approximations made in determining the (approximate) solutions for the mode functions $\chi_{\boldsymbol{\kappa}}(t',t)$ [eq. (6.2.21)].

[21] To arrive at eq. (6.2.35) we have made use of the fact that we are free to make the change of variables $\boldsymbol{\kappa}_i \to -\boldsymbol{\kappa}_i$ in each of the integrals in the expressions for $\Pi^{\lessgtr}$. Furthermore, we have used that $\int_{\mathbf{x}} e^{-i(\mathbf{k}+\mathbf{p}-\mathbf{q}-\mathbf{l})} = 2\pi\,\delta^{(3)}(\mathbf{k}+\mathbf{p}-\mathbf{q}-\mathbf{l})$.

Following the same procedure for the collision integral for $M_{\mathbf{k}}(t)$, we have from eq. (6.2.31), that

$$\begin{aligned}
C^{(M)}_{\mathbf{k}}[N,M;t] \;&=\; \frac{1}{2}\int_{-\infty}^{+\infty}\mathrm{d}t'\int_{\mathbf{x},\mathbf{y}}\Big[\Pi^{<}(x,y)G^{>}_{M_{\mathbf{k}}}(y,x) \;-\; \Pi^{>}(x,y)G^{<}_{M_{\mathbf{k}}}(y,x)\Big] \\
=\; \frac{1}{2}\int_{-\infty}^{+\infty}\mathrm{d}t'\int_{\mathbf{x},\mathbf{y}}&\frac{1}{\mathrm{Vol}\sqrt{2\bar{\omega}_{\mathbf{k}}}}\Big[\Pi^{<}(x,y)\Big(\big\langle\hat{\chi}(y)\hat{a}_{\mathbf{k}}(t)\big\rangle_t e^{+i\mathbf{k}\cdot\mathbf{x}} - \big\langle\hat{\chi}(y)\hat{a}_{-\mathbf{k}}(t)\big\rangle_t e^{-i\mathbf{k}\cdot\mathbf{x}}\Big) \\
&\qquad - \Pi^{>}(x,y)\Big(\big\langle\hat{a}_{\mathbf{k}}(t)\hat{\chi}(y)\big\rangle_t e^{+i\mathbf{k}\cdot\mathbf{x}} - \big\langle\hat{a}_{-\mathbf{k}}(t)\hat{\chi}(y)\big\rangle_t e^{-i\mathbf{k}\cdot\mathbf{x}}\Big)\Big] \\
=\; \frac{\lambda^2}{3!}\int_{-\infty}^{+\infty}\mathrm{d}t'\int_{\mathbf{p},\mathbf{q}}&\frac{1}{\sqrt{2\bar{\omega}_{\mathbf{k}}}\sqrt{2\bar{\omega}_{\mathbf{k}+\mathbf{p}-\mathbf{q}}}}\tilde{f}^{(M)}_{\mathbf{k},\mathbf{p},\mathbf{q}}(t,t')\;, \qquad (6.2.37)
\end{aligned}$$

where the function $\tilde{f}^{(M)}_{\mathbf{k},\mathbf{p},\mathbf{q}}(t,t')$ is defined as,

$$\begin{aligned}
\tilde{f}^{(M)}_{\mathbf{k},\mathbf{p},\mathbf{q}}(t,t') \;&=\; \big[\mathrm{Tr}\big[\mathbb{X}_{\mathbf{k}}(t',t)\mathbb{N}_{\mathbf{k}}(t)\mathbb{A}_{(2)}\big]\big]^{*}\prod_{j=1}^{3}\big[\mathrm{Tr}\big[\mathbb{X}_{\boldsymbol{\kappa}'_j}(t',t)\mathbb{N}_{\boldsymbol{\kappa}'_j}(t)\mathbb{A}_{(1)}\big]\big]^{*} \\
&\quad - \mathrm{Tr}\big[\mathbb{X}_{\mathbf{k}}(t',t)\mathbb{N}_{\mathbf{k}}(t)\mathbb{A}_{(3)}\big]\prod_{j=1}^{3}\mathrm{Tr}\big[\mathbb{X}_{\boldsymbol{\kappa}'_j}(t',t)\mathbb{N}_{\boldsymbol{\kappa}'_j}(t)\mathbb{A}_{(1)}\big]\,. \qquad (6.2.38)
\end{aligned}$$

The current form of the collision integrals $C^{(N)}_{\mathbf{k}}[N,M;t]$ and $C^{(M)}_{\mathbf{k}}[N,M;t]$ [Eqs. (6.2.35) and (6.2.37)] can be further reduced. Indeed, upon inserting the approximate solution for the mode functions $\chi_{\boldsymbol{\kappa}}(t',t)$ [eq. (6.2.21)] into Eqs. (6.2.36) and (6.2.38), both collision integrals can be expressed in the following general form,

$$\mathcal{C}^{(\mathcal{O})}_{\mathbf{k}}[N,M;t] \;\simeq\; \frac{\lambda^2}{2}\sum_{j\,=\,1}^{4}\int\mathrm{d}\Pi^{(j)}_{\mathbf{p},\mathbf{q},\mathbf{k}}\; f^{(\mathcal{O})}_{(j);\mathbf{p},\mathbf{q},\mathbf{k}}[N,M;t]\;, \qquad (6.2.39)$$

where we have introduced the modified phase-space measure

$$\mathrm{d}\Pi^{(j)}_{\mathbf{p},\mathbf{q},\mathbf{k}} \;=\; \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3}\,2\pi\delta\left(\Delta\omega_j\right)\prod_{\boldsymbol{\kappa}}\frac{1}{2\bar{\omega}_{\boldsymbol{\kappa}}}\;, \qquad (6.2.40)$$

and defined the set of functions $\{\Delta\omega_j | j = 1, 2, 3, 4\}$, with

$$\Delta\omega_1 = \bar{\omega}_{\mathbf{k}} + \bar{\omega}_{\mathbf{p}+\mathbf{q}-\mathbf{k}} - \bar{\omega}_{\mathbf{p}} - \bar{\omega}_{\mathbf{q}} \,, \tag{6.2.41a}$$

$$\Delta\omega_2 = \bar{\omega}_{\mathbf{k}} + \bar{\omega}_{-\mathbf{p}-\mathbf{q}-\mathbf{k}} + \bar{\omega}_{\mathbf{p}} + \bar{\omega}_{\mathbf{p}} \,, \tag{6.2.41b}$$

$$\Delta\omega_3 = \bar{\omega}_{\mathbf{p}} + \bar{\omega}_{\mathbf{q}} + \bar{\omega}_{\mathbf{k}-\mathbf{p}-\mathbf{q}} - \bar{\omega}_{\mathbf{k}} \,, \tag{6.2.41c}$$

$$\Delta\omega_4 = \bar{\omega}_{\mathbf{k}} + \bar{\omega}_{\mathbf{q}} + \bar{\omega}_{\mathbf{p}-\mathbf{k}-\mathbf{q}} - \bar{\omega}_{\mathbf{p}} \,. \tag{6.2.41d}$$

The set of functions $\{f^{(\mathcal{O})}_{(j);\mathbf{p},\mathbf{q},\mathbf{k}}[N, M; t] | j = 1, 2, 3, 4\}$ contain the statistical factors, and their explicit expressions (for $\mathcal{C}^{(N)}_{\mathbf{k}}[N, M; t]$ and $\mathcal{C}^{(M)}_{\mathbf{k}}[N, M; t]$ respectively) are defined in the appendix.

In deriving eq. (6.2.39) we have made use of the fact that, within the Wigner-Weisskopf (or Markovian) approximation, only the mode functions depend on t', and therefore their products (appearing in Eqs. (6.2.36) and (6.2.38)) can be reduced to δ-functions upon integrating over t', for example,

$$\begin{aligned}
\int_{-\infty}^{+\infty} \mathrm{d}t' \, \chi_{\mathbf{k}}\chi_{\mathbf{p}}\chi^*_{\mathbf{q}}\chi^*_{\mathbf{k}+\mathbf{p}-\mathbf{q}} &\simeq \frac{1}{\sqrt{2\bar{\omega}_{\mathbf{k}}}\sqrt{2\bar{\omega}_{\mathbf{p}}}\sqrt{2\bar{\omega}_{\mathbf{q}}}\sqrt{2\bar{\omega}_{\mathbf{k}+\mathbf{p}-\mathbf{q}}}} \int_{-\infty}^{+\infty} \mathrm{d}t'' \, e^{-i\Delta\omega_1 t''} \\
&= \frac{1}{\sqrt{2\bar{\omega}_{\mathbf{k}}}\sqrt{2\bar{\omega}_{\mathbf{p}}}\sqrt{2\bar{\omega}_{\mathbf{q}}}\sqrt{2\bar{\omega}_{\mathbf{k}+\mathbf{p}-\mathbf{q}}}} \, 2\pi\delta\left(\Delta\omega_1\right) \,,
\end{aligned} \tag{6.2.42}$$

where we have made a change of variables $t' \to t'' = t - t'$. We see therefore, that the Wigner-Weisskopf approximation restores quasi-energy conservation[22] at each interaction vertex. Consequently, the only kinematically viable process is the two-to-two scattering, corresponding to the case $j = 1$ in eq. (6.2.39). Accordingly, eq. (6.2.39) reduces to

$$\mathcal{C}^{(\mathcal{O})}_{\mathbf{k}}[N, M, t] \simeq \frac{\lambda^2}{2} \int \mathrm{d}\Pi_{\mathbf{p},\mathbf{q},\mathbf{k}} \, f^{(\mathcal{O})}_{\mathbf{p},\mathbf{q},\mathbf{k}}[N, M; t] \,, \tag{6.2.43}$$

where

$$\mathrm{d}\Pi_{\mathbf{p},\mathbf{q},\mathbf{k}} = \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3} \, 2\pi\delta(\bar{\omega}_{\mathbf{k}} + \bar{\omega}_{\mathbf{p}+\mathbf{q}-\mathbf{k}} - \bar{\omega}_{\mathbf{p}} - \bar{\omega}_{\mathbf{q}}) \prod_{\boldsymbol{\kappa}} \frac{1}{2\bar{\omega}_{\boldsymbol{\kappa}}} \,. \tag{6.2.44}$$

[22] The "quasi-energy" conservation refers to the fact that it is not the instantaneous time-dependent energies that are conserved but rather their approximate time averages.

where $\boldsymbol{\kappa} \in \{\mathbf{k},\, \mathbf{p},\, \mathbf{q},\, (\mathbf{p}+\mathbf{q}-\mathbf{k})\}$. The real and imaginary parts of the statistical weights $f^{(N)}_{\mathbf{p},\mathbf{q},\mathbf{k}}[N,M;t]$ and $f^{(M)}_{\mathbf{p},\mathbf{q},\mathbf{k}}[N,M;t]$ are given by

$$\begin{aligned}
&\mathrm{Re} f^{(N)}_{\mathbf{p},\mathbf{q},\mathbf{k}}[N,M;t] \;=\; -\,\mathrm{Re} f^{(M)}_{\mathbf{p},\mathbf{q},\mathbf{k}}[N,M;t] \\
&=\; \mathrm{Re}\Big[(1+N_{\mathbf{k}}-M_{\mathbf{k}})\left(N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)(1+N_{\mathbf{p}+\mathbf{q}-\mathbf{k}}+M_{\mathbf{p}+\mathbf{q}-\mathbf{k}}) \\
&\quad -\;(N_{\mathbf{k}}-M^*_{\mathbf{k}})(1+N_{\mathbf{p}}+M_{\mathbf{p}})(1+N_{\mathbf{q}}+M_{\mathbf{q}})\left(N_{\mathbf{p}+\mathbf{q}-\mathbf{k}}+M^*_{\mathbf{p}+\mathbf{q}-\mathbf{k}}\right)\Big]\,, \qquad (6.2.45a)
\end{aligned}$$

$$\begin{aligned}
&\mathrm{Im} f^{(M)}_{\mathbf{p},\mathbf{q},\mathbf{k}}[N,M;t] \\
&=\; \mathrm{Im}\Big[(N_{\mathbf{k}}+M^*_{\mathbf{k}})(1+N_{\mathbf{p}}+M_{\mathbf{p}})(1+N_{\mathbf{q}}+M_{\mathbf{q}})\left(N_{\mathbf{p}+\mathbf{q}-\mathbf{k}}+M^*_{\mathbf{p}+\mathbf{q}-\mathbf{k}}\right) \\
&\quad -\;(1+N_{\mathbf{k}}+M_{\mathbf{k}})\left(N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)(1+N_{\mathbf{p}+\mathbf{q}-\mathbf{k}}+M_{\mathbf{p}+\mathbf{q}-\mathbf{k}})\Big]\,. \qquad (6.2.45b)
\end{aligned}$$

These comprise the gain and loss terms that arise from two-to-two scattering processes, which are shown diagrammatically in fig. 6.2.3, where we have treated the contributions from $N_{\mathbf{k}}$ and $M^{(*)}_{\mathbf{k}}$ separately.We draw attention to the relative signs between the N's and M's in the factors corresponding to the external momentum $\mathbf{k}$ in eq. (6.2.45), which do not occur for the internal statistical factors. These have arisen because the gain and loss terms are interchanged in the contributions from the external pair correlation $M_{\mathbf{k}}$ relative to those arising from the number density $N_{\mathbf{k}}$.

Given this extensive analysis, upon inserting the form of the collision integral [eq. (6.2.43)] into the Markovian master equations [eq. (6.1.71)], and neglecting the $\mathcal{O}(\lambda)$ contributions[23], we arrive at a (self-consistent) set of Boltzmann equations describing the evolution of the number density of scalar particles during the post-inflationary preheating phase, in which we capture the leading-order two-to-two collisional processes driving thermalistion:

$$\dot{N}_{\mathbf{k}}(t) \;\simeq\; \frac{\dot{\omega}_{\mathbf{k}}(t)}{\omega_{\mathbf{k}}(t)}\,\mathrm{Re}\, M_{\mathbf{k}}(t) \;+\; \mathrm{Re}\,\mathcal{C}^{(N)}_{\mathbf{k}}[N,M;t]\,, \qquad (6.2.46a)$$

$$\mathrm{Re}\,\dot{M}_{\mathbf{k}}(t) \;\simeq\; +\,2\omega_{\mathbf{k}}(t)\,\mathrm{Im}\, M_{\mathbf{k}}(t) \;+\; \frac{1}{2}\,\frac{\dot{\omega}_{\mathbf{k}}(t)}{\omega_{\mathbf{k}}(t)}\Big(2N_{\mathbf{k}}(t) \;+\; 1\Big) \;+\; \mathrm{Re}\,\mathcal{C}^{(M)}_{\mathbf{k}}[N,M;t]\,, \qquad (6.2.46b)$$

$$\mathrm{Im}\,\dot{M}_{\mathbf{k}}(t) \;\simeq\; -\,2\omega_{\mathbf{k}}(t)\,\mathrm{Re}\, M_{\mathbf{k}}(t) \;+\; \mathrm{Im}\,\mathcal{C}^{(M)}_{\mathbf{k}}[N,M;t]\,, \qquad (6.2.46c)$$

[23] As previously discussed, these correspond to radiative corrections to the χ mass, which are not of relevance here as they do not drive thermalisation.

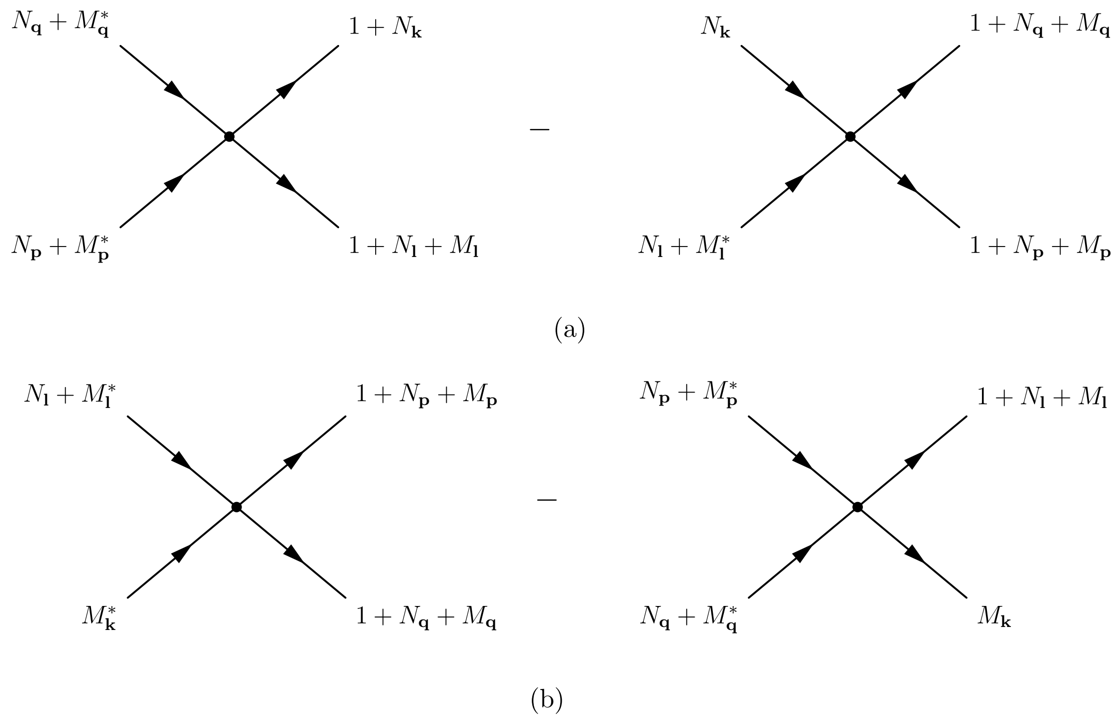
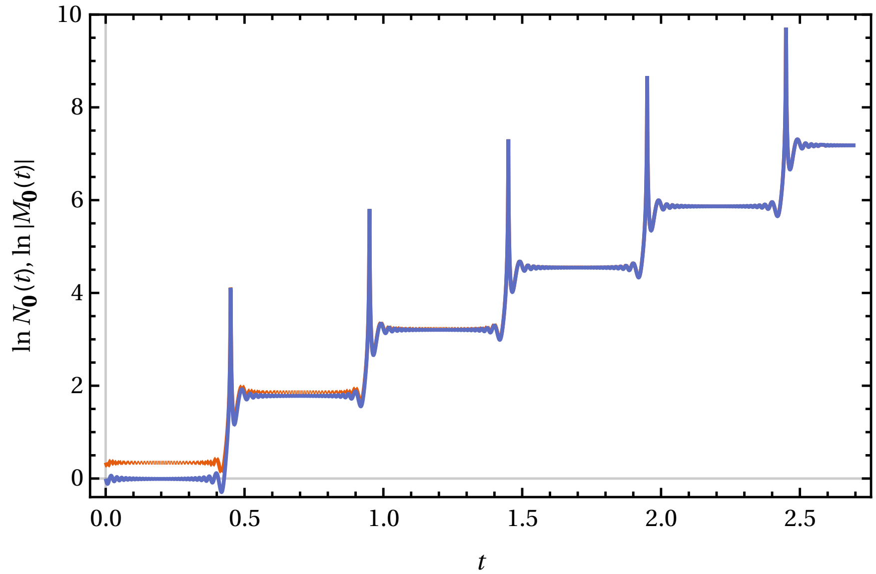


Figure 6.2.3: Feynman diagrams of the two-to-two scattering processes in the number density collision integral $\mathcal{C}^{(N)}_{\mathbf{k}}[N, M, t]$: (a) gain and loss terms, where the "external" momentum $\mathbf{k}$ is associated with the number density; (b) gain and loss terms, where the "external" momentum $\mathbf{k}$ is associated with the (complex conjugate of the) pair correlation. Note that the momentum $\mathbf{l} = \mathbf{p} + \mathbf{q} - \mathbf{k}$ is determined by three-momentum conservation at each vertex.

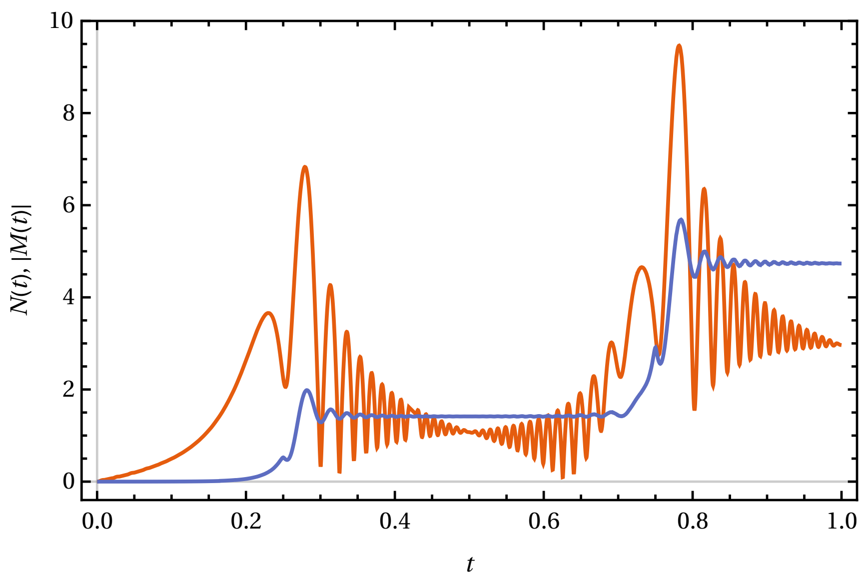
where we have written the equation for $M_{\mathbf{k}}(t)$ in terms of its real and imaginary parts, which is sufficient to determine the evolution of both $M_{\mathbf{k}}(t)$ and $M^*_{\mathbf{k}}(t)$. These Boltzmann equations cannot be solved analytically, and thus in order to study the evolution of scalar particle number density during preheating, one must adopt a numerical approach, the details of which we will discuss in the next section.

6.3 Evolving the system: a numerical analysis

In this section, we present numerical solutions to the Boltzmann equations in eq. (6.2.46) for the very early stages of preheating. Throughout this analysis, we have focused on the case of broad resonance (i.e. that which occurs over a broad range of momenta and corresponds to the condition $q = \frac{g\varphi_0^2)}{8m_\phi^2} \gg 1$), since reheating becomes extremely efficient, thereby enabling a relatively large occupancy for each momentum mode to build up in just a few oscillations of the inflaton field. In this case, one expects the effect of the collision terms to be more pronounced at these early stages compared to the regime of narrow resonance. With this in mind, we choose



(a)



(b)

Figure 6.2.4: (a) Time evolution (measured in units of $m_\phi/2\pi$) of the χ particle number density $N_{\mathbf{k}}$ (blue) and pair correlation $|M_{\mathbf{k}}|$ (orange) for the mode $\mathbf{k} = \mathbf{0}$ in the regime of broad resonance ($q \sim 10^3$) for the collisionless case $\lambda = 0$. (b) Evolution of the integrated number density (blue) and pair correlation (orange), i.e. $N(t) = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} N_{\mathbf{k}}$ and $|M(t)| = \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} |M_{\mathbf{k}}|$ up to $t = 2\pi/m_\phi$.

the model parameters as follows: $\varphi_0 = \frac{M_{\rm Pl}}{10} = 10^5\, m_\phi$, $m_\chi = \frac{m_\phi}{10}$ and $g = 5 \times 10^{-7}$, such that $q \sim \mathcal{O}(10^3)$.

The results that are discussed in this section were obtained by means of a fourth-order Runge-Kutta differential solver implemented in Mathematica and involving five non-trivial phase-space integrals over the magnitudes of the momenta $\mathbf{p}$ and $\mathbf{q}$, the relative angle between them and the relative angles of one of these momenta to the external momentum $\mathbf{k}$.[24] We remark that the approximations made in order to reduce the solutions for the mode functions $\chi_{\boldsymbol{\kappa}}(t', t)$ (cf. eq. (6.2.21)) to a form yielding (quasi-)energy-conserving Dirac delta functions introduce an error of at most $\sim 15\%$ to the collision integrals (see §6.2.2). This is, however, anticipated to be a global error, rather than a relative error between the contributions to each collision integral, and therefore is expected to have little impact on the inferences that follow. Let us first analyse the collision-free (i.e. in the absence of thermalisation) evolution

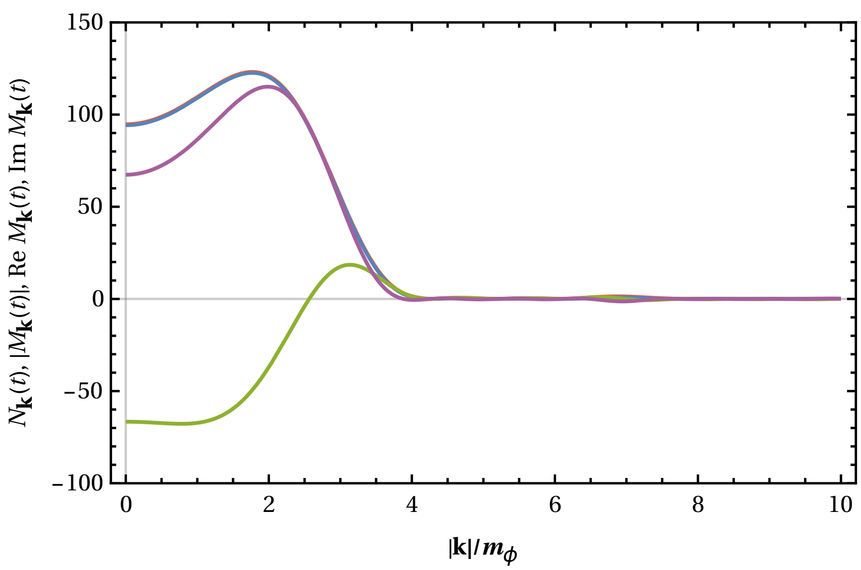


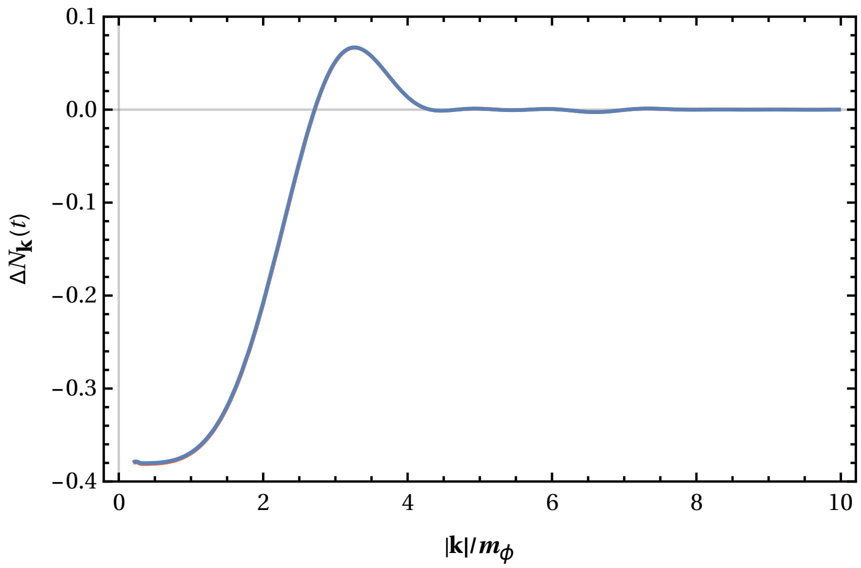
Figure 6.3.1: Plots of $N_{\mathbf{k}}$ (blue), $|M_{\mathbf{k}}|$ (orange), $\mathrm{Re}\, M_{\mathbf{k}}$ (green) and $\mathrm{Im}\, M_{\mathbf{k}}$ (purple) as a function of $|\mathbf{k}|/m_\phi$ at $t = 4\pi/m_\phi$.

[24]The $SO(3)$ symmetry of the collision integrals [eq. (6.2.43)] to further reduce the complexity of the problem. Indeed, we can always orient our reference frame such that $\mathbf{k}$ lies purely along the z-axis, and $\mathbf{q}$ lies purely in the xz-plane. In doing so, by working in spherical polar coordinates, we can reduce the number of variables that we need to integrate over from six to five.

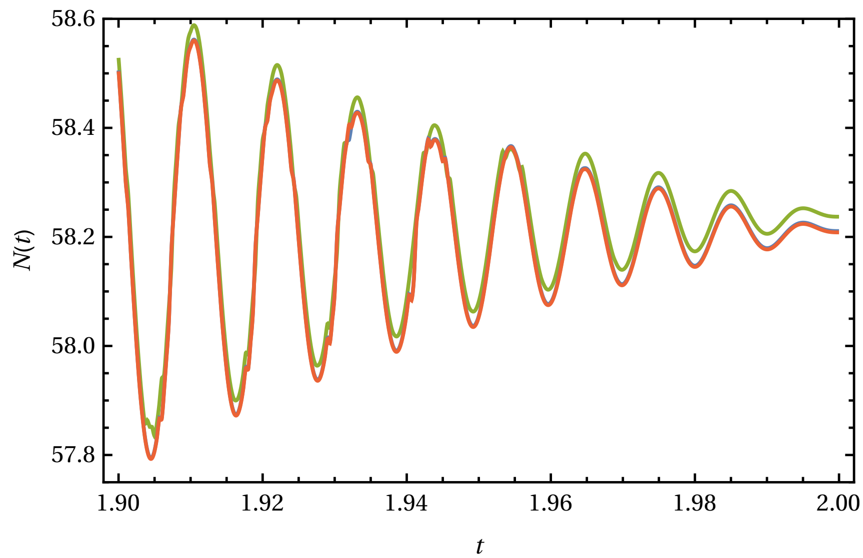
of the number density and pair correlations. Figure 6.2.4a illustrates the evolution of the (natural) logarithm of the number density and the pair correlation for the zero mode $|\mathbf{k}| = 0$ over the first three inflaton oscillations for the collisionless case, i.e. with $\lambda = 0$. We see that the density matrix approach correctly captures the resonant particle production, and it therefore provides a framework within which to study non-adiabatic particle production that is complementary to existing methods based on solving the Mathieu equation for the field modes. In particular, we see the characteristic jumps in the number density, occurring each time the inflaton field passes through zero. Note that the adiabatic approximation is satisfied between each jump, i.e. $\frac{|\dot{\omega}_{\mathbf{k}}|}{\omega_{\mathbf{k}}^2} < 1$, such that $N_{\mathbf{k}}$ is an approximate adiabatic invariant and remains roughly constant. From eq. (6.2.46), it is clear that the pair correlations act to source the growth in the number density. This is corroborated by the numerics, where, in fig. 6.2.4b, we see that the growth in the pair correlations precedes the growth in the number density. We reiterate that the presence of the pair correlations plays a crucial role in the non-adiabatic particle production.

Figure 6.3.1 shows plots of $N_{\mathbf{k}}$, $|M_{\mathbf{k}}|$, $\mathrm{Re}\, M_{\mathbf{k}}$ and $\mathrm{Im}\, M_{\mathbf{k}}$ as a function of $|\mathbf{k}|/m_\phi$ at $t = 4\pi/m_\phi$. We see that the number density is non-zero for a continuous range of momenta, typically within the interval $0 \leq |\mathbf{k}| \lesssim m_\phi(q/4)^{1/4} \approx\ 4m_\phi$, as is expected for broad resonance (cf. §5.1.2). Importantly, fig. 6.3.1 also confirms the expected result that $N_{\mathbf{k}}$ and $|M_{\mathbf{k}}|$ are the same order of magnitude throughout the preheating phase. In fact, it is evident from fig. 6.3.1 that they are almost identical.

We now turn our attention to the collisional cases, i.e. $\lambda \neq 0$. In the first instance, we set the pair correlations M and M^* to zero in the collision terms in eq. (6.2.46), so as to be able to isolate their impact. Figure 6.3.2a shows the number density as a function of $|\mathbf{k}|/m_\phi$ for the collisionless case and collisional cases with $\lambda \in \{0.1, 0.2\}$, neglecting the pair correlations. While the maximum difference is at the sub-percent level ($\sim 0.3\%$) for both the $\lambda = 0.1$ and $\lambda = 0.2$ cases, we see that the collisions lead to a suppression of the particle production, corresponding to a reduction in the efficiency of the resonance, as we might expect. This suppression is also visible in fig. 6.3.2b, where we show the time-evolution of the collisionless and collisional number densities for the same case. These results illustrate that the collisions have an effect (albeit initially small) fairly soon after the onset of preheating. Importantly, one would expect these effects to become more pronounced as preheating proceeds and as the number density grows.



(a)



(b)

Figure 6.3.2: (a) Difference $\Delta N_{\mathbf{k}} = N_{\mathbf{k}}^{\lambda} - N_{\mathbf{k}}^{\lambda=0}$ between the collisional and collisionless number densities as a function of $|\mathbf{k}|/m_\phi$ for $\lambda = 0.1$ (blue) and $\lambda = 0.2$ (orange) at $t = 4\pi/m_\phi$ and in the case where only the number density $N_{\mathbf{k}}$ participates in the collision integral. (b) Time evolution of the total number $N(t)$ for $\lambda = 0$ (green), $\lambda = 0.1$ (blue) and $\lambda = 0.2$ (orange), where the time evolution is shown in units of $m_\phi/2\pi$ (corresponding to the number of periods of the inflaton condensate). We have truncated the graph near to $t \sim 4\pi/m_\phi$ to make the suppression visible.

In fig. 6.3.3, we plot the difference between the cases without and with the pair correlations for the collisional case with $\lambda = 0.1$. The impact of the pair correlations is negligible, despite the magnitudes of $N_{\mathbf{k}}$ and $|M_{\mathbf{k}}|$ being almost identical. However, by plotting the time-evolution of the integrated collision term

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\Delta N(t) \;=\; \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\,\mathrm{Re}\,\mathcal{C}^{(N)}_{\mathbf{k}}[N,M;t] \tag{6.3.1}$$

for the cases with and without the pair correlations (see fig. 6.3.4), the reason for this negligible impact becomes apparent. Specifically, the terms involving $M_{\mathbf{k}}$ and $M^*_{\mathbf{k}}$ re-

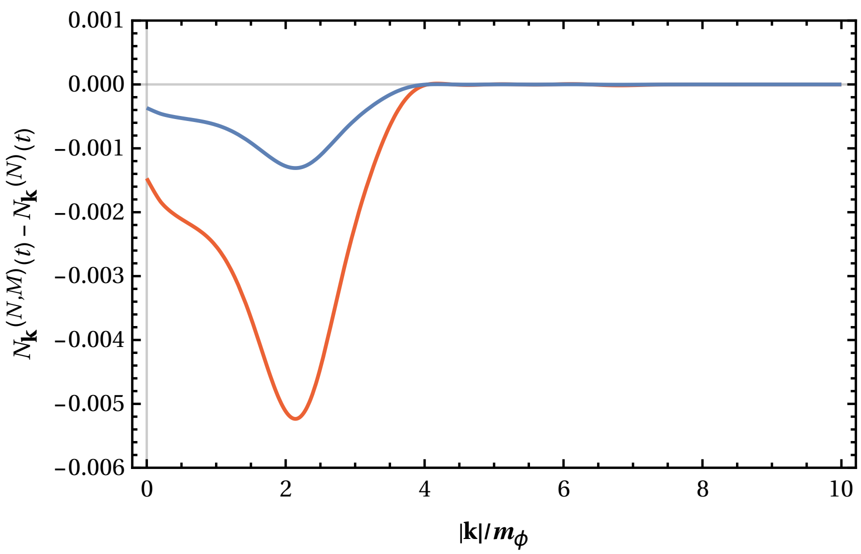


Figure 6.3.3: Difference $N^{(N,M)}_{\mathbf{k}} - N^{(N)}_{\mathbf{k}}$ between the number densities with both $N_{\mathbf{k}}$ and $M_{\mathbf{k}}$ participating ($N^{(N,M)}_{\mathbf{k}}$) in the collision terms, and with only $N_{\mathbf{k}}$ participating ($N^{(N)}_{\mathbf{k}}$) for $\lambda = 0.1$ (blue) and $\lambda = 0.2$ (orange) at $t = 4\pi/m_\phi$.

sult in highly oscillatory contributions, which fluctuate about an average value that is only negligibly different from that of the case where only $N_{\mathbf{k}}$ contributes. This can be traced back to the master equations for $M_{\mathbf{k}}$ and $M^*_{\mathbf{k}}$, obtained from Eqs. (6.2.46b) and (6.2.46c). Both contain an oscillatory contribution with instantaneous period $T \sim 2\pi/\omega_{\mathbf{k}}(t) \ll 1/\Gamma_{\chi\chi\to\chi\chi}$. As such, $M_{\mathbf{k}}$ and $M^*_{\mathbf{k}}$ oscillate on time-scales much shorter than those over which collisional processes take place, and their contribution effectively averages to zero.

Therefore, with the present separation of scales, we can safely ignore any contributions from $M_{\mathbf{k}}$ and $M^*_{\mathbf{k}}$ to the collision integrals in the master equations. On the

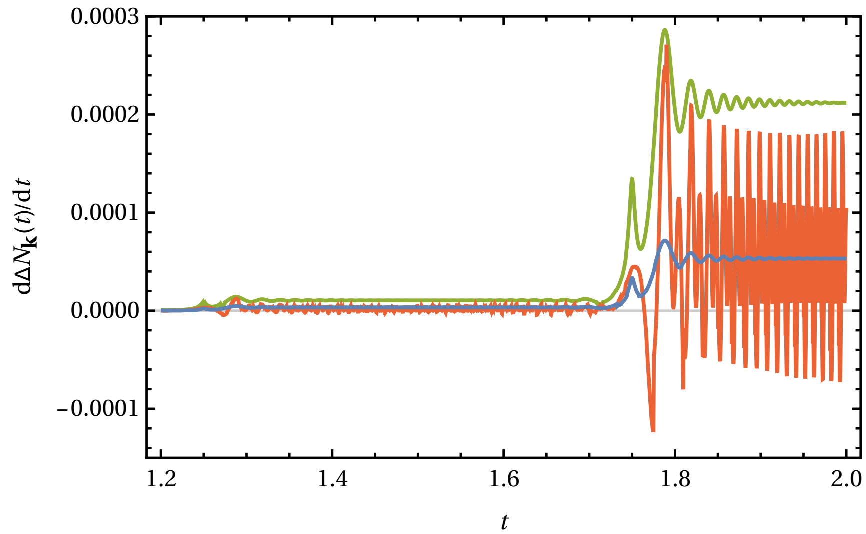
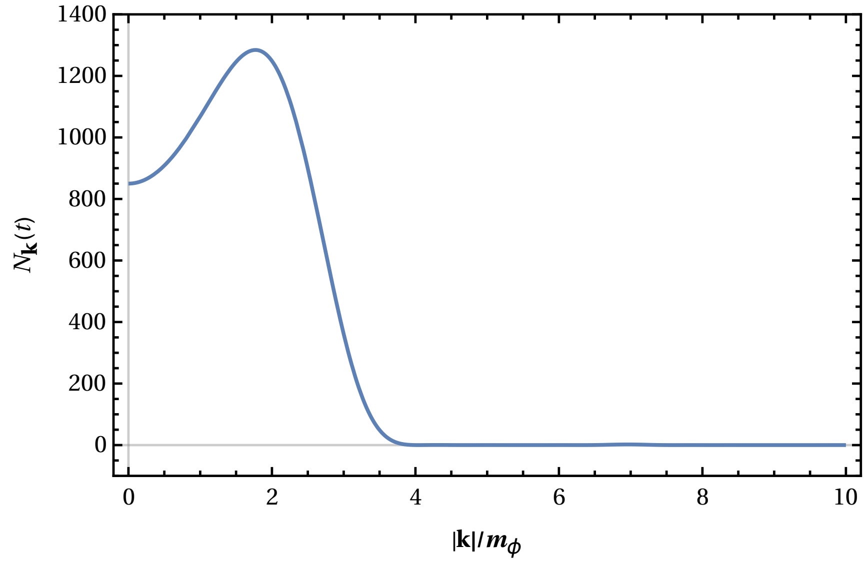


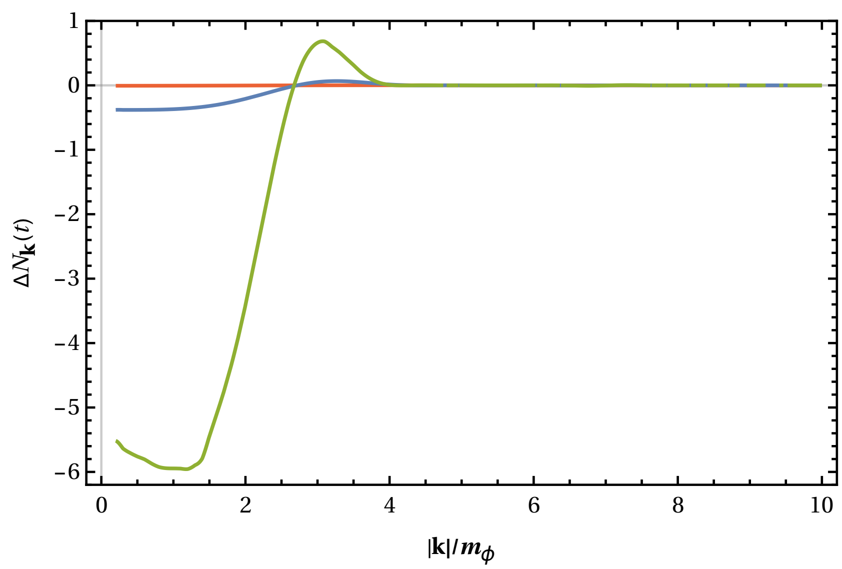
Figure 6.3.4: Behaviour of the collision integral $\mathrm{d}\Delta N(t)/\mathrm{d}t$ as a function of time t (in units of $m_\phi/2\pi$). The blue and green curves correspond to the case where only $N_{\mathbf{k}}$ contributes, for $\lambda = 0.1$ and $\lambda = 0.2$ respectively, and the orange curve, to the case in which $M_{\mathbf{k}}$ and $M^*_{\mathbf{k}}$ are also included. The graph has been truncated at $t = 1.2$, since the collision integral is negligible beforehand. In particular, it is found that the magnitude of $\mathrm{d}\Delta N(t)/\mathrm{d}t$ for $\lambda = 0.2$ is four times larger than the $\lambda = 0.1$, as would be expected.

other hand, if the collision rate becomes comparable to the rate of oscillation of the pair correlations, one might expect a greater residual effect on the evolution of the number density. However, one might then doubt the applicability of the approximations used here to treat the time-dependence of the phase space, and we leave dedicated studies to future work.

By neglecting the contributions from $M_{\mathbf{k}}$ and $M^*_{\mathbf{k}}$ to the collision integrals, as we have shown to be appropriate for the present choice of parameters, the stability of the numerics is improved. By this means, we were able to evolve the system reliably for three full oscillations of the inflaton field. The collisionless number density is shown in fig. 6.3.5a, having increased in amplitude by an order of magnitude compared with the previous inflaton oscillation shown in fig. 6.3.1. The comparison with the collisional case is presented in fig. 6.3.5b. We see that the suppression of the number density increases with each full oscillation of the inflaton field over the three periods. In particular, we see that after three periods, the maximum relative difference has increased from the sub-percent level ($\sim 0.3\%$) to order 1%, that is the effect of the collision integrals has essentially doubled after only one additional cycle.



(a)



(b)

Figure 6.3.5: (a) Plot of the number density in the collisionless case $N_{\mathbf{k}}^{\lambda=0}$ at $t = 0.9 \times 6\pi/m_\phi$. (b) The difference $\Delta N_{\mathbf{k}} = N_{\mathbf{k}}^{\lambda} - N_{\mathbf{k}}^{\lambda=0}$ between the collisional and collisionless number density for $\lambda = 0.2$ at $t = 2\pi/m_\phi$ (orange), $t = 4\pi/m_\phi$ (blue) and $t = 0.9 \times 6\pi/m_\phi$ (green) in the case where only the number density $N_{\mathbf{k}}$ participates in the collision integral.

6.4 Discussion

In this chapter we have used a density matrix approach to construct a toy model preheating theory, in which we simultaneously account for the effects of both resonant particle production and the leading-order two-to-two scattering processes that drive thermalisation. In particular, within the density matrix formalism, we have shown that one can derive a (self-consistent) set of quantum Boltzmann equations, which are able to describe the evolution of an ensemble of self-interacting scalar particles that are subject to an oscillating mass term. During the preheating phase of the early universe, these equations can be used to determine the evolution of the number density and pair correlations for a scalar field coupled to the inflaton, while accounting also for collisional processes.

Through a numerical analysis of this toy model of preheating, we solved for this evolution over the first few inflaton oscillations, through which we illustrated the importance of the pair correlations during preheating. Specifically, it was evident from the analysis in §6.3 that they play a crucial role in mediating the non-adiabatic particle production. Importantly, from this observation, we were able to establish an important generalization of the previous observation by Morikawa and Sasaki [262] that small perturbations to such a system would destroy coherences between particle and anti-particle states. Indeed, it follows that any processes that cause such pair correlations to decohere will suppress, or shut off, the resonant particle production. Moreover, we demonstrated that the pair correlations are of comparable magnitude to the number density throughout preheating and cannot, therefore, be neglected a priori in the collision integrals that precipitate thermalisation. In spite of the latter observation, we noted from the Boltzmann equations, that the pair correlations can safely be neglected in the collision integrals when the collision rate is much slower than the rate of free-phase oscillations of the pair correlations. This was confirmed in the numerical data, in which we showed that the contributions of the pair correlations are effectively time-averaged away.

Perhaps most importantly, the numerical analysis conducted in §6.3 suggests that the collision terms have an impact on the resulting number density even in the very early stages of preheating. Specifically, after only three oscillations of the inflaton condensate, we found that the number density starts to become suppressed relative to the collisionless case, with the data indicating an $\mathcal{O}(1\%)$ deviation between their respective magnitudes. This suppression is expected to increase significantly as preheating progresses, which would lead to appreciable deviations between the (magnitudes of) the particle number densities in the collisionful and collisionless cases. This motivates further numerical studies beyond what we have discussed in §6.3 to establish the effect of accounting fully for the thermalisation processes during preheating (and reheating) on the thermal history of the early universe.

We should note, that in the work presented in this chapter, we have considered a simplified toy model of preheating. In a more realistic scenario (as was briefly discussed in §5.2), one would need to account for the effects of the Hubble expansion both on the decay of the inflaton condensate and the structure of the resonance bands. In addition, one would want to account for perturbative inflaton decays, as well as the backreaction of the particle-production processes on the inflaton condensate. In the case of broad resonance, the cosmological expansion results in a stochastic resonance behaviour, wherein the number density increases exponentially on average [98, 114]. This background expansion competes with the effects of backreactions and rescatterings in determining the efficiency of the resonant particle production [114]. One may therefore anticipate the effects of collisional processes during the production phases to remain significant in more realistic scenarios. Before exploring these possibilities, we note that it would be constructive to make direct comparisons between the present density matrix approach, where the relation to canonical quantities such as the number density is manifest, and others based on the closed-time-path formalism of non-equilibrium quantum field theory, where one must instead employ, e.g., quasi-particle approximations in order to extract physical observables. With this is mind, in future research it would be interesting, albeit technically challenging, to reintroduce the effects of Hubble expansion and backreaction, with the aim of working towards a more realistic model.

Chapter 7

Epilogue

7.1 Summary

The research elaborated on in this thesis is separated into two parts: the first focusing on the CCP, a problem in late universe cosmology; and the second on early universe cosmology, analysing the nature of post-inflationary preheating, in particular the thermalisation process of the produced particles.

In §1, we gave a brief introduction to our current best theory of gravity, GR, an important ingredient in our endeavour to understand our universe. We touched upon its theoretical robustness, as well as its successes and failures, both theoretically and experimentally. We then finished off the chapter by briefly examining some possible solutions to the problems faced by GR, and how, in the case of the CCP, the lack of a compelling solution in the field theory sector, i.e. from the SM (or even Supersymmetry), has prompted research into modifications of GR.

Following on from this, in §2 we discussed problems pertaining to the early universe, in particular the necessity for inflation, and the subsequent reheating phase. We highlighted the intricacies of post-inflationary reheating, specifically the non-perturbative nature of the inflaton condensate in the early stages, and its effects on particle-production. In §3, we moved on to analyse the CCP in more detail. Indeed, the CCP remains a thorn in the side of modern physics, presently still evading a complete solution. We elucidated on how it is intimately related to the radiative instability of vacuum energy loop contributions from massive particles. Moreover, on how this results in higher order contributions not being significantly suppressed relative to leading order terms, requiring one to repeatedly fine-tune the classical parameter appearing in the action for GR, in order for our theory to match the present day observable universe. There is clearly something wrong on the theory side here, and there have been many attempts to provide an answer.

As we went on to discuss in §4, one particular (minimal) approach is to introduce a new scalar degree of freedom ϕ into the gravity sector, that is able to self-tune, so as to screen the effects of the vacuum energy on spacetime curvature, from the matter sector. We elaborated on the subtleties that are involved here, most importantly, that in order to evade Weinberg's famous no-go theorem, preventing such self-adjustment mechanisms, one has to allow ϕ to dynamically self-tune, i.e., break Poincaré invariance of the vacuum solution at the level of the scalar field, such that $\phi = \phi(t)$. Such an approach naturally lead to the construction of a class of self-tuning theories, known as the "Fab-Four".

After reviewing the Fab-Four, we moved on to consider a generalisation, in which matter interacts with the gravitational sector via a disformal coupling. In this scenario, matter becomes directly coupled to the self-tuning field, however, one can always recast the theory in a different representation, the so-called Horndeski frame, such that this direct coupling is removed. Evaluating the Horndeski Lagrangian on an FRW background, we ascertained expressions for the equation of motion for ϕ, as well as the gravitational Hamiltonian. Upon passing the theory through the self-tuning filter (established in the derivation of the Fab-Four), we were able to determine the necessary conditions for our disformally coupled theory to self-tune. Importantly, we were able to show that the theory reduces to the Fab-Four, in the conformal limit. Moreover, we determined a strong condition on the form of the conformal function in the disformal transformation of the metric, namely, that it cannot be a function of the canonical momentum $X = -\frac{1}{2}(\partial\phi)^2$ of the scalar field. Given this, we were able to recast the FRW Horndeski Lagrangian in manifestly self-tuning form. Finally, we provided a particular solution to the self-tuning equations of motion, proving that the self-tuning solution set is non-trivial.

In the second half of this thesis, we moved on to an analysis of early universe cosmology. In particular, we focused on the post-inflationary reheating phase, in which the vast energy density stored in the inflaton condensate is transferred back to matter fields. This precipitates a radiation dominated epoch, and enables a transition to the standard HBB model, whose description of the latter stages of the early universe agrees well with experimental data.

In §5, we discussed in detail the non-perturbative behaviour of the inflaton condensate at the start of reheating, and how this can induce a period of highly efficient, rapid particle production via parametric resonance of fields coupled to the inflaton. First reviewing previous analyses of this early stage of reheating, known as preheating, we then commented on the necessity of thermalisation during this period, and throughout reheating, in order for the produced particles to equilibriate, eventually reaching a so-called reheat temperature and enabling a transition to the standard HBB model. This lead us into §6, where we started by motivating the need for a thorough understanding of the post-inflationary reheating an thermalisation processes. Specifically, their importance in determining the early thermal history of the universe. Whilst acknowledging the research that has already been conducted in understanding the thermalisation process, we noted that there is still some way to go to attain a more complete knowledge. In particular, the role and impact of the (pair) correlations between particle and anti-particle states, which are present during the preheating phase, is less well understood.

With this in mind, we were motivated to develop a better understanding of the thermalisation process during preheating. In §6.1, we adopted a density matrix approach that enabled us to derive a (self-consistent) system of quantum Boltzmann equations. In doing so, one can go beyond the usual mode-function analysis, accounting for both the resonant particle production and the collisional processes simultaneously. Moreover, we were able to show the pivotal role of the pair correlations in the particle production, and that without them resonant production simply cannot occur. Importantly, this means that any processes that cause such pair correlations to decohere will suppress, or shut off, the resonant particle production. The Boltzmann equations further highlighted the fact that the magnitudes of the number density and pair correlations are similar in value during preheating. As such, one would expect them to play equally important roles in the collision integral precipitating thermalisation.

An analytic solution to the Boltzmann equations is not possible, thus necessitating a numerical analysis. In §6.3 we discussed the results of this analysis, in which we were able to solve the system of equations over the first few inflaton oscillations for an illustrative set of benchmark parameters. We first considered the collisionless case, in which we showed that the density matrix approach correctly captures the resonant particle production. Accordingly, it provides a framework within which to study non-adiabatic particle production that is complementary to existing methods based on solving the Mathieu equation for the field modes. Importantly, the data illustrated the crucial presence of the pair correlations in order for non-adiabatic particle production to occur. Furthermore, we confirmed that their magnitude is negligibly different from that of the number density N during the preheating phase.

We then turned our attention to the collision cases, initially setting the pair correlations M and M^* to zero in the collision terms so as to isolate their impact. We found that, while the maximum difference between the collisionless and collisionful number densities is at the sub-percent level, the collisions nevertheless lead to a suppression of the particle production, corresponding to a reduction in the efficiency of the resonance. This served to show that the collisions have an effect (albeit initially small) fairly soon after the onset of preheating. Importantly, one would expect these effects to become more pronounced as preheating proceeds and as the number density grows.

Interestingly, upon including the pair correlations in the collision integrals, we found that their presence had a negligible effect on evolution of the number density. The reason for this is that the contributions from M and M^* are highly oscillatory, fluctuating about an average value that is only negligibly different from that of the case where only N contributes. This arises from their free phase evolution being much faster than the collision rate of the produced particles, resulting in their contributions to the collision integral essentially averaging to zero. With this hierarchy in place one can safely ignore any contributions from M and M^* to the collision integrals. That being said, if the collision rate becomes comparable to the rate of oscillation of the pair correlations, one might expect a greater residual effect on the evolution of the number density.

Finally, upon neglecting contributions from M and M^* to the collision integrals, we found that the stability of the numerics is improved. As a result, we were able to evolve the system reliably for three full oscillations of the inflaton field. In doing so, we showed how the suppression of the number density increases with each full oscillation of the inflaton field over the three periods. In particular, we found that after one additional (full) oscillation of the inflaton field, the the maximum relative difference between the collisionless and collisionful number densities increased from the sub-percent level ($\sim 0.3\%$) to order 1%. This suggests that a much larger deviation might accumulate as preheating progresses.

7.2 Future directions

Both of the research projects in this thesis present interesting results, but also pose some intriguing questions, opening up possible future avenues of research. We shall discuss some of these briefly in this section.

In §4.2, we developed a generalisation of Fab-Four theory, with the aim of providing a disformally self-tuning solution to the CCP. We showed that, on a cosmological FRW background, it is possible to achieve this, so long as certain constraints are placed on the form of the disformal transformation of the Jordan-frame metric. However, as remarked upon in §4.3, there is still work to be done on this disformal generalisation of Fab-Four theory. Unlike the Fab-Four, we have yet to determine a fully covariant form for the theory. As such, one avenue of future research would be to promote the theory to covariant form. This would be a more technically challenging task than in the Fab-Four case, due to the presence of the arbitrary disformal function $B(\phi, X)$ (which is zero in Fab-Four). Consequently, the path is less clear, nevertheless, it would be worth analysing the plethora of curvature invariants that one can construct in Horndeski theory, and ascertain whether it is possible to massage the disformally self-tuning Lagrangian into a form that is manifestly a combination of a certain set of invariants (evaluated on an FRW background).

There is also the question of whether the theory is still relevant, due to the recent gravitational wave data. As discussed in §4.3, the results have severely reduced the viable regions of parameter space for many classes of Horndeski theories. That being said, there are loop-holes in the constraints. One possible approach would be to use the EOM for the scalar field ϕ, to express $\ddot{\phi}$ in terms of $\dot{\phi}$, H and $\dot{H}$. One can then use the analog Friedmann equations to express this in terms of the energy density arising from the matter sector. By inserting this into the expression in α_T, that is required to vanish, the problem will then be recast into determining choices of the matter content of the universe for which H evolves in a particular fashion, such that it forces $\alpha_T \approx 0$. Of course, one might argue that such choices for the matter sector may need to be finally tuned, nonetheless, it is still worth investigating. Another avenue to consider, would be to search for solutions for the Horndeski functions K, G_3, G_4 and G_5, such that α_T vanishes independently of matter content of the universe. It may well be the case that such solutions do not exist, or are simply trivial, but again, one should not discard this possiblility without further exploration. Of course, even if non-trivial solutions for the Horndeski functions, it remains to be seem whether they are able to describe a sensible evolution of the universe that matches observations.

Finally, it is known that any disformally coupled Horndeski theory cannot be brought into Horndeski form under a set of field re-definitions. Since the disformally self-tuning Lagrangian is degenerate, and is at least in some sense an extension of Horndeski theory, it may well be that it lies within a certain class of so-called DHOST theory. This would be an interesting hypothesis to test, since DHOST theories are less constrained by the gravitational wave data, thus potentially opening up a larger viable region of parameter space for the disformal Fab-Four. However, one would also need to adhere to the graviton decay constraints placed by Creminelli *et al.* [207], which may well rule out the disformal Fab-Four as a viable model.

In the latter half of this thesis, we moved on to consider the reheating phase of the early universe. In particular, we were interested in its early stages, so-called preheating, and how the effects of thermalisation play a role on the evolution of particle number densities during this phase. In §6, we derived a set of quantum Boltzmann equations describing the evolution of a scalar particle number density throughout preheating. To conduct a numercial analysis we made several assumptions, in order to reduce the complexity of the system. This enabled us to determine some interesting results, in particular highlighting the importance of pair correlations, and the effects of thermalisation during the preheating phase, but also left some open questions.

So far, we have evolved the system for three full oscillations of the inflaton condensate. In future research, one would aim to push this further and determine, to a fuller extent, how the effects thermalisation impact the evolution of particle number densities throughout preheating. This would necessarily involve the construction of a more robust code that is able to handle the increasingly computationally taxing numerical integrations that arise as preheating progresses. Moreover, the numerical results determined in §6.3 were done under the assumption of a separation of time-scales between the particle production and collisional processes, namely, that the latter occurs at a much slower rate than the former. This enabled us to employ a Wigner-Weisskoff approximation for the collision integrals, resulting in WKB-like approximate solutions for the mode functions. An initial extension would therefore be to relax this approximation, and determine full solutions for the mode functions, such that one can better approximate the behaviour of the collision integrals during intervals of non-adiabaticity. This would, however, necessiate one to numerical solve the evolution equations for the mode functions.

Assuming that this can be done, it would then be interesting to relax the hierarchy between the production and collision processes. One expects that the pair correlations would play a more significant role in the collision processes as the particle collision rate and free phase oscillation rate of the the pair correlations become the same order of magnitude. Although the analysis in such a scenario would be technically challenging, one might expect to see much larger deviations between the evolution of the collisionless and collisionful number densities developing during the preheating phase.

Finally, in this thesis, we have considered a toy model preheating theory to study the effects of thermalisation. Whilst this was certainly sufficient to provide us with some enlightening results, it is nevertheless a simplification of the actual theory. With the aim of working towards a more realistic scenario in mind, in future work, one could consider introducing some of the additional effects present in the early universe. For example, a starting point would be to account for Hubble expansion, which would cause the inflaton condensate to redshift, as well as affecting the resonant particle production process. Furthermore, one could introduce the backreactions of the produced particles on the condensate, and in addition account for its decay due to the transfer of energy to the coupled matter fields.[1] In the regime of broad resonance, although complicating the process, such effects serve to stabilise preheating, and so one may therefore anticipate the effects of collisional processes during the production phases to remain significant in more realistic scenarios.

[1] In reality, this is much easier said than done, and so would have to be a gradual process, involving several stages in which one progresses by introducing one effect at a time.

Appendix

Further details on the preheating Boltzmann equations

Here we give the full expressions for the component statistical weight functions $f^{(N)}_{(j);\mathbf{p},\mathbf{q},\mathbf{k}}[N,M;t]$ and $f^{(M)}_{(j);\mathbf{p},\mathbf{q},\mathbf{k}}[N,M;t]$ $(j=1,2,3,4)$, in the collision integrals (given by eq. (6.2.39)), which we neglected in §6.2.2 for brevity. These are as follows (for N and M respectively),

$$\begin{aligned}
f^{(N)}_{(1);\mathbf{p},\mathbf{q},\mathbf{k}} \;=\;& (1+N_{\mathbf{k}})\left(N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)(1+N_{\mathbf{p+q-k}}+M_{\mathbf{p+q-k}}) \\
&- N_{\mathbf{k}}\left(1+N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(1+N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)(N_{\mathbf{p+q-k}}+M_{\mathbf{p+q-k}}) \\
&+ M^*_{\mathbf{k}}(1+N_{\mathbf{p}}+M_{\mathbf{p}})(1+N_{\mathbf{q}}+M_{\mathbf{q}})\left(N_{\mathbf{p+q-k}}+M^*_{\mathbf{p+q-k}}\right) \\
&- M^*_{\mathbf{k}}(N_{\mathbf{p}}+M_{\mathbf{p}})(N_{\mathbf{q}}+M_{\mathbf{q}})\left(1+N_{\mathbf{p+q-k}}+M^*_{\mathbf{p+q-k}}\right)\ ,
\end{aligned}$$

$$\begin{aligned}
f^{(N)}_{(2);\mathbf{p},\mathbf{q},\mathbf{k}} \;=\;& \frac{1}{3}\Big[(1+N_{\mathbf{k}})(1+N_{\mathbf{p}}+M_{\mathbf{p}})(1+N_{\mathbf{q}}+M_{\mathbf{q}})(1+N_{\mathbf{-k-p-q}}+M_{\mathbf{-k-p-q}}) \\
&- N_{\mathbf{k}}(N_{\mathbf{p}}+M_{\mathbf{p}})(N_{\mathbf{q}}+M_{\mathbf{q}})(N_{\mathbf{-k-p-q}}+M_{\mathbf{-k-p-q}}) \\
&+ M^*_{\mathbf{k}}\left(N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)\left(N_{\mathbf{-k-p-q}}+M^*_{\mathbf{-k-p-q}}\right) \\
&- M^*_{\mathbf{k}}\left(1+N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(1+N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)\left(1+N_{\mathbf{-k-p-q}}+M^*_{\mathbf{-k-p-q}}\right)\Big]\ ,
\end{aligned}$$

$$\begin{aligned}
f^{(N)}_{(3);\mathbf{p},\mathbf{q},\mathbf{k}} \;=\;& \frac{1}{3}\Big[(1+N_{\mathbf{k}})\left(N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)\left(N_{\mathbf{k-p-q}}+M^*_{\mathbf{k-p-q}}\right) \\
&- N_{\mathbf{k}}\left(1+N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(1+N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)\left(1+N_{\mathbf{k-p-q}}+M^*_{\mathbf{k-p-q}}\right) \\
&+ M^*_{\mathbf{k}}(1+N_{\mathbf{p}}+M_{\mathbf{p}})(1+N_{\mathbf{q}}+M_{\mathbf{q}})(1+N_{\mathbf{k-p-q}}+M_{\mathbf{k-p-q}}) \\
&- M^*_{\mathbf{k}}(N_{\mathbf{p}}+M_{\mathbf{p}})(N_{\mathbf{q}}+M_{\mathbf{q}})(N_{\mathbf{k-p-q}}+M_{\mathbf{k-p-q}})\Big]\ ,
\end{aligned}$$

$$
\begin{aligned}
f^{(N)}_{(4);\mathbf{p},\mathbf{q},\mathbf{k}} \;=\;& (1+N_{\mathbf{k}})\left(N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)(1+N_{\mathbf{q}}+M_{\mathbf{q}})(1+N_{\mathbf{p}-\mathbf{k}-\mathbf{q}}+M_{\mathbf{p}-\mathbf{k}-\mathbf{q}}) \\
&-\; N_{\mathbf{k}}\left(1+N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)(N_{\mathbf{q}}+M_{\mathbf{q}})(N_{\mathbf{p}-\mathbf{k}-\mathbf{q}}+M_{\mathbf{p}-\mathbf{k}-\mathbf{q}}) \\
&+\; M^*_{\mathbf{k}}(1+N_{\mathbf{p}}+M_{\mathbf{p}})\left(N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)\left(N_{\mathbf{p}-\mathbf{k}-\mathbf{q}}+M^*_{\mathbf{p}-\mathbf{k}-\mathbf{q}}\right) \\
&-\; M^*_{\mathbf{k}}(N_{\mathbf{p}}+M_{\mathbf{p}})\left(1+N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)\left(1+N_{\mathbf{p}-\mathbf{k}-\mathbf{q}}+M^*_{\mathbf{p}-\mathbf{k}-\mathbf{q}}\right) \;,
\end{aligned}
$$

$$
\begin{aligned}
f^{(M)}_{(1);\mathbf{p},\mathbf{q},\mathbf{k}} \;=\;& N_{\mathbf{k}}(1+N_{\mathbf{p}}+M_{\mathbf{p}})(1+N_{\mathbf{q}}+M_{\mathbf{q}})\left(N_{\mathbf{p}+\mathbf{q}-\mathbf{k}}+M^*_{\mathbf{p}+\mathbf{q}-\mathbf{k}}\right) \\
&-\; (1+N_{\mathbf{k}})(N_{\mathbf{p}}+M_{\mathbf{p}})(N_{\mathbf{q}}+M_{\mathbf{q}})\left(1+N_{\mathbf{p}+\mathbf{q}-\mathbf{k}}+M^*_{\mathbf{p}+\mathbf{q}-\mathbf{k}}\right) \\
&+\; M_{\mathbf{k}}\left(N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)(1+N_{\mathbf{p}+\mathbf{q}-\mathbf{k}}+M_{\mathbf{p}+\mathbf{q}-\mathbf{k}}) \\
&-\; M_{\mathbf{k}}\left(1+N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(1+N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)(N_{\mathbf{p}+\mathbf{q}-\mathbf{k}}+M_{\mathbf{p}+\mathbf{q}-\mathbf{k}}) \;,
\end{aligned}
$$

$$
\begin{aligned}
f^{(M)}_{(2);\mathbf{p},\mathbf{q},\mathbf{k}} \;=\;& \frac{1}{3}\Big[N_{\mathbf{k}}\left(N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)\left(N_{-\mathbf{k}-\mathbf{p}-\mathbf{q}}+M^*_{-\mathbf{k}-\mathbf{p}-\mathbf{q}}\right) \\
&-\; (1+N_{\mathbf{k}})\left(1+N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(1+N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)\left(1+N_{-\mathbf{k}-\mathbf{p}-\mathbf{q}}+M^*_{-\mathbf{k}-\mathbf{p}-\mathbf{q}}\right) \\
&+\; M_{\mathbf{k}}(1+N_{\mathbf{p}}+M_{\mathbf{p}})(1+N_{\mathbf{q}}+M_{\mathbf{q}})(1+N_{-\mathbf{k}-\mathbf{p}-\mathbf{q}}+M_{-\mathbf{k}-\mathbf{p}-\mathbf{q}}) \\
&-\; M_{\mathbf{k}}(N_{\mathbf{p}}+M_{\mathbf{p}})(N_{\mathbf{q}}+M_{\mathbf{q}})(N_{-\mathbf{k}-\mathbf{p}-\mathbf{q}}+M_{-\mathbf{k}-\mathbf{p}-\mathbf{q}})\Big] \;,
\end{aligned}
$$

$$
\begin{aligned}
f^{(M)}_{(3);\mathbf{p},\mathbf{q},\mathbf{k}} \;=\;& \frac{1}{3}\Big[N_{\mathbf{k}}(1+N_{\mathbf{p}}+M_{\mathbf{p}})(1+N_{\mathbf{q}}+M_{\mathbf{q}})(1+N_{\mathbf{k}-\mathbf{p}-\mathbf{q}}+M_{\mathbf{k}-\mathbf{p}-\mathbf{q}}) \\
&-\; (1+N_{\mathbf{k}})(N_{\mathbf{p}}+M_{\mathbf{p}})(N_{\mathbf{q}}+M_{\mathbf{q}})(N_{\mathbf{k}-\mathbf{p}-\mathbf{q}}+M_{\mathbf{k}-\mathbf{p}-\mathbf{q}}) \\
&+\; M_{\mathbf{k}}\left(N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)\left(N_{\mathbf{k}-\mathbf{p}-\mathbf{q}}+M^*_{\mathbf{k}-\mathbf{p}-\mathbf{q}}\right) \\
&-\; M_{\mathbf{k}}\left(1+N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)\left(1+N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)\left(1+N_{\mathbf{k}-\mathbf{p}-\mathbf{q}}+M^*_{\mathbf{k}-\mathbf{p}-\mathbf{q}}\right)\Big] \;,
\end{aligned}
$$

$$
\begin{aligned}
f^{(M)}_{(4);\mathbf{p},\mathbf{q},\mathbf{k}} \;=\;& N_{\mathbf{k}}(1+N_{\mathbf{p}}+M_{\mathbf{p}})\left(N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)\left(N_{\mathbf{p}-\mathbf{k}-\mathbf{q}}+M^*_{\mathbf{p}-\mathbf{k}-\mathbf{q}}\right) \\
&-\; (1+N_{\mathbf{k}})(N_{\mathbf{p}}+M_{\mathbf{p}})\left(1+N_{\mathbf{q}}+M^*_{\mathbf{q}}\right)\left(1+N_{\mathbf{p}-\mathbf{k}-\mathbf{q}}+M^*_{\mathbf{p}-\mathbf{k}-\mathbf{q}}\right) \\
&+\; M_{\mathbf{k}}\left(N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)(1+N_{\mathbf{q}}+M_{\mathbf{q}})(1+N_{\mathbf{p}-\mathbf{k}-\mathbf{q}}+M_{\mathbf{p}-\mathbf{k}-\mathbf{q}}) \\
&-\; M_{\mathbf{k}}\left(1+N_{\mathbf{p}}+M^*_{\mathbf{p}}\right)(N_{\mathbf{q}}+M_{\mathbf{q}})(N_{\mathbf{p}-\mathbf{k}-\mathbf{q}}+M_{\mathbf{p}-\mathbf{k}-\mathbf{q}}) \;.
\end{aligned}
$$

Bibliography

[1] A. Einstein, "Die grundlage der allgemeinen relativitätstheorie," *Annalen der Physik* **354** (1916), no. 7 769–822.

[2] S. M. Carroll, "Lecture notes on general relativity," `gr-qc/9712019`.

[3] N. Straumann, *General relativity*. Springer Science & Business Media, 2012.

[4] G. W. Gibbons and S. W. Hawking, "Action Integrals and Partition Functions in Quantum Gravity," *Phys. Rev.* **D15** (1977) 2752–2756.

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

[6] J. Norton, "Einstein, the hole argument and the reality of space," in *Measurement, Realism and Objectivity*, pp. 153–188. Springer, 1987.

[7] C. M. Will, "The Confrontation between General Relativity and Experiment," *Living Rev. Rel.* **17** (2014) 4, `1403.7377`.

[8] M. P. Haugan and C. Lammerzahl, "Principles of equivalence: their role in gravitation physics and experiments that test them," *Lect. Notes Phys.* **562** (2001) 195–212, `gr-qc/0103067`. [,195(2001)].

[9] E. G. Adelberger, B. R. Heckel, and A. Nelson, "Tests of the gravitational inverse-square law," *Annual Review of Nuclear and Particle Science* **53** (2003), no. 1 77–121.

[10] S. S. Shapiro, J. L. Davis, D. E. Lebach, and J. Gregory, "Measurement of the solar gravitational deflection of radio waves using geodetic very-long-baseline interferometry data, 1979–1999," *Physical Review Letters* **92** (2004), no. 12 121101.

[11] **Virgo, LIGO Scientific** Collaboration, B. P. Abbott *et. al.*, "Observation of Gravitational Waves from a Binary Black Hole Merger," *Phys. Rev. Lett.* **116** (2016), no. 6 061102, `1602.03837`.

[12] **LIGO Scientific Collaboration and Virgo Collaboration** Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, *et. al.*, "Gw151226: Observation of gravitational waves from a 22-solar-mass binary black hole coalescence," *Phys. Rev. Lett.* **116** (Jun, 2016) 241103.

[13] **LIGO Scientific and Virgo Collaboration** Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, *et. al.*, "Gw170104: Observation of a 50-solar-mass binary black hole coalescence at redshift 0.2," *Phys. Rev. Lett.* **118** (Jun, 2017) 221101.

[14] B. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, V. Adya, *et. al.*, "Gw170608: Observation of a 19 solar-mass binary black hole coalescence," *The Astrophysical Journal Letters* **851** (2017), no. 2 L35.

[15] **LIGO Scientific Collaboration and Virgo Collaboration** Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, *et. al.*, "Gw170814: A three-detector observation of gravitational waves from a binary black hole coalescence," *Phys. Rev. Lett.* **119** (Oct, 2017) 141101.

[16] **Virgo, Fermi-GBM, INTEGRAL, LIGO Scientific** Collaboration, B. P. Abbott *et. al.*, "Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A," *Astrophys. J.* **848** (2017), no. 2 L13, `1710.05834`.

[17] I. I. Shapiro, "Fourth test of general relativity," *Physical Review Letters* **13** (1964), no. 26 789.

[18] D. Lovelock, "The Einstein tensor and its generalizations," *J. Math. Phys.* **12** (1971) 498–501.

[19] D. Lovelock, "The four-dimensionality of space and the einstein tensor," *J. Math. Phys.* **13** (1972) 874–876.

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

[21] R. M. Wald, "Spin-two fields and general covariance," *Phys. Rev. D* **33** (Jun, 1986) 3613–3625.

[22] S. N. Gupta, "Gravitation and electromagnetism," *Phys. Rev.* **96** (Dec, 1954) 1683–1685.

[23] S. Weinberg, "Photons and gravitons in perturbation theory: Derivation of maxwell's and einstein's equations," *Phys. Rev.* **138** (May, 1965) B988–B1002.

[24] S. Deser, "Self-interaction and gauge invariance," *General Relativity and gravitation* **1** (1970), no. 1 9–18.

[25] R. H. Kraichnan, "Special-relativistic derivation of generally covariant gravitation theory," *Physical Review* **98** (1955), no. 4 1118.

[26] D. G. Boulware and S. Deser, "Classical general relativity derived from quantum gravity," *Annals of Physics* **89** (1975), no. 1 193–240.

[27] S. Weinberg, "Photons and gravitons in s-matrix theory: Derivation of charge conservation and equality of gravitational and inertial mass," *Phys. Rev.* **135** (Aug, 1964) B1049–B1056.

[28] S. Weinberg, *The Quantum theory of elds. Vol. 1: Foundations*. Cambridge University Press, 2005.

[29] G. t Hooft and M. J. Veltman, "One-loop divergencies in the theory of gravitation," *Ann. Henri Poincaré* **20** (1974), no. CERN-TH-1723 69–94.

[30] M. H. Goroff and A. Sagnotti, "The ultraviolet behavior of einstein gravity," *Nuclear Physics B* **266** (1986), no. 3-4 709–736.

[31] A. E. van de Ven, "Two-loop quantum gravity," *Nuclear Physics B* **378** (1992), no. 1 309–366.

[32] J. F. Donoghue, "General relativity as an effective field theory: The leading quantum corrections," *Phys. Rev.* **D50** (1994) 3874–3888, `gr-qc/9405057`.

[33] B. S. DeWitt and G. Esposito, "An introduction to quantum gravity," *International Journal of Geometric Methods in Modern Physics* **5** (2008), no. 01 101–156.

[34] S. Talaganis, T. Biswas, and A. Mazumdar, "Towards understanding the ultraviolet behavior of quantum loops in infinite-derivative theories of gravity," *Class. Quant. Grav.* **32** (2015), no. 21 215017, `1412.3467`.

[35] C. Rovelli, "Loop quantum gravity," *Living reviews in relativity* **11** (2008), no. 1 5.

[36] J. Polchinski, "String theory. vol. 1: An introduction to the bosonic string.," *Cambridge, UK: Univ. Pr* **402** (1998).

[37] J. Polchinski, "Introduction to cosmic f-and d-strings," in *String theory: From gauge interactions to cosmology*, pp. 229–253. Springer, 2005.

[38] D. Tong, "String Theory," `0908.0333`.

[39] D.-W. Chiou, "Loop Quantum Gravity," *Int. J. Mod. Phys.* **D24** (2014), no. 01 1530005, `1412.4362`.

[40] V. C. Rubin and W. K. Ford Jr, "Rotation of the andromeda nebula from a spectroscopic survey of emission regions," *The Astrophysical Journal* **159** (1970) 379.

[41] A. Bosma, "21-cm line studies of spiral galaxies. ii. the distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types.," *The Astronomical Journal* **86** (1981) 1825–1846.

[42] G. L. Bryan and M. L. Norman, "Statistical properties of x-ray clusters: Analytic and numerical comparisons," *Astrophys. J.* **495** (1998) 80, `astro-ph/9710107`.

[43] R. Carlberg, H. K. C. Yee, E. Ellingson, R. Abraham, P. Gravel, S. Morris, and C. J. Pritchet, "Galaxy cluster virial masses and Omega," *Astrophys. J.* **462** (1996) 32, `astro-ph/9509034`.

[44] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich, R. L. Gilliland, C. J. Hogan, S. Jha, R. P. Kirshner, *et. al.*, "Observational evidence from supernovae for an accelerating universe and a cosmological constant," *The Astronomical Journal* **116** (1998), no. 3 1009.

[45] **Supernova Cosmology Project** Collaboration, S. Perlmutter *et. al.*, "Measurements of Omega and Lambda from 42 high redshift supernovae," *Astrophys. J.* **517** (1999) 565–586, `astro-ph/9812133`.

[46] W. Press and M. Davis, "How to identify and weigh virialized clusters of galaxies in a complete redshift catalog," *The Astrophysical Journal* **259** (1982) 449–473.

[47] M. Davis and P. Peebles, "Evidence for local anisotropy of the hubble flow," *Annual Review of Astronomy and Astrophysics* **21** (1983), no. 1 109–130.

[48] A. H. Guth, "The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems," *Phys. Rev.* **D23** (1981) 347–356.

[49] G. Hinshaw, D. Larson, E. Komatsu, D. Spergel, C. Bennett, J. Dunkley, M. Nolta, M. Halpern, R. Hill, N. Odegard, *et. al.*, "Nine-year wilkinson microwave anisotropy probe (wmap) observations: cosmological parameter results," *The Astrophysical Journal Supplement Series* **208** (2013), no. 2 19.

[50] J. L. Sievers, R. A. Hlozek, M. R. Nolta, V. Acquaviva, G. E. Addison, P. A. Ade, P. Aguirre, M. Amiri, J. W. Appel, L. F. Barrientos, *et. al.*, "The atacama cosmology telescope: Cosmological parameters from three seasons of data," *Journal of Cosmology and Astroparticle Physics* **2013** (2013), no. 10 060.

[51] **Planck** Collaboration, N. Aghanim *et. al.*, "Planck 2018 results. VI. Cosmological parameters," `1807.06209`.

[52] R. M. Wald, "Asymptotic behavior of homogeneous cosmological models in the presence of a positive cosmological constant," *Physical Review D* **28** (1983), no. 8 2118.

[53] A. Padilla, "Lectures on the Cosmological Constant Problem," `1502.05296`.

[54] M. Milgrom, "A modification of the newtonian dynamics as a possible alternative to the hidden mass hypothesis," *The Astrophysical Journal* **270** (1983) 365–370.

[55] J. Bekenstein, "Relativistic gravitation theory for the mond paradigm, 2004," *Phys. Rev. D* **70** no. 8.

[56] R. Sundrum, "Fat gravitons, the cosmological constant and submillimeter tests," *Phys. Rev.* **D69** (2004) 044014, `hep-th/0306106`.

[57] Y. Aghababaie, C. P. Burgess, S. L. Parameswaran, and F. Quevedo, "Towards a naturally small cosmological constant from branes in 6-D supergravity," *Nucl. Phys.* **B680** (2004) 389–414, `hep-th/0304256`.

[58] N. Kaloper and A. Padilla, "Sequestering the Standard Model Vacuum Energy," *Phys. Rev. Lett.* **112** (2014), no. 9 091304, `1309.6562`.

[59] N. Kaloper, A. Padilla, D. Stefanyszyn, and G. Zahariade, "Manifestly Local Theory of Vacuum Energy Sequestering," *Phys. Rev. Lett.* **116** (2016), no. 5 051302, `1505.01492`.

[60] C. Charmousis, E. J. Copeland, A. Padilla, and P. M. Saffin, "Self-tuning and the derivation of a class of scalar-tensor theories," *Phys. Rev.* **D85** (2012) 104040, `1112.4866`.

[61] W. T. Emond and P. M. Saffin, "Disformally self-tuning gravity," *Journal of High Energy Physics* **2016** (2016), no. 3 161.

[62] T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, "Modified gravity and cosmology," *arXiv preprint arXiv:1106.2476* (2011).

[63] A. Joyce, B. Jain, J. Khoury, and M. Trodden, "Beyond the Cosmological Standard Model," *Phys. Rept.* **568** (2015) 1–98, `1407.0059`.

[64] G. Lemaître, "Un univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques," in *Annales de la Société scientifique de Bruxelles*, vol. 47, pp. 49–59, 1927.

[65] E. Hubble, "A relation between distance and radial velocity among extra-galactic nebulae," *Proceedings of the National Academy of Sciences* **15** (1929), no. 3 168–173.

[66] A. R. Liddle, "An Introduction to cosmological inflation," in *Proceedings, Summer School in High-energy physics and cosmology: Trieste, Italy, June 29-July 17, 1998*, pp. 260–295, 1999. `astro-ph/9901124`.

[67] S. Dodelson, *Modern cosmology.* Elsevier, 2003.

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

[69] M. S. Turner, "Cosmology: standard and inflationary," in *Particle Physics, Astrophysics and Cosmology, XXIIth SLAC Summer Institute, Ed. J. Chan and L. De Porcel (Conf-9408100)*, 1994.

[70] B. Pandey and S. Sarkar, "Testing homogeneity in the Sloan Digital Sky Survey Data Release Twelve with Shannon entropy," *Mon. Not. Roy. Astron. Soc.* **454** (2015), no. 3 2647–2656, `1507.03124`.

[71] S. Birrer *et. al.*, "H0LiCOW - IX. Cosmographic analysis of the doubly imaged quasar SDSS 1206+4332 and a new measurement of the Hubble constant," `1809.01274`.

[72] A. G. Riess *et. al.*, "Milky Way Cepheid Standards for Measuring Cosmic Distances and Application to Gaia DR2: Implications for the Hubble Constant," *Astrophys. J.* **861** (2018), no. 2 126, `1804.10655`.

[73] **LIGO Scientific, VINROUGE, Las Cumbres Observatory, DES, DLT40, Virgo, 1M2H, Dark Energy Camera GW-E, MASTER** Collaboration, B. P. Abbott *et. al.*, "A gravitational-wave standard siren measurement of the Hubble constant," *Nature* **551** (2017), no. 7678 85–88, `1710.05835`.

[74] W. L. Freedman, "Cosmology at a Crossroads," *Nat. Astron.* **1** (2017) 0121, `1706.02739`.

[75] A. D. Linde, "A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems," *Phys. Lett.* **108B** (1982) 389–393.

[76] A. Albrecht and P. J. Steinhardt, "Cosmology for grand unified theories with radiatively induced symmetry breaking," *Phys. Rev. Lett.* **48** (Apr, 1982) 1220–1223.

[77] E. W. Kolb and M. S. Turner, "The Early Universe," *Front. Phys.* **69** (1990) 1–547.

[78] A. Golovnev, V. Mukhanov, and V. Vanchurin, "Vector Inflation," *JCAP* **0806** (2008) 009, `0802.2068`.

[79] M. R. Setare and V. Kamali, "Warm Vector Inflation," *Phys. Lett.* **B726** (2013) 56–65, `1309.2452`.

[80] R. Emami, S. Mukohyama, R. Namba, and Y.-l. Zhang, "Stable solutions of inflation driven by vector fields," *JCAP* **1703** (2017), no. 03 058, 1612.09581.

[81] J. P. Conlon, "The QCD axion and moduli stabilisation," *JHEP* **05** (2006) 078, hep-th/0602233.

[82] **ATLAS** Collaboration, G. Aad *et. al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett.* **B716** (2012) 1–29, 1207.7214.

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

[84] K. Nakayama and F. Takahashi, "Higgs Chaotic Inflation in Standard Model and NMSSM," *JCAP* **1102** (2011) 010, 1008.4457.

[85] C. Germani and A. Kehagias, "New Model of Inflation with Non-minimal Derivative Coupling of Standard Model Higgs Boson to Gravity," *Phys. Rev. Lett.* **105** (2010) 011302, 1003.2635.

[86] D. S. Goldwirth and T. Piran, "Inhomogeneity and the onset of inflation," *Phys. Rev. Lett.* **64** (Jun, 1990) 2852–2855.

[87] D. S. Goldwirth and T. Piran, "Spherical inhomogeneous cosmologies and inflation: Numerical methods," *Phys. Rev. D* **40** (Nov, 1989) 3263–3279.

[88] D. S. Goldwirth and T. Piran, "Initial conditions for inflation," *Physics Reports* **214** (1992), no. 4 223–292.

[89] K. Clough, E. A. Lim, B. S. DiNunno, W. Fischler, R. Flauger, and S. Paban, "Robustness of Inflation to Inhomogeneous Initial Conditions," *JCAP* **1709** (2017), no. 09 025, 1608.04408.

[90] A. D. Linde, "Chaotic inflation," *Physics Letters B* **129** (sep, 1983) 177–181.

[91] A. D. Linde, "Inflationary Cosmology," *Lect. Notes Phys.* **738** (2008) 1–54, 0705.0164.

[92] K. Dimopoulos, "The quantum origin of cosmic structure: theory and observations," *J. Phys. Conf. Ser.* **283** (2011) 012010, 1009.5466.

[93] C. Armendariz-Picon, T. Damour, and V. F. Mukhanov, "k - inflation," *Phys. Lett.* **B458** (1999) 209–218, hep-th/9904075.

[94] A. D. Linde, "Hybrid inflation," *Phys. Rev.* **D49** (1994) 748–754, `astro-ph/9307002`.

[95] S. Tsujikawa, "Introductory review of cosmic inflation," in *2nd Tah Poe School on Cosmology: Modern Cosmology Phitsanulok, Thailand, April 17-25, 2003*, 2003. `hep-ph/0304257`.

[96] J. L. Cervantes-Cota and G. Smoot, "Cosmology today-A brief review," *AIP Conf. Proc.* **1396** (2011) 28–52, `1107.1789`.

[97] B. A. Bassett, S. Tsujikawa, and D. Wands, "Inflation dynamics and reheating," *Rev. Mod. Phys.* **78** (2006) 537–589, `astro-ph/0507632`.

[98] Kofman, Lev and Linde, Andrei D. and Starobinsky, Alexei A., "Towards the theory of reheating after inflation," *Phys. Rev.* **D56** (1997) 3258–3295, `hep-ph/9704452`.

[99] A. D. Linde, "Particle physics and inflationary cosmology," *Contemp. Concepts Phys.* **5** (1990) 1–362, `hep-th/0503203`.

[100] R. Allahverdi, R. Brandenberger, F.-Y. Cyr-Racine, and A. Mazumdar, "Reheating in Inflationary Cosmology: Theory and Applications," *Ann. Rev. Nucl. Part. Sci.* **60** (2010) 27–51, `1001.2600`.

[101] L. F. Abbott, E. Farhi, and M. B. Wise, "Particle Production in the New Inflationary Cosmology," *Phys. Lett.* **117B** (1982) 29.

[102] A. D. Dolgov and A. D. Linde, "Baryon Asymmetry in Inflationary Universe," *Phys. Lett.* **116B** (1982) 329.

[103] A. Albrecht, P. J. Steinhardt, M. S. Turner, and F. Wilczek, "Reheating an Inflationary Universe," *Phys. Rev. Lett.* **48** (1982) 1437.

[104] J. Braden, L. Kofman, and N. Barnaby, "Reheating the Universe After Multi-Field Inflation," *JCAP* **1007** (2010) 016, `1005.2196`.

[105] S. Davidson, E. Nardi, and Y. Nir, "Leptogenesis," *Phys. Rept.* **466** (2008) 105–177, `0802.2962`.

[106] H. Baer, K.-Y. Choi, J. E. Kim, and L. Roszkowski, "Dark matter production in the early Universe: beyond the thermal WIMP paradigm," *Phys. Rept.* **555** (2015) 1–60, `1407.0017`.

[107] K. A. Olive, G. Steigman, and T. P. Walker, "Primordial nucleosynthesis: Theory and observations," *Phys. Rept.* **333** (2000) 389–407, astro-ph/9905320.

[108] J. McDonald, "Reheating temperature and inflaton mass bounds from thermalization after inflation," *Phys. Rev.* **D61** (2000) 083513, hep-ph/9909467.

[109] M. Yu. Khlopov and A. D. Linde, "Is It Easy to Save the Gravitino?," *Phys. Lett.* **138B** (1984) 265–268.

[110] J. R. Ellis, J. E. Kim, and D. V. Nanopoulos, "Cosmological Gravitino Regeneration and Decay," *Phys. Lett.* **145B** (1984) 181–186.

[111] M. Gleiser and R. O. Ramos, "Microphysical approach to nonequilibrium dynamics of quantum fields," *Phys. Rev.* **D50** (1994) 2441–2455, hep-ph/9311278.

[112] Y.-K. E. Cheung, M. Drewes, J. U. Kang, and J. C. Kim, "Effective Action for Cosmological Scalar Fields at Finite Temperature," *JHEP* **08** (2015) 059, 1504.04444.

[113] Kofman, Lev and Linde, Andrei D. and Starobinsky, Alexei A., "Reheating after inflation," *Phys. Rev. Lett.* **73** (1994) 3195–3198, hep-th/9405187.

[114] V. Mukhanov, *Physical Foundations of Cosmology*. Cambridge University Press, Oxford, 2005.

[115] P. B. Greene and L. Kofman, "Preheating of fermions," *Phys. Lett.* **B448** (1999) 6–12, hep-ph/9807339.

[116] P. B. Greene and L. Kofman, "On the theory of fermionic preheating," *Phys. Rev.* **D62** (2000) 123516, hep-ph/0003018.

[117] J. Berges, D. Gelfand, and J. Pruschke, "Quantum theory of fermion production after inflation," *Phys. Rev. Lett.* **107** (2011) 061301, 1012.4632.

[118] G. F. Giudice, M. Peloso, A. Riotto, and I. Tkachev, "Production of massive fermions at preheating and leptogenesis," *JHEP* **08** (1999) 014, hep-ph/9905242.

[119] I. I. Tkachev, "Inflation," *Nucl. Phys. Proc. Suppl.* **110** (2002) 144–150, `hep-ph/0112136`.

[120] D. Mattingly, "Modern tests of Lorentz invariance," *Living Rev. Rel.* **8** (2005) 5, `gr-qc/0502097`.

[121] M. D. Schwartz, *Quantum field theory and the standard model.* Cambridge University Press, 2014.

[122] W. E. Lamb Jr and R. C. Retherford, "Fine structure of the hydrogen atom by a microwave method," *Physical Review* **72** (1947), no. 3 241.

[123] M. Bordag, "M. bordag, u. mohideen, and vm mostepanenko, phys. rep. 353, 1 (2001).," *Phys. Rep.* **353** (2001) 1.

[124] H. A. Bethe, "The electromagnetic shift of energy levels," *Physical Review* **72** (1947), no. 4 339.

[125] J. Schwinger, "Quantum electrodynamics. ii. vacuum polarization and self-energy," *Physical Review* **75** (1949), no. 4 651.

[126] R. P. Feynman, "Relativistic cut-off for quantum electrodynamics," *Physical Review* **74** (1948), no. 10 1430.

[127] V. Braginskii and V. Panov, "Verification of equivalence of inertial and gravitational masses," *Zh. Eksp. Teor. Fiz.* **61** (1971) 463–466.

[128] S. Weinberg, "The cosmological constant problem," *Reviews of Modern Physics* **61** (1989), no. 1 1.

[129] R. Bousso, "The cosmological constant," *General Relativity and Gravitation* **40** (2008), no. 2 607–637.

[130] J. Martin, "Everything you always wanted to know about the cosmological constant problem (but were afraid to ask)," *Comptes Rendus Physique* **13** (2012), no. 6-7 566–665.

[131] C. Burgess, "The cosmological constant problem: why its hard to get dark energy from microphysics," *Post-Planck Cosmology: Lecture Notes of the Les Houches Summer School: Volume 100, July 2013* **100** (2015) 149.

[132] **ATLAS, CMS** Collaboration, G. Aad *et. al.*, "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments," *Phys. Rev. Lett.* **114** (2015) 191803, `1503.07589`.

[133] M. Holthausen, K. S. Lim, and M. Lindner, "Planck scale Boundary Conditions and the Higgs Mass," *JHEP* **02** (2012) 037, `1112.2415`.

[134] N. Craig, "Naturalness and new approaches to the hierarchy problem,".

[135] G. t Hooft, "Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking," *NATO Sci. Ser. B* **59** (1979), no. PRINT-80-0083 (UTRECHT) 135–157.

[136] J. Polchinski, "Renormalization and effective lagrangians," *Nuclear Physics B* **231** (1984), no. 2 269–295.

[137] C. P. Burgess, "Introduction to Effective Field Theory," *Ann. Rev. Nucl. Part. Sci.* **57** (2007) 329–362, `hep-th/0701053`.

[138] J. D. Lykken, "Introduction to supersymmetry," in *Fields, strings and duality. Proceedings, Summer School, Theoretical Advanced Study Institute in Elementary Particle Physics, TASI'96, Boulder, USA, June 2-28, 1996*, pp. 85–153, 1996. `hep-th/9612114`.

[139] K. A. Olive, P. D. Group, *et. al.*, "Review of particle physics," *Chinese physics C* **38** (2014), no. 9 090001.

[140] M. V. Ostrogradski, *Mémoire sur les équations différentielles relatives au problème des isopérimetres.. Lu le 17 (29) novembre 1848.* l'Académie impériale des sciences, 1850.

[141] R. P. Woodard, "Avoiding dark energy with 1/r modifications of gravity," *Lect. Notes Phys.* **720** (2007) 403–433, `astro-ph/0601672`.

[142] A. V. Smilga, "Ghost-free higher-derivative theory," *Phys. Lett.* **B632** (2006) 433–438, `hep-th/0503213`.

[143] G. Allemandi and M. Francaviglia, "The variational approach to alternative theories of gravity," *International Journal of Geometric Methods in Modern Physics* **4** (2007), no. 01 1–23.

[144] A. Nicolis and R. Rattazzi, "Classical and quantum consistency of the DGP model," *JHEP* **06** (2004) 059, hep-th/0404159.

[145] K. Hinterbichler and J. Khoury, "Symmetron Fields: Screening Long-Range Forces Through Local Symmetry Restoration," *Phys. Rev. Lett.* **104** (2010) 231301, 1001.4525.

[146] M. Pietroni, "Dark energy condensation," *Phys. Rev.* **D72** (2005) 043535, astro-ph/0505615.

[147] K. A. Olive and M. Pospelov, "Environmental dependence of masses and coupling constants," *Phys. Rev.* **D77** (2008) 043524, 0709.3825.

[148] T. Damour and A. M. Polyakov, "The String dilaton and a least coupling principle," *Nucl. Phys.* **B423** (1994) 532–558, hep-th/9401069.

[149] P. Brax, C. van de Bruck, A.-C. Davis, B. Li, and D. J. Shaw, "Nonlinear Structure Formation with the Environmentally Dependent Dilaton," *Phys. Rev.* **D83** (2011) 104026, 1102.3692.

[150] J. Khoury and A. Weltman, "Chameleon fields: Awaiting surprises for tests of gravity in space," *Phys. Rev. Lett.* **93** (2004) 171104, astro-ph/0309300.

[151] J. Khoury and A. Weltman, "Chameleon cosmology," *Phys. Rev.* **D69** (2004) 044026, astro-ph/0309411.

[152] E. Babichev, C. Deffayet, and R. Ziour, "k-Mouflage gravity," *Int. J. Mod. Phys.* **D18** (2009) 2147–2154, 0905.2943.

[153] E. Babichev, C. Deffayet, and G. Esposito-Farese, "Improving relativistic MOND with Galileon k-mouflage," *Phys. Rev.* **D84** (2011) 061502, 1106.2538.

[154] P. Brax, C. Burrage, and A.-C. Davis, "Screening fifth forces in k-essence and DBI models," *JCAP* **1301** (2013) 020, 1209.1293.

[155] C. Burrage and J. Khoury, "Screening of scalar fields in Dirac-Born-Infeld theory," *Phys. Rev.* **D90** (2014), no. 2 024001, 1403.6120.

[156] A. I. Vainshtein, "To the problem of nonvanishing gravitation mass," *Phys. Lett.* **39B** (1972) 393–394.

[157] C. Deffayet, G. R. Dvali, G. Gabadadze, and A. I. Vainshtein, "Nonperturbative continuity in graviton mass versus perturbative discontinuity," *Phys. Rev.* **D65** (2002) 044026, `hep-th/0106001`.

[158] P. Brax, C. Van de Bruck, A. Davis, J. Khoury, and A. Weltman, "Chameleon dark energy," in *AIP conference proceedings*, vol. 736, pp. 105–110, AIP, 2004.

[159] N. Kaloper, "Challenging the cosmological constant," *Physics Letters B* **653** (2007), no. 2 109–115.

[160] J. Wang, L. Hui, and J. Khoury, "No-go theorems for generalized chameleon field theories," *Physical Review Letters* **109** (2012), no. 24 241301.

[161] A. L. Erickcek, N. Barnaby, C. Burrage, and Z. Huang, "Catastrophic Consequences of Kicking the Chameleon," *Phys. Rev. Lett.* **110** (2013) 171101, `1304.0009`.

[162] A. L. Erickcek, N. Barnaby, C. Burrage, and Z. Huang, "Chameleons in the Early Universe: Kicks, Rebounds, and Particle Production," *Phys. Rev.* **D89** (2014), no. 8 084074, `1310.5149`.

[163] A. Padilla, E. Platts, D. Stefanyszyn, A. Walters, A. Weltman, and T. Wilson, "How to Avoid a Swift Kick in the Chameleons," *JCAP* **1603** (2016), no. 03 058, `1511.05761`.

[164] H. van Dam and M. Veltman, "Massive and mass-less yang-mills and gravitational fields," *Nuclear Physics B* **22** (1970), no. 2 397–411.

[165] V. I. Zakharov, "Linearized gravitation theory and the graviton mass," *JETP Lett.* **12** (1970) 312.

[166] A. Nicolis, R. Rattazzi, and E. Trincherini, "The Galileon as a local modification of gravity," *Phys. Rev.* **D79** (2009) 064036, `0811.2197`.

[167] C. Deffayet, G. Esposito-Farese, and A. Vikman, "Covariant galileon," *Physical Review D* **79** (2009), no. 8 084003.

[168] C. Deffayet, O. Pujolas, I. Sawicki, and A. Vikman, "Imperfect dark energy from kinetic gravity braiding," *Journal of Cosmology and Astroparticle Physics* **2010** (2010), no. 10 026.

[169] F. P. Silva and K. Koyama, "Self-accelerating universe in galileon cosmology," *Physical Review D* **80** (2009), no. 12 121301.

[170] T. Kobayashi, M. Yamaguchi, and J. Yokoyama, "G-inflation: Inflation driven by the Galileon field," *Phys. Rev. Lett.* **105** (2010) 231302, `1008.0603`.

[171] E. Babichev and C. Deffayet, "An introduction to the vainshtein mechanism," *Classical and Quantum Gravity* **30** (2013), no. 18 184001.

[172] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis, and R. Rattazzi, "Causality, analyticity and an ir obstruction to uv completion," *Journal of High Energy Physics* **2006** (2006), no. 10 014.

[173] N. Kaloper, A. Padilla, P. M. Saffin, and D. Stefanyszyn, "Unitarity and the vainshtein mechanism," *Physical Review D* **91** (2015), no. 4 045017.

[174] A. De Felice and S. Tsujikawa, "Generalized Galileon cosmology," *Phys. Rev.* **D84** (2011) 124029, `1008.4236`.

[175] T. Chiba, A. De Felice, and S. Tsujikawa, "Cosmological Scaling Solutions for Multiple Scalar Fields," *Phys. Rev.* **D90** (2014), no. 2 023516, `1403.7604`.

[176] G. Gubitosi and E. V. Linder, "Purely Kinetic Coupled Gravity," *Phys. Lett.* **B703** (2011) 113–118, `1106.2815`.

[177] R. de Putter and E. V. Linder, "Kinetic k-essence and Quintessence," *Astropart. Phys.* **28** (2007) 263–272, `0705.0400`.

[178] C. van de Bruck, T. Koivisto, and C. Longden, "Disformally coupled inflation," *JCAP* **1603** (2016), no. 03 006, `1510.01650`.

[179] G. W. Horndeski, "Second-order scalar-tensor field equations in a four-dimensional space," *International Journal of Theoretical Physics* **10** (1974), no. 6 363–384.

[180] C. Deffayet, X. Gao, D. A. Steer, and G. Zahariade, "From k-essence to generalized galileons," *Physical Review D* **84** (2011), no. 6 064039.

[181] C. Charmousis, E. J. Copeland, A. Padilla, and P. M. Saffin, "General second order scalar-tensor theory, self tuning, and the Fab Four," *Phys. Rev. Lett.* **108** (2012) 051101, `1106.2000`.

[182] E. J. Copeland, A. Padilla, and P. M. Saffin, "The cosmology of the Fab-Four," *JCAP* **1212** (2012) 026, `1208.3373`.

[183] E. Babichev, C. Charmousis, D. Langlois, and R. Saito, "Beyond Fab Four," *Class. Quant. Grav.* **32** (2015) 242001, `1507.05942`.

[184] D. Bettoni and S. Liberati, "Disformal invariance of second order scalar-tensor theories: Framing the Horndeski action," *Phys. Rev.* **D88** (2013) 084020, `1306.6724`.

[185] T. Koivisto, D. Wills, and I. Zavala, "Dark D-brane Cosmology," *JCAP* **1406** (2014) 036, `1312.2597`.

[186] C. de Rham and A. J. Tolley, "DBI and the Galileon reunited," *JCAP* **1005** (2010) 015, `1003.5917`.

[187] G. Goon, K. Hinterbichler, A. Joyce, and M. Trodden, "Gauged Galileons From Branes," *Phys. Lett.* **B714** (2012) 115–119, `1201.0015`.

[188] C. de Rham, "Massive Gravity," *Living Rev. Rel.* **17** (2014) 7, `1401.4173`.

[189] J. D. Bekenstein, "The Relation between physical and gravitational geometry," *Phys. Rev.* **D48** (1993) 3641–3647, `gr-qc/9211017`.

[190] J. Sherman and W. J. Morrison, "Adjustment of an inverse matrix corresponding to a change in one element of a given matrix," *The Annals of Mathematical Statistics* **21** (1950), no. 1 124–127.

[191] M. S. Bartlett, "An inverse matrix adjustment arising in discriminant analysis," *The Annals of Mathematical Statistics* **22** (1951), no. 1 107–111.

[192] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, "Healthy theories beyond Horndeski," *Phys. Rev. Lett.* **114** (2015), no. 21 211101, `1404.6495`.

[193] C. Lin, S. Mukohyama, R. Namba, and R. Saitou, "Hamiltonian structure of scalar-tensor theories beyond Horndeski," *JCAP* **1410** (2014), no. 10 071, `1408.0670`.

[194] J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, "Exploring gravitational theories beyond Horndeski," *JCAP* **1502** (2015) 018, `1408.1952`.

[195] X. Gao, "Hamiltonian analysis of spatially covariant gravity," *Phys. Rev.* **D90** (2014) 104033, `1409.6708`.

[196] T. Kobayashi, Y. Watanabe, and D. Yamauchi, "Breaking of Vainshtein screening in scalar-tensor theories beyond Horndeski," *Phys. Rev.* **D91** (2015), no. 6 064013, 1411.4130.

[197] E. Babichev, K. Koyama, D. Langlois, R. Saito, and J. Sakstein, "Relativistic Stars in Beyond Horndeski Theories," *Class. Quant. Grav.* **33** (2016), no. 23 235014, 1606.06627.

[198] D. Langlois and K. Noui, "Degenerate higher derivative theories beyond Horndeski: evading the Ostrogradski instability," *JCAP* **1602** (2016), no. 02 034, 1510.06930.

[199] D. Langlois, "Degenerate Higher-Order Scalar-Tensor (DHOST) theories," in *Proceedings, 52nd Rencontres de Moriond on Gravitation (Moriond Gravitation 2017): La Thuile, Italy, March 25-April 1, 2017*, pp. 221–228, 2017. 1707.03625.

[200] N. Kaloper, A. Padilla, and D. Stefanyszyn, "Sequestering effects on and of vacuum decay," *Phys. Rev.* **D94** (2016), no. 2 025022, 1604.04000.

[201] G. D'Amico, N. Kaloper, A. Padilla, D. Stefanyszyn, A. Westphal, and G. Zahariade, "An tude on global vacuum energy sequester," *JHEP* **09** (2017) 074, 1705.08950.

[202] J. M. Ezquiaga and M. Zumalacrregui, "Dark Energy After GW170817: Dead Ends and the Road Ahead," *Phys. Rev. Lett.* **119** (2017), no. 25 251304, 1710.05901.

[203] D. Langlois, R. Saito, D. Yamauchi, and K. Noui, "Scalar-tensor theories and modified gravity in the wake of GW170817," 1711.07403.

[204] J. Sakstein and B. Jain, "Implications of the Neutron Star Merger GW170817 for Cosmological Scalar-Tensor Theories," *Phys. Rev. Lett.* **119** (2017), no. 25 251303, 1710.05893.

[205] 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), no. 25 251301, 1710.06394.

[206] P. Creminelli and F. Vernizzi, "Dark Energy after GW170817 and GRB170817A," *Phys. Rev. Lett.* **119** (2017), no. 25 251302, 1710.05877.

[207] P. Creminelli, M. Lewandowski, G. Tambalo, and F. Vernizzi, "Gravitational Wave Decay into Dark Energy," `1809.03484`.

[208] N. Narozhnyi and A. Nikishov, "Pair production by aperiod electric field," *Zh. Eksp. Teor. Fiz* **65** (1973) 862–874.

[209] V. Mostepanenko and V. Frolov, "Particle production from vacuum by homogeneous electric field with periodical time dependence," *Yadernaya Fizika* **19** (1974), no. 4 885–896.

[210] A. A. Grib, V. Mostepanenko, and S. Mamayev, "Vacuum quantum effects in strong fields,".

[211] A. Dolgov and D. Kirilova, "Production of particles by a variable scalar field," *Soviet Journal of Nuclear Physics* **51** (1990) 172–177.

[212] J. H. Traschen and R. H. Brandenberger, "Particle production during out-of-equilibrium phase transitions," *Physical Review D* **42** (1990), no. 8 2491.

[213] Amin, Mustafa A. and Hertzberg, Mark P. and Kaiser, David I. and Karouby, Johanna, "Nonperturbative Dynamics Of Reheating After Inflation: A Review," *Int. J. Mod. Phys.* **D24** (2014) 1530003, `1410.3808`.

[214] K. Schmitz and G. Vertongen, "Reheating and preheating after inflation: an introduction," *Research workshop seminar notes, Deutsches Elektronen-Synchrotron (DESY)* (November, 2010).

[215] N. W. McLachlan, "Theory and application of mathieu functions,".

[216] A. Taruya, "Parametric amplification of density perturbation in the oscillating inflation," *Phys. Rev.* **D59** (1999) 103505, `hep-ph/9812342`.

[217] L. Kofman, "Reheating and preheating after inflation," in *Particle cosmology. Proceedings, 3rd RESCEU International Symposium, Tokyo, Japan, November 10-13, 1997*, pp. 1–8, 1997. `hep-ph/9802285`.

[218] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, vol. 55. Courier Corporation, 1964.

[219] S. Y. Khlebnikov and I. I. Tkachev, "Classical decay of the inflaton," *Physical Review Letters* **77** (1996), no. 2 219.

[220] S. Yu. Khlebnikov and I. I. Tkachev, "The Universe after inflation: The Wide resonance case," *Phys. Lett.* **B390** (1997) 80–86, `hep-ph/9608458`.

[221] S. Yu. Khlebnikov and I. I. Tkachev, "Resonant decay of Bose condensates," *Phys. Rev. Lett.* **79** (1997) 1607–1610, `hep-ph/9610477`.

[222] T. Prokopec and T. G. Roos, "Lattice study of classical inflaton decay," *Phys. Rev.* **D55** (1997) 3768–3775, `hep-ph/9610400`.

[223] R. Lerner and A. Tranberg, "Thermal blocking of preheating," *JCAP* **15** (2015), no. 04 014, `1502.01718`.

[224] D. I. Podolsky, G. N. Felder, L. Kofman, and M. Peloso, "Equation of state and beginning of thermalization after preheating," *Phys. Rev.* **D73** (2006) 023501, `hep-ph/0507096`.

[225] D. V. Semikoz, "The kinetic stage of the universe reheating," *Helv. Phys. Acta* **69** (1996) 207–210.

[226] S. Davidson and S. Sarkar, "Thermalization after inflation," *JHEP* **11** (2000) 012, `hep-ph/0009078`.

[227] R. Allahverdi, "Thermalization after inflation and reheating temperature," *Phys. Rev.* **D62** (2000) 063509, `hep-ph/0004035`.

[228] R. Allahverdi and M. Drees, "Thermalization after inflation and production of massive stable particles," *Phys. Rev.* **D66** (2002) 063513, `hep-ph/0205246`.

[229] I. G. Moss and C. M. Graham, "Particle production and reheating in the inflationary universe," *Phys. Rev.* **D78** (2008) 123526, `0810.2039`.

[230] K. Harigaya and K. Mukaida, "Thermalization after/during Reheating," *JHEP* **05** (2014) 006, `1312.3097`.

[231] K. Mukaida and M. Yamada, "Thermalization Process after Inflation and Effective Potential of Scalar Field," *JCAP* **1602** (2016), no. 02 003, `1506.07661`.

[232] D. Boyanovsky, M. D'Attanasio, H. J. de Vega, R. Holman, and D. S. Lee, "Reheating and thermalization: Linear versus nonlinear relaxation," *Phys. Rev.* **D52** (1995) 6805–6827, hep-ph/9507414.

[233] D. T. Son, "Reheating and thermalization in a simple scalar model," *Phys. Rev.* **D54** (1996) 3745–3761, hep-ph/9604340.

[234] S. Yu. Khlebnikov and I. I. Tkachev, "Classical decay of inflaton," *Phys. Rev. Lett.* **77** (1996) 219–222, hep-ph/9603378.

[235] L. Kofman, "Preheating, thermalization and supergravity," in *Proceedings, 3rd International Heidelberg Conference on Dark matter in astro- and particle physics (DARK 2000): Heidelberg, Germany, July 10-14, 2000*, pp. 135–149, 2000. hep-ph/0012298.

[236] R. Micha and I. I. Tkachev, "Relativistic turbulence: A Long way from preheating to equilibrium," *Phys. Rev. Lett.* **90** (2003) 121301, hep-ph/0210202.

[237] R. Micha and I. I. Tkachev, "Preheating and thermalization after inflation," in *Proceedings, 5th Internationa Conference on Strong and Electroweak Matter (SEWM 2002): Heidelberg, Germany, October 2-5, 2002*, pp. 210–219, 2003. hep-ph/0301249.

[238] R. Micha and I. I. Tkachev, "Turbulent thermalization," *Phys. Rev.* **D70** (2004) 043538, hep-ph/0403101.

[239] A. Kaya, "Fluctuations of Quantum Fields in a Classical Background and Reheating," *Phys. Rev.* **D81** (2010) 023521, 0909.2712.

[240] M. Drewes and J. U. Kang, "The Kinematics of Cosmic Reheating," *Nucl. Phys.* **B875** (2013) 315–350, 1305.0267. [Erratum: Nucl. Phys.B888,284(2014)].

[241] M. Drewes, "On finite density effects on cosmic reheating and moduli decay and implications for Dark Matter production," *JCAP* **1411** (2014), no. 11 020, 1406.6243.

[242] G. Sigl and G. Raffelt, "General kinetic description of relativistic mixed neutrinos," *Nucl. Phys.* **B406** (1993) 423–451.

[243] M. Herranen, K. Kainulainen, and P. M. Rahkila, “Towards a kinetic theory for fermions with quantum coherence,” *Nucl. Phys.* **B810** (2009) 389–426, `0807.1415`.

[244] M. Herranen, K. Kainulainen, and P. M. Rahkila, “Quantum kinetic theory for fermions in temporally varying backgrounds,” *JHEP* **09** (2008) 032, `0807.1435`.

[245] M. Herranen, K. Kainulainen, and P. M. Rahkila, “Kinetic theory for scalar fields with nonlocal quantum coherence,” *JHEP* **05** (2009) 119, `0812.4029`.

[246] G. Baym and L. P. Kadanoff, “Conservation Laws and Correlation Functions,” *Phys. Rev.* **124** (1961) 287–299.

[247] J. S. Schwinger, “Brownian motion of a quantum oscillator,” *J. Math. Phys.* **2** (1961) 407–432.

[248] L. V. Keldysh, “Diagram technique for nonequilibrium processes,” *Zh. Eksp. Teor. Fiz.* **47** (1964) 1515–1527. [Sov. Phys. JETP20,1018(1965)].

[249] J.-P. Blaizot and E. Iancu, “The Quark gluon plasma: Collective dynamics and hard thermal loops,” *Phys. Rept.* **359** (2002) 355–528, `hep-ph/0101103`.

[250] J. Berges, “Introduction to nonequilibrium quantum field theory,” *AIP Conf. Proc.* **739** (2005) 3–62, `hep-ph/0409233`. [,3(2004)].

[251] B. Garbrecht, T. Prokopec, and M. G. Schmidt, “Particle number in kinetic theory,” *Eur. Phys. J.* **C38** (2004) 135–143, `hep-th/0211219`.

[252] M. Herranen, K. Kainulainen, and P. M. Rahkila, “Coherent quantum Boltzmann equations from cQPA,” *JHEP* **12** (2010) 072, `1006.1929`.

[253] M. Herranen, K. Kainulainen, and P. M. Rahkila, “Flavour-coherent propagators and Feynman rules: Covariant cQPA formulation,” *JHEP* **02** (2012) 080, `1108.2371`.

[254] C. Fidler, M. Herranen, K. Kainulainen, and P. M. Rahkila, “Flavoured quantum Boltzmann equations from cQPA,” *JHEP* **02** (2012) 065, `1108.2309`.

[255] P. Millington and A. Pilaftsis, “Perturbative nonequilibrium thermal field theory,” *Phys. Rev.* **D88** (2013), no. 8 085009, `1211.3152`.

[256] P. S. Bhupal Dev, P. Millington, A. Pilaftsis, and D. Teresi, "Flavour covariant transport equations: an application to resonant leptogenesis," *Nucl. Phys.* **B886** (2014) 569–664, 1404.1003.

[257] C. Volpe, D. Vnnen, and C. Espinoza, "Extended evolution equations for neutrino propagation in astrophysical and cosmological environments," *Phys. Rev.* **D87** (2013), no. 11 113010, 1302.2374.

[258] D. Vnnen and C. Volpe, "Linearizing neutrino evolution equations including neutrino-antineutrino pairing correlations," *Phys. Rev.* **D88** (2013) 065003, 1306.6372.

[259] A. Vlasenko, G. M. Fuller, and V. Cirigliano, "Neutrino Quantum Kinetics," *Phys. Rev.* **D89** (2014), no. 10 105004, 1309.2628.

[260] J. Serreau and C. Volpe, "Neutrino-antineutrino correlations in dense anisotropic media," *Phys. Rev.* **D90** (2014), no. 12 125040, 1409.3591.

[261] A. Kartavtsev, G. Raffelt, and H. Vogel, "Neutrino propagation in media: Flavor-, helicity-, and pair correlations," *Phys. Rev.* **D91** (2015), no. 12 125020, 1504.03230.

[262] M. Morikawa and M. Sasaki, "Entropy Production in the Inflationary Universe," *Prog. Theor. Phys.* **72** (1984) 782.

[263] **BICEP2, Planck** Collaboration, P. A. R. Ade *et. al.*, "Joint Analysis of BICEP2/Keck Array and Planck Data," *Phys. Rev. Lett.* **114** (2015) 101301, 1502.00612.

[264] L. A. Kofman and A. D. Linde, "Generation of Density Perturbations in the Inflationary Cosmology," *Nucl. Phys.* **B282** (1987) 555.

[265] H. Goldstein, C. Poole, and J. Safko, *Classical mechanics*. AAPT, 2002.

[266] V. F. Weisskopf and E. P. Wigner, "Calculation of the natural brightness of spectral lines on the basis of dirac's theory," *Z. Phys.* **63** (1930) 54–73.

[267] N. Bogoljubov, V. V. Tolmachov, and D. Širkov, "A new method in the theory of superconductivity," *Fortschritte der physik* **6** (1958), no. 11-12 605–682.

[268] R. L. Kobes and G. W. Semenoff, "Discontinuities of Green Functions in Field Theory at Finite Temperature and Density," *Nucl. Phys.* **B260** (1985) 714–746.

[269] R. L. Kobes and G. W. Semenoff, "Discontinuities of Green Functions in Field Theory at Finite Temperature and Density. 2," *Nucl. Phys.* **B272** (1986) 329–364.

[270] R. E. Cutkosky, "Singularities and discontinuities of Feynman amplitudes," *J. Math. Phys.* **1** (1960) 429–433.

[271] G. 't Hooft and M. J. G. Veltman, "DIAGRAMMAR," *NATO Sci. Ser. B* **4** (1974) 177–322.