university of groningen

faculty of science and engineering

Primordial Non-Gaussianity in the Single- and Multi-Field Inflationary Scenarios

Bachelor Thesis

in

Theoretical Physics

Author: T.W.J de Wild

Supervisors: Prof. Dr. D. Roest

Prof. Dr. A. Mazumdar

MSc. P. Christodoulidis

July 5, 2018

Dedicated to my dad, Harry de Wild

Table of Contents

Introduction

> *"Cosmology is among the oldest subjects to captivate our species. And that's no wonder. We're storytellers, and what could be more grand than the story of creation?"*
>
> — Brian Greene

Currently, the inflationary paradigm is one of the most promising candidates for early universe physics [2, 3, 14, 15, 47, 60]. In essence, the theory of inflation conjectures that the universe underwent a period of exponential expansion shortly after the Big Bang, before it transits into the radiation and matter dominated era's which are described by conventional Big Bang theory. In this thesis, we aim to constrain different inflationary scenarios based on precision observations of the cosmic microwave background (CMB). However, before going into details, the relevance and context of the research outlined in this thesis will be further motivated and established.

Historically, a period of cosmic inflation was conjectured in order to account for a number of observed properties of the universe, which could not be explained in the framework of conventional Big Bang theory. In particular, conventional Big Bang theory ceases to explain two observational facts. First of all, the universe is observed to geometrically flat today [49]. In other words, the geometry of the universe possesses no intrinsic curvature. The issue with this observation is that, in order to obtain the observed flatness of the universe today, the geometry of the universe must have been extremely flat at the earliest stages after the Big Bang. Conventional Big Bang theory does not *a priori* provide an explanation why the universe should start out in such a flat initial state, thereby giving rise to a fine-tuning problem [14, 15, 47, 60].

Secondly, conventional Big Bang theory fails to account for the observed uniformity of the CMB. In all directions along the sky, the CMB radiation is observed to have the same temperature to a very high degree. This uniformity of temperature requires that different patches in the CMB have been in causal contact, as otherwise a thermal equilibrium resulting in the observed uniform temperature cannot be established. However, conventional Big Bang theory predicts that CMB patches on the sky separated by more than two degrees have never been in causal contact. Therefore, at least within the framework of conventional Big Bang theory, the observed uniformity of the CMB cannot be explained. This mystery is known as the horizon problem [14, 15, 60].

Inflation dynamically solves these Big Bang puzzles in an intuitive way [47, 60]. Even if the universe starts out in a state which is not at all flat, the inflationary expansion will drive the universe to a flat state. Therefore, the observed flatness becomes a logical dynamical consequence, rather than a puzzle. Similarly, inflation solves the horizon problem as the seemingly causally disconnected patches of the CMB have been in causal contact at the beginning of inflation. Therefore, in the inflationary paradigm, the observed uniformity of the CMB becomes a prediction, instead of an unexplained feature of the universe.

In addition to solving the puzzles of conventional Big Bang theory, it was later discovered that inflation also naturally provides the primordial seeds required for structure formation [25, 60, 65, 69, 75]. During the inflationary stage, quantum fluctuations are rapidly stretched to cosmological scales, freeze-in and become classical. Those quantum fluctuations thus transform into classical perturbations. In the radiation dominated era, those perturbations originated during inflation manifest as perturbations in the energy density throughout the universe. Subsequently, the density inhomogeneities serve as the primordial seeds out of which all observed large-scale structure is formed according to the mechanism of gravitational instability. Furthermore, the quantum fluctuations during inflation induce minute, but observable, temperature anisotropies in the CMB.

Although inflation serves as an extremely powerful paradigm for explaining several features of the observable universe, the microscopic mechanism behind inflation is still to be revealed [17, 60]. Usually, inflation is modeled using scalar fields characterized by a flat potential, on which the fields slowly roll down to the minimum. The flatness of the potential is required since otherwise exponential expansion of space-time cannot be established. Furthermore, the flatness of the potential puts strong constraints on the type of interactions between the scalar fields and possibly other fields present during inflation. Constrained by the flatness of the potential, those interactions are small and can to good approximation be treated as Gaussian fluctuations around a free field theory. Therefore, the fluctuations generated during inflation, and subsequently present in the CMB as temperature anisotropies, are predicted to be nearly Gaussian distributed.

Roughly, we can discriminate between two classes of inflation models: single- and multi-field models of inflation. On the one hand, single field models of inflation, also referred to as the *vanilla* scenario of inflation, are simple and intuitive. On the other hand, multi-field models of inflation are well motivated within the context of candidate theories for high energy physics, such as string theory [4, 15, 17].

Up till now, the two classes of models are both still in agreement with the current cosmological data. In particular, the quantum fluctuations during inflation and temperature fluctuations in the CMB are predicted to be small, nearly Gaussian and adiabatic. More formally, the so-called power spectrum of the fluctuations is expected to be featureless, almost scale-invariant and its residual scale dependence is predicted to be described by a power-law. Those generic predictions hold for most of the single- and multi-field models and are verified by CMB observations, for instance, made by the Planck satellite [3].

From one side, this observational verification may be regarded as a huge success of inflation. However, based on those observational results we cannot discriminate between different microscopic scenarios that are proposed to underlie the inflationary phase.[1] Therefore, we have to search for other observational differences between competing inflationary scenarios. In this thesis, we aim to derive observational differences between the single- and multi-field scenarios of inflation imprinted in the CMB.

In particular, we will focus on the (non-)gaussianity feature of the temperature fluctuations in the CMB. As mentioned above, due to the flatness of the potential, the quantum fluctuations generated during inflation are expected to be nearly Gaussian, as well as the temperature anisotropies in the CMB. However, some level of non-gaussianity is expected to be present, due to the necessary coupling of the fields to the gravitational field. In addition, the non-gaussian signal can be enhanced depending on the specific details of the inflationary

[1] In more technical terms, which will be introduced properly in this thesis, different models of inflation are often degenerate in terms of the spectral index and tensor-to-scalar ratio.

model. To this end, an important discovery was made over the last decade [2]: different classes of inflationary models would produce different non-gaussianity signals. Those non-gaussianity signals would also be imprinted in the CMB anisotropies. Hence, by actually measuring the type of non-gaussianity signal contained in the CMB, classes of inflationary models predicting a different signal can potentially be ruled out.

Summarizing, in this thesis we examine how non-gaussianity contained in the CMB can be exploited as an observational window to constrain and possibly even rule out (classes of) inflationary models. In particular, we will aim to derive the predicted level of non-gaussianity in the case of single- and multi-field inflation. Ultimately, those predictions can be compared with (future) observations on the non-gaussianity contained in the CMB to potentially rule out the single- or multi-field scenario in favor of the other.

Outline of this Thesis

In order to introduce different aspects discussed in this thesis in a clear and structured manner, we have divided this thesis into five main parts. In part I, we will extensively review the key aspects of conventional Big Bang. In particular, we will discuss the flatness and horizon puzzles in great detail. Subsequently, we will introduce the classical mechanism of inflation as a possible solution to these problems. In part II, we will leave the classical regime of the inflationary paradigm and introduce, at a qualitative level, the quantum effects responsible for the primordial seeds of large-scale structure formation and the CMB anisotropies. Furthermore, we will introduce the formalism, called cosmological perturbation theory, that we will use to study the quantum fluctuations around the homogeneous background universe. In part III, we will study the quantum effects during (single-field) inflation in great detail. In particular, we will derive the power spectrum of the quantum fluctuations and compare with CMB observations. Furthermore, we will introduce the general concepts related to non-gaussianity. Finally, in part IV and V, the level of non-gaussianity is predicted for the single- and multi-field inflationary scenarios.

Different Routes Through the Thesis

Depending on the reader's interest, this thesis can be approached via different routes. For a reader new to the field of inflationary cosmology, the first three parts provide an extensive introduction to both the classical and quantum aspects of the theory. Readers solely interested in the quantum effects of inflation can omit the first part. Finally, readers focussing on the calculation of the non-gaussianity level in single- and multi-field inflation, the first three parts can be skipped.

It is worthwhile to mention that most of the material in the first three parts of this thesis can also be found in well-written lecture notes and books on inflation, see e.g. [14, 15, 60, 74]. Nevertheless, those parts are included in order to make this work as complete and self-contained as possible. This is not the case for part IV and V, which are merely based on actual research papers (e.g. [64, 76, 77]) and complement them by providing detailed derivations of the results stated in those papers. In that sense, the last two parts serve as the core of this work, since they contribute significantly to the existing literature.

Acknowledgements

Although this thesis is presented as the work of one individual, I have, and am willing to, acknowledge that it is instead a collaborative effort to which many people have contributed in different ways. I found myself in the wonderful and fruitful position to learn from my various different people, including my supervisors, friends and family. In particular, I am grateful for the support I received from the following people.

My Supervisors.—First of all, I would like to thank my supervisors, D. Roest, A. Mazumdar and P. Christodoulidis. Diederik, I am grateful for the opportunity you gave me to enter the field of theoretical physics and the possibility to perform this research project at the Van Swinderen Institute. Perseas, thank you for supervising on a daily basis during this project. Finally, I would like to thank Anupam for being available as the second supervisor and attending my final presentation.

My High-School Science Teachers.—Without the aid of my high school science teachers, A. Brouwer and A. Linthorst, I would probably not have chosen to study physics in the first place, thank you. Arjen and Albert, I've learned a lot of fundamental lessons about science from you. Thank you for constantly triggering my curiosity and giving me the freedom to figure out things myself. Without your patience and enthusiasm in teaching the principles of physics and chemistry, I would never have come this far in my educative journey in the natural sciences.

My Family.—Dear Corinne and Stijn, I would like to thank you both for the support you provided me with during this project. In particular, I would like to thank you for showing me the rest of the world during times I got stuck in minus signs, and factors of 2 and π. You helped me to put difficulties encountered during this project into the right perspective. Dear Harry, I remember playing with wooden blocks with you as a small child during times you worked in the garage. In this way, you triggered my curiosity for the way the world around me works, already in the earliest stages. Although you will possible never know, you were my first science teacher. I love you all!

My Friends.—I am grateful for being surrounded by lovely friends, in particular Floris, Sytze, Lena, Lotte, Marit. It is hard to imagine friends that are more supportive and caring than you guys. Floris, I'd like to thank you for your emotional and technical support during this project: at times the Latex errors took control over my frustrations you were there to help me out. Sytze, I guess it's safe to state that I never encountered someone that takes life as positive as you do. In that respect, I have learned a lot from you and I still do, thank you. Lena, I want to thank you for your friendship. It does not happen often in life that you

meet someone that shares the same kind of humor, music preferences and (most importantly) eating habits. Furthermore, I am thankful for your understanding and patient support at the times my humor disappeared. Lotte, thank you for the unlimited kindness you share with the people around you. You inspire me.

Shannah.—Although you don't like to be in spotlights, I guess it is justified to shed some light on the kind and patient role you have taken in this project. First of all, thanks for your confidence in me during times I lost confidence in myself. Secondly, I would like to use this opportunity to apologize for the busy times during this project at which it may have appeared that you where standing in second place. From now on, I'll make sure this changes. Finally and most importantly, I would sincerely like to thank you for being the person you are. I love you!

Notation and Conventions

"Notation is a complete nightmare."
— Sytze Tirion

Throughout this work, the author aimed to be as consistent as possible in terms of notation and conventions. Below, the main notations and conventions are introduced point-wise.

▷ *Natural Units.*—The reduced Planck constant $\hbar = h/2\pi$ and the speed of light c are set equal to one, that is:

$$\hbar = c \equiv 1.$$

▷ *Reduced Planck Mass.*—The fundamental constants of nature are conventionally combined into the so-called reduced Planck mass:

$$M_{\rm pl} \equiv \sqrt{\frac{\hbar c}{8\pi G}} = 2.435 \times 10^{18}\,\text{GeV},$$

where G is the Newtonian gravitational constant. In natural units, the reduced Planck mass becomes:

$$M_{\rm pl} = (8\pi G)^{-1/2}.$$

▷ *Indices.*—Space-time indices are given by Greek letters, e.g. μ, ν, spatial indices are given by roman letters, e.g. i, j. Since 4-dimensional space-time is assumed throughout, latin indices run from 0 to 3 and roman indices run from 1 to 3. That is:

$$x^\mu = (x^0, x^1, x^2, x^3) = (t, \boldsymbol{x}), \qquad x^i = (x^1, x^2, x^3) = \boldsymbol{x},$$

where the $\mu = 0$ component corresponds to the temporal coordinate: $x^0 = t$. In a cartesian coordinate system, the spatial components are $x^1 = x$, $x^2 = y$ and $x^3 = z$.

▷ *Metric Signature.*—The metric signature in this work is $(-+++)$, such that the invariant differential line element ds^2 for a flat Minkowski spacetime becomes:

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = -dt^2 + d\boldsymbol{x}^2.$$

In accordance with the chosen metric signature, the Minkowski metric reads $\eta_{\mu\nu} = \text{diag}(-1, +1, +1, +1)$.

▷ *Vectors.*—Four-vectors are usually represented using index notation, e.g. p^μ denotes the energy-momentum 4-vector. Physical 3-vectors are represented with bold letters or in index notation, e.g. $\boldsymbol{x}, x^i$ and $\boldsymbol{p}, p^i$ for the position and 3-momentum vector, respectively. Boldface letters and index notation will be used interchangeably. The magnitude of a vector is given in unbolded notation, e.g. $x \equiv \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}}$. Lastly, vectors can be written in terms of their magnitude and the corresponding unit vector. For instance, the 3-momentum can be written as $\boldsymbol{p} \equiv p\hat{\boldsymbol{p}}$, where $\hat{\boldsymbol{p}}$ denotes the unit vector.

▷ *Spatial Derivatives and Gradients.*—In the literature, different definitions of the Laplacian are used, resulting form different conventions on contracting (spatial) indices with the Kronecker delta function or the spatial metric. In chapters 1-9, we take the convention to contract indices using the Kronecker function, such that the Laplacian is defined as:

$$\partial^2 = \partial_i \partial^i = \delta^{ij} \partial_i \partial_j.$$

From chapter 10 and onwards, it proves convenient to adopt the convention that spatial indices are solely contracted using the spatial part of the metric g_{ij}, which for a flat background FRW universe is defined as: $g_{ij} = a^2 \delta_{ij}$. The Laplacian is then defined as:

$$\nabla^2 \equiv \partial_i \partial^i = g^{ij} \partial_i \partial_j = \frac{\partial^2}{a^2}.$$

In Fourier space, the operator ∂^2 is replaced by $-k^2$, so that we have the prescription $\nabla^2 \to -k^2/a^2$. To make the difference between $\partial^2 = \delta^{ij} \partial_i \partial_j$ and $\nabla^2 \equiv \partial_i \partial^i$ clear from chapter 10 and onwards, we define ∂^2 in those chapters as:

$$\partial^2 \equiv \partial_i \partial_i \equiv \delta^{ij} \partial_i \partial_j,$$

emphasizing that the indices i and j are summed over in the expression for ∂^2 using the Kronecker delta δ^{ij}, but *not* contracted, which is to be done using the metric. It should be noted that our notation is different from for instance Riotto [74], which defines ∇^2 as our ∂^2. In practice, to compare our expressions to literature, one has to pay attention to the factors of a^2.

▷ *Einstein Summation Convention.*—Repeated indices are summed over, that is:

$$A_\mu B_\mu \equiv \sum_{\mu=0}^{3} A_\mu B_\mu = A_0 B_0 + A_1 B_1 + A_2 B_2 + A_3 B_3.$$

The summation convention appears frequently for the Kronecker delta function δ^m_n, which is zero for $m \neq n$ and one for $m = n$. For the upper index identical to the lower index, the summation convention yields:

$$\delta^i_i = \delta^1_1 + \delta^2_2 + \delta^3_3 = 3.$$

Lastly, note that using the summation convention and index notation, the dot product of the position vector $\boldsymbol{x}$ can be expressed as:

$$\boldsymbol{x} \cdot \boldsymbol{x} = x^i x_i = x_1^2 + x_2^2 + x_3^2.$$

▷ *Coordinate, Proper and Conformal Time.*—Coordinate or cosmic time will be denoted as t. and proper time as η. In contrast to coordinate time, we will often use conformal time as the evolution variable. The conformal time differential $d\tau$ is related to the coordinate time differential dt as:

$$d\tau = dt/a(t), \tag{0.0.1}$$

where $a(t)$ is the scale factor. Conformal time is useful as it factorizes the FRW metric of an expanding universe into a static (Minkowski) component and a single function of time (the scale factor).

▷ *Derivation Boxes.*—Technical derivations of various results are performed in so-called derivations boxes, indicated by frames such as the one around this section.

$$\boxed{\text{Boxes contain supplementary material or derivations.}}$$

Omitting the content in these boxes while reading will not cause any obstruction in following the main message or common thread.

▷ *Fourier Convention.*—The Fourier convention in this work is:

$$\mathcal{R}(\boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \mathcal{R}_{\boldsymbol{k}}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \qquad \mathcal{R}_{\boldsymbol{k}} = \int d^3\boldsymbol{x}\, \mathcal{R}(\boldsymbol{x})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}},$$

where $\boldsymbol{k}$ is the wavevector and $k = |\boldsymbol{k}|$ is its magnitude, which is also known as the wavenumber. Sometimes the bold notation $d^3\boldsymbol{k}$ is omitted and d^3k is written instead. The volume element d^3k (in Fourier space) is defined as:

$$d^3k = k^2 \sin\theta\, d\theta\, d\phi\, dk$$

In cases, where the direction of $\boldsymbol{k}$ is relevant, the integral is often separated into a radial integral over the magntiude k and an angular integral over the direction of the wavevector $\hat{\boldsymbol{k}}$ as:

$$\int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \to \int \frac{dk}{k} \frac{k^3}{2\pi^2} \int \frac{d^2\hat{\boldsymbol{k}}}{4\pi}$$

▷ *Dirac Delta Functions.*—Using the Fourier convention stated above, the 3-dimensional Dirac delta functions in real and momentum space are given by:

$$\delta^{(3)}(\boldsymbol{k}) = \int d^3\boldsymbol{x}\, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}, \qquad \delta^{(3)}(\boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}.$$

▷ *Partial Derivatives in Fourier Space.*—Often, we will switch between the original equations and their Fourier-equivalents. To do this, typically two actions should be performed: (a) the original variable $\mathcal{Q}$ is replaced by $\mathcal{Q}_{\boldsymbol{k}}$ and (b) possible spatial partial derivatives ∂_j are replaced by ik_j, where i is the imaginary unit.

$$\partial_j \mathcal{Q}(\boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \mathcal{Q}_{\boldsymbol{k}}\, \partial_j e^{ik_j x^j} = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} (ik_j \mathcal{Q}_{\boldsymbol{k}})\, e^{ik_j x^j}.$$

▷ *Power Spectrum.*—In accordance with the obeyed Fourier convention, we define the power spectrum $\mathcal{P}_{\mathcal{Q}}$ of a generic field $\mathcal{Q}$ as follows:

$$\mathcal{P}_{\mathcal{Q}} \equiv \frac{k^3}{2\pi^2} |\mathcal{Q}_{\boldsymbol{k}}|^2,$$

where $\mathcal{Q}_{\boldsymbol{k}}$ is the considered Fourier mode in the expansion of the field.

▷ *Perturbed Variables.*—A generic variable $\mathcal{Q}$ perturbed to first order is written as:

$$\mathcal{Q} = \overline{\mathcal{Q}} + \delta\mathcal{Q},$$

where $\overline{\mathcal{Q}}$ and $\delta\mathcal{Q}$ denote the zeroth order (background) quantity and the perturbation, respectively. The bar over the background variable will be omitted whenever the difference between the background variable and the perturbation is evident. The perturbation can always be recognized by means of the δ-notation.

Below, we will list the most commonly used symbols in this thesis. In case multiple meanings are assigned to the same symbol, its meaning should be evident from the context.

Symbol	Description
$\hbar$	Planck's constant (set to unity)
c	Speed of light (set to unity)
k_{B}	Boltzmann's constant
G	Gravitational constant
M_{pl}	Planck mass in natural units $M_{\mathrm{pl}}^2 \equiv 1/8\pi G$
t, η, τ	Cosmic, conformal and proper time, $d\tau \equiv dt/a$
$\dot{A}$	Cosmic time derivative of A, $\dot{A} \equiv \partial_t A$
A'	Conformal time derivative of A, $A' \equiv \partial_\tau A$
ds	Invariant space-time line element
x^μ	Space-time coordinate
$\boldsymbol{x}, x^i$	Spatial coordinate
$\hat{\boldsymbol{x}}$	Spatial unit vector
$\boldsymbol{n}$	Directional unit vector
$\boldsymbol{k}, k^i$	Momentum 3-vector
$\hat{\boldsymbol{k}}$	Momentum unit vector
k	Momentum magnitude, $k \equiv \lvert\boldsymbol{k}\rvert$
$\boldsymbol{K}$	Total vector momentum, $\boldsymbol{K} \equiv \boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3$
K	Total scalar momentum, $K \equiv k_1 + k_2 + k_3$
k_{123}	Product of scalar momenta, $k_{123} \equiv k_1 k_2 k_3$
a	Scale factor
H	Hubble parameter, $H \equiv \dot{a}/a$
H_0	Present day value of H
H_*	Value of H during inflation
$\mathcal{H}$	Conformal Hubble parameter, $\mathcal{H} \equiv aH$
$r_{\mathrm{EH}}, r_{\mathrm{PH}}$	Event and particle horizon
$g_{\mu\nu}$	Space-time metric tensor
g	Determinant metric tensor, $g \equiv \det g_{\mu\nu}$
$\Gamma^\rho_{\mu\nu}$	Christoffel symbol
$G_{\mu\nu}$	Einstein tensor
$R^\sigma_{\mu\nu\rho}$	Riemann curvature tensor
$R_{\mu\nu}, R$	Ricci tensor and scalar
γ_{ij}	Spatial 3-metric
$T_{\mu\nu}$	Energy-momentum tensor
Σ_{ij}	Anisotropic stress tensor
∂_μ	Partial space-time derivative
∇_μ	Covariant space-time derivative
$\Box A$	D'Alembertian operator acting on A
∇^2	Laplacian for FRW metric, $\nabla^2 \equiv g^{ij}\partial_i\partial j = \partial^2/a^2$
∂^2	Differential operator, $\partial^2 \equiv \delta^{ij}\partial_i\partial_j$
$\mathcal{L}_b A_{\mu\nu},\ \Delta_b A_{\mu\nu}$	Lie derivatives of field $A_{\mu\nu}$ w.r.t. b^ρ, $\mathcal{L}_b \equiv -\Delta_b$.
U^μ	Four velocity, $U^\mu \equiv dx^\mu/d\eta$
$\boldsymbol{v}, v^i$	Velocity

Symbol	Description
v	Magnitude of velocity, $v^2 \equiv g^{ij} v_i v_j$
γ	Lorentz or gamma factor, $\gamma(v) \equiv 1/\sqrt{1-v^2}$
ρ_i, P_i	Energy density and pressure of constituent i
w	Equation of state, $w \equiv P/\rho$
c_s	Speed of sound, $c_s^2 \equiv \dot{P}/\dot{\rho}$
Ω_i	Density parameter constituent i, $\Omega_i \equiv \rho_i/\rho$
K	Curvature parameter
Λ	Cosmological constant
ε	First Hubble parameter, $\varepsilon \equiv -\dot{H}/H^2$
η	Second Hubble parameter, $\eta \equiv \dot{\varepsilon}/H\varepsilon$
ϕ	Inflaton scalar field
χ	Additional scalar field (e.g. spectator)
$V(\chi)$	Potential function of inflaton field
$V_\phi, V_{\phi\phi}$	Frist and second field derivatives of $V(\phi)$
m_A	Mass of field A
$\varepsilon_{\rm v}$	Potential slow-roll parameter, $\varepsilon_{\rm v} \equiv M_{\rm pl}^2/2(V_\phi/V)^2$
$\eta_{\rm v}$	Second potential slow-roll, $\eta_{\rm v} \equiv M_{\rm pl}^2 V_{\phi\phi}/V$
δ	Dimensionless ratio of $\ddot{\phi}$ and $\dot{\phi}$, $\delta \equiv -\ddot{\phi}/H\dot{\phi}$
X	Random variable
$\rho(x)$	Probability density function
$\mathrm{E}[X]$	Expectation value, X
$\mathrm{Var}[X]$	Variance of X, related to σ as $\sigma^2 \equiv \mathrm{Var}[X]$
ξ_{ij}	Two-point correlation function of i and j
$\langle A(\boldsymbol{x})A(\boldsymbol{y})\rangle$	Two-point correlation function of field A
$\langle A(\boldsymbol{x})A(\boldsymbol{y})A(\boldsymbol{z})\rangle$	Three-point correlation function of field A
$\bar{A}$	Background value of A
δA	First order perturbation of A
δT	Temperature fluctuation field CMB
$\delta\phi$	Fluctuation in inflaton field
δ_i	Fractional density perturbation, $\delta_i \equiv \delta\rho_i/\rho$
$\mathcal{R}$	Comoving curvature perturbation
ζ	Curvature perturbation on uniform slices
$\mathcal{S}$	Isocurvature of entropy perturbation
$\delta\Sigma_{ij}$	Anisotropic stress perturbation
Φ	Gravitational lapse
Ψ	Grivatational potential or curvature perturbation
B_i	Shift vector
E_{ij}	Shear tensor
$\hat{A}$	Quantum operator of A, hat often omitted
$\langle\hat{A}\rangle$	Expectation value of operator $\hat{A}$
$[\hat{A}, \hat{B}]$	Commutator of operators $\hat{A}$ and $\hat{B}$
$\hat{a}_{\boldsymbol{k}}^\dagger, \hat{a}_{\boldsymbol{k}}$	Creation and annihilation operators
$W[f_k, f_k^*]$	Wroskian of mode function f_k
f	Rescaled quantum fluctuation $f \equiv a\delta\phi$

Symbol	Description
z	Auxilary variable defined as $z \equiv a\dot{\phi}/H$
f_k	Mode function of f in Fourier space
$H_\nu^{(1,2)}$	First and second Hankel functions
$\mathcal{P}_A$	Power spectrum of A
$n_{\rm s}$	Scalar spectral index
$n_{\rm t}$	Tensor spectral index
r	Tensor to scalar ratio, $r \equiv A_{\rm T}^2/A_{\rm S}^2$
$A_{\rm T}, A_{\rm S}$	Tensor and scalar amplitudes
S	Action
H	Hamiltonian
$\mathcal{H}$	Hamiltonian density
L	Lagrangian
$\mathcal{L}$	Lagrangian density
$H_0, \mathcal{H}_0$	Free field Hamiltonian (density)
$H_{\rm int}, \mathcal{H}_{\rm int}$	Interaction Hamiltonian (density)
π	Momentum conjugate, $\pi \equiv \partial\mathcal{L}/\partial\dot{A}$
$\delta S/\delta A$	Variational derivative
$\Theta(\boldsymbol{n}) \equiv \delta T(\boldsymbol{n})/\bar{T}$	CMB temperature anisotropy field
$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle$	Three point correlation function of $\mathcal{R}$
$B_{\mathcal{R}}(k_1, k_2, k_3)$	Bispectrum of $\mathcal{R}$
$f_{\rm NL}$	Non-linearity parameter
$a_{\ell m}$	Multipole moments
$Y_{\ell m}$	Spherical harmonics
C_ℓ	Angular power spectrum
U	Unitary evolution operator associated with H
U_0	Unitary evolution operator associated with H_0
F	Unitary evolution operator associated with $H_{\rm int}$
$\lvert 0\rangle$	Free vacuum state
$\lvert \Omega\rangle$	Interaction vacuum state
E_0, E_Ω	Energy of free and interaction vacua
N	Lapse function
N^i	Shift function
Σ_t	Constant time spatial hypersurface
n^μ	Time-like normal vector to Σ_t
$h_{\mu\nu}$	Induced metric on hypersurfaces
h	Determinant induced metric, $h \equiv \det h_{\mu\nu}$
$\mathcal{D}_\mu$	Covariant derivative associated with $h_{\mu\nu}$
$K_{\mu\nu}$	Extrinsic curvature tensor
$\mathcal{O}_i$	Generic quantum field operator
δ_i	Mass dimension of operator $\mathcal{O}_i$
c_i, d_i	Wilsonian coefficients
Λ	Energy cut-off of Effective Field Theory
g	Generic interaction coupling
$\partial^2\chi$	Redefinition of $\dot{\mathcal{R}}$, $\partial^2\chi \equiv a^2\varepsilon\dot{\mathcal{R}}$

Symbol	Description
$S_{1,2,3}$	Action to first, second and third order in perturbations
∂S	Spatial or temporal boundary term in the action
G_{IJ}	Field space metric, often set to $G_{IJ} \equiv \delta_{IJ}$
$\mathcal{M}_{IJ}$	Mass-matrix
$U(\theta), S(\theta)$	Rotation matrices, representations of $SO(n)$
λ_I	I-th eigenvalue
σ	Adiabatic field
s_I	I-th entropic field
θ	Two-field angle
$\delta\sigma$	Adiabatic field perturbation
δs	Entropic field perturbation
$T(t_*, t)$	Transfer function
$\mathcal{R}$	Curvature perturbation for two-field model, $\mathcal{R} = H(\delta\sigma/\dot{\sigma})$
S	Entropy perturbation for two-field model, $S = H(\delta s/\dot{\sigma})$
Δ	Adiabicity-entropy correlation angle
Q^I	Quantum fluctuation in I-th field
$J(x)$	Source function
$\delta/\delta J(x)$	Functional derivative
$Z[J]$	Generating functional
$D_Q(x_1 - x_2)$	Propagator field Q

Part I

Background Evolution and Inflation

Chapter 1

Conventional Big Bang Theory

"There is a theory which states that if ever anyone discovers exactly what the universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable. There is another theory which states that this has already happened."

— Douglas Adams

In this first chapter, the main framework of Big Bang theory will be introduced from first principles.[1] That is, the main results of conventional Big Bang theory – such as the metric of the universe – will be derived using General Relativity. The main goal of this chapter will be to set up the essential ingredients of Big Bang theory that will underly the content treated in the coming chapters.

This chapter is organized as follows. In section 1.1 the assumptions at the heart of cosmology will be introduced. These assumptions are often referred as the Cosmological Principle and they put strong constraints on the mathematical description of the universe. Furthermore, the metric expansion of space on large scales and the corresponding Hubble law will be discussed. Subsequently, in section 1.2, the Cosmological Principle and metric expansion of space will be used to derive the metric of the universe, known as the Friedmann-Robertson-Walker (FRW) metric, which we use throughout this thesis to describe the (background) geometry of the universe.[2] Following up on this, the dynamical evolution of the universe will be examined in section 1.3. The main results of this section will be the Friedmann equations, which form one of the cornerstones of modern cosmology. Then, in the final section, the main results of conventional Big Bang theory will be compared with observations.

1.1 Foundations of Cosmology

In order to be able to describe the universe theoretically by means of mathematical models, a number of assumptions should be made to start from and build on. In particular, there are three fundamental assumptions forming the foundation of modern cosmology. Two of these assumptions are concerned with the uniform structure of the universe on large scales and are together commonly referred to as the *Cosmological Principle*. The third assumption relates to the dynamics of the universe on these large scales. In this section, these assumptions

[1] It should be mentioned that this chapter and the next are adapted versions of chapters 4,5 and 6 in previous work [94], performed under supervision of D. Roest and A Chatzistavrakidis and supported by the Honours College department of the University of Groningen (RUG).

[2] The implications of the FRW metric for freely falling particles, which move along geodesics, will be studied in section A.3.

will be introduced and discussed. Furthermore, observational evidence will be provided to support the introduced assumptions.

1.1.1 The Cosmological Principle

As mentioned, the *Cosmological Principle* (CP) is based on two assumptions.[3] First, the distribution of matter in the universe is assumed to *isotropic* on large scales, i.e. for scales of order 100 Mpc.[4] An isotropic space is rotational invariant: the same in every direction. In other words, the universe is assumed to appear the same in all directions when averaged over distance scales of about 100 Mpc or larger. Second, the matter distribution in the universe is assumed to be *homogeneous* on these large scales. A homogeneous space is one that is translationally invariant. That is, averaged on scales of order 100 Mpc, the universe is the same at every point. The requirements of isotropy and homogeneity put strong constraints on the mathematical description of the universe. For instance, the CP singles out a single form for the metric of the universe, as described in section 1.2.

Observational Support

The strongest observational support for the CP comes from (a) galaxy redshift surveys and (b) temperature measurements on the cosmic microwave background (CMB). Below, results from both types of observations are presented and discussed briefly.

Galaxy Redshifts Surveys.—Galaxy redshift surveys determine the distribution of galaxies as a function of distance from earth. Between 1997 and 2002 the Anglo-Australian Observatory conducted the 2dF Galaxy Redshift Survey [39], the obtained distribution of galaxies as function of distance is shown in Fig. 1.1. As can be seen, the distribution of galaxies – represented by blue dots – becomes more and more homogeneous as the distance increases. This pattern occurs for every radial path outward starting at the center, which supports isotropy of the universe.

Temperature CMB.—At early times, the universe was filled with a hot dense plasma of electrons, protons and photons. This plasma was opaque to the photons, since they constantly scattered off electrons via Thomson scattering. Due to this constant scattering, the effective distance travelled by photons was negligibly small. However, in course of time, the universe cooled and eventually the energy scale was reached at which neutral hydrogen could be formed from the electrons and protons: this process is called *recombination*.[5] The stage of recombination took place when the temperature dropped to about $T_{\text{recomb}} = 13.6$ eV (in units where $k_{\text{B}} \equiv 1$), corresponding to the binding energy between the electron and proton in Hydrogen. At that moment, the photons decoupled from the matter and the universe became transparent to them. Since then, the photons streamed freely through the universe and are nowadays observed as background radiation called the Cosmic Microwave Background (CMB).

[3] For a more mathematical discussion of the Cosmological Principle, see chapter 14.1 of [89].

[4] The *Megaparsec*, abbreviated as Mpc, is the unit of length used to quantify distances to objects outside the solar system: $1\,\text{Mpc} = 3.0857 \times 10^{16}\,\text{m}$.

[5] The *re-* in the term *recombination* is misleading, since this really is the first time electrons and protons combine into neutral hydrogen. Hence, the term *combination* would be more appropriate. Unfortunately, the term recombination became standard terminology in literature.

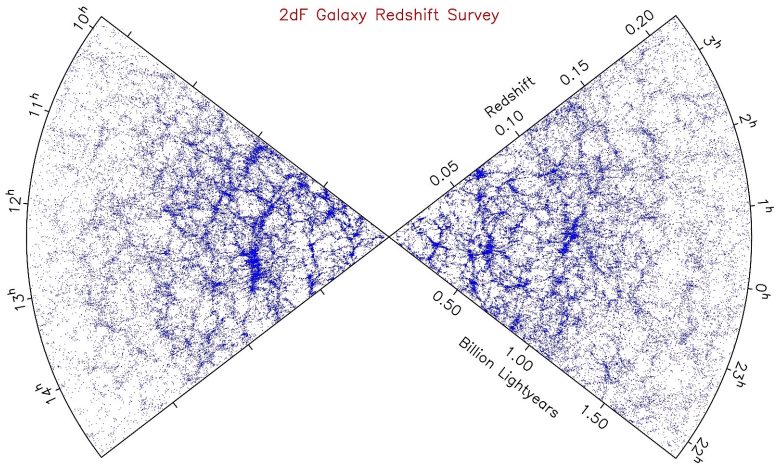


Figure 1.1: Distribution of galaxies (blue dots) as a function of radial distance and direction according to the 2dF Galaxy Redshift Survey [39]. In each direction the distribution of galaxies gets more homogeneous as distance increases: this supports homogeneity and isotropy.

Among other satellites, the Wilkinson Microwave Anisotropy Probe (WMAP) has measured the temperature of the CMB in all directions along the sky. The temperature of the CMB is measured to be the same in every direction to very high accuracy [49]:

$$\bar{T}_{\mathrm{CMB}} = 2.725 \pm 0.002\,\mathrm{K}. \tag{1.1.1}$$

Anisotropies in the temperature as a function of direction are only observed on small relative scales of $\delta T/\bar{T}_{\mathrm{CMB}} = \mathcal{O}(10^{-5})$. Hence, the temperature of the CMB strongly favors the assumption that the universe is isotropic.

1.1.2 The Expanding Universe

The third and final assumption that lies at the core of modern cosmology is concerned with the dynamics of the universe. On sufficiently large scales, again of order 100 Mpc and larger, it is assumed that the universe expands. Mathematically, this implies that the spatial part of the space-time metric, known as the 3-metric, changes as a function time. In particular, the expansion of the universe constrains the components of the 3-metric to be increasing functions of time. Then, since the 3-metric itself increases over time, the physical distance between two points in the universe increases with time: this is called the metric expansion of space.

Mathematically, the metric expansion of space is modelled by the Friedmann-Robertson-Walker metric, which will be derived in the following the section. However, this model of the universe is only valid on large scales: roughly the scale of galaxy clusters and larger. On smaller scales, the metric expansion is suppressed by the gravitational attraction between the present matter.

Hubble's Law

Observationally, the metric expansion of space is described by *Hubble's law*, which states that:

- ▷ Galaxies observed in extragalactic space, at distances of 10 Mpc or more away, are found to have a Doppler shift analogous to a relative velocity away from Earth, known as the recession velocity.
- ▷ The relative velocity of these galaxies away from earth is approximately proportional to their distance to Earth, up to a few 100 Mpc away from earth. For larger distance scales, the notion of distance itself becomes less well-defined and the relation becomes model-dependent: i.e. the matter content of the universe should be taken into account.

Theoretically, the relation between the recession velocity v_r and distance was first derived from General Relativity by G. Lemaître in 1927 [59]. For distances up to approximately 100 Mpc, the relationship between distance d and recession velocity v_r is linear and often expressed as:

$$v_r = H_0 d, \tag{1.1.2}$$

where the constant of proportionality, H_0, is called the Hubble parameter, named after E. Hubble, who was the first to confirm the relation based on observations [51]. Sometimes, H_0 is called the Hubble constant. This terminology is misleading, since the Hubble parameter is emphatically *not* constant over time: see section 1.2.2. In particular, the subscript 0 in H_0 is used to indicate the present day value of the Hubble parameter. The terminology Hubble *constant* originates from the fact that the Hubble parameter is constant over space, as imposed by the CP. The most recent present-day value of the Hubble parameter is:

$$H_0 = 67.31 \pm 0.96\,\mathrm{km\ s^{-1}\ Mpc^{-1}}, \tag{1.1.3}$$

as obtained from observations made by the Planck satellite in 2015 [3].

1.2 Geometry of the Universe

By the Cosmological Principle, the number of possibilities for the geometry of the universe reduces significantly. The constraints of homogeneity and isotropy allow to classify three different geometries for the universe: a Euclidean, spherical or hyperbolic geometry. In this section, the spatial metric encompassing those three geometries will be introduced. Finally, we will discuss Friedmann-Robertson-Walker metric, which describes the expanding geometry of the universe on large scales.

1.2.1 Spatial Geometry and 3-Metric of the Universe

As mentioned, based on spatial homogeneity and isotropy, the spatial geometry of the universe is described (at a specific instant in time) by a constant 3-curvature K. The background evolution of the universe can then be represented as the sequence of constant time hypersurfaces $\mathcal{M}_t$, each of which is homogeneous and isotropic and has a constant 3-curvature. The 3-curvature falls naturally into three different classes: $K < 0$, corresponding to a negatively curved spatial geometry, $K = 0$ (flat geometry) and $K > 0$ (positively curved). Note that only the *sign* of the 3-curvature K is relevant here, since by appropriate rescaling of coordinates, the precise value of K (except for its sign) can be chosen arbitrarily. In what follows, the coordinates will be rescaled such that the curvature parameter K takes on values -1, 0 and $+1$ for negatively curved space, flat space or positively curved space, respectively.

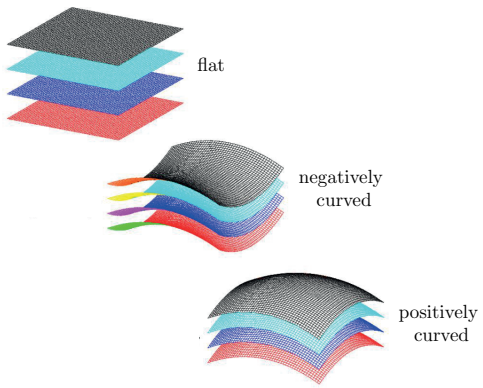


Figure 1.2: Time-ordered spatial slices or hypersurfaces with the possible curvatures: $K = 0$ (flat), $K < 0$ (negatively curved) and $K > 0$ (positively curved).

As we will derive explicitly in Appendix A, the spatial metric γ_{ij} (3-metric) corresponding to the three geometries can be written as:

$$\boxed{\gamma_{ij}(\boldsymbol{x}) = \delta_{ij} + K\frac{x_i x_j}{1 - K(x_k x^k)},} \tag{1.2.1}$$

where K takes on the values -1 0 and $+1$ for negatively curved, flat and positively curved space, respectively. The spatial line element $d\ell^2$ can be written as:

$$d\ell^2 = a^2 d\boldsymbol{x}^2 = a^2 \gamma_{ij} dx^i dx^j, \tag{1.2.2}$$

where the factor a is an increasing function of time (only). In particular, it will take the role of a scale factor accounting for the metric expansion of the universe by stretching the coordinate system over time. Using the above result, the invariant space-time interval $ds^2 \equiv g_{\mu\nu} dx^\mu dx^\nu = -dt^2 + d\ell^2$ becomes:

$$ds^2 = -dt + a^2(t)\left[d\boldsymbol{x}^2 + K\frac{(\boldsymbol{x}\cdot d\boldsymbol{x})^2}{1 - K\boldsymbol{x}^2}\right] = -dt^2 + a^2(t)\gamma_{ij}(\boldsymbol{x})\, dx^i dx^j, \tag{1.2.3}$$

according to the $(-+++)$ metric signature. This result is known as the Friedmann-Robertson-Walker (FRW) metric,[6] and a universe described by this metric is often called a FRW universe.

From the above expression for ds^2, the space-time metric $g_{\mu\nu}$ of the universe can be extracted as:

$$g_{\mu\nu} = \begin{bmatrix} -1 & 0 \\ 0 & a^2\gamma_{ij} \end{bmatrix} \tag{1.2.4}$$

Observe that the symmetry constraints on the geometry of the universe, as imposed by the Cosmological Principle, reduce the ten independent components of the metric into a function of time and a constant: the scale factor $a(t)$ and the curvature parameter K, respectively. Furthermore, note that the metric indeed does not posses non-trivial spatial dependence.

[6]Strictly speaking, Eq. 1.2.3 is *not* the metric, but the invariant space-time interval ds^2 which is related to the *actual* metric $g_{\mu\nu}$ via $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$. However, because of this close relation between ds^2 and $g_{\mu\nu}$, both are often referred to as the metric.

This can be seen most easily for a universe with zero curvature $K = 0$, in that case $\gamma_{ij} = \delta_{ij}$ (by Eq. A.1.10) and the metric becomes:

$$g_{\mu\nu} = \begin{bmatrix} -1 & 0 \\ 0 & a^2\delta_{ij} \end{bmatrix}, \tag{1.2.5}$$

which is clearly independent of spatial coordinates. In the rest of this work, the flat metric will be used mostly as it is favored by observations (see section 1.4).

1.2.2 Comoving and Physical Quantities

Describing distances and velocities in an expanding universe is not straightforward, since the background in which these quantities are defined – the FRW metric – itself changes with time via the scale factor $a(t)$. Therefore, two types of coordinates are used: *comoving* and *physical* coordinates. The comoving coordinates are fixed with the expansion of the universe and hence do not posses time dependence. Physical coordinates, on the other hand, describe the real positions in space, which change as function of time due to the expansion of the universe.

In the FRW metric line element (Eq. 1.2.3):

$$ds^2 = -dt^2 + a^2(t)\gamma_{ij}(\boldsymbol{x})\, dx^i dx^j, \tag{1.2.6}$$

the coordinates $x^i \equiv \{x^1, x^2, x^3\}$ are comoving coordinates. The comoving coordinates $\boldsymbol{x}$ can be transformed to physical coordinates $\boldsymbol{x}_{\text{phys}}$ and vice versa via the relationship:

$$\boldsymbol{x}_{\text{phys}}(t) = a(t)\, \boldsymbol{x}. \tag{1.2.7}$$

The above relation between physical and comoving coordinates explicitly shows that the physical coordinates are time-dependent via the scale factor, whereas the comoving coordinates are not.

In an FRW space-time, the physical velocity of an object is defined as the time derivative of the physical coordinate x^i_{phys}, that is:

$$\boxed{\boldsymbol{v}_{\text{phys}} \equiv \frac{d\boldsymbol{x}_{\text{phys}}}{dt} = a(t)\frac{d\boldsymbol{x}}{dt} + \frac{da}{dt}\boldsymbol{x} = \boldsymbol{v}_{\text{pec}} + H\boldsymbol{x}_{\text{phys}},} \tag{1.2.8}$$

where H is the Hubble parameter in terms of the scale factor and its first time derivative:

$$H(t) \equiv \frac{1}{a}\frac{da}{dt} = \frac{\dot{a}}{a}, \tag{1.2.9}$$

and it follows that H is time-dependent. Notice that the physical velocity consists of two contributions: the *peculiar* velocity $\boldsymbol{v}_{\text{pec}} = a(t)\dot{\boldsymbol{x}}$ and the Hubble *flow*, $H\boldsymbol{x}_{\text{phys}}$. This Hubble flow represents the part of the velocity inherent to the expansion of the universe. The peculiar velocity of an object is velocity relative to the comoving coordinate system. That is, the peculiar velocity is the velocity measured by an observer following the Hubble flow.[7]

[7] More on the kinematics of particles in the FRW universe can be found in Appendix A.3.

Recovering Hubble's Law

Hubble's law (Eq. 1.1.2) can be extracted from Eq. 1.2.8 by assuming $\boldsymbol{v}_{\text{pec}} = 0$ for galaxies, i.e. the galaxies are fixed with the expansion of the universe. In that case, the physical velocity of the galaxies is given by solely the Hubble flow contribution:

$$\boldsymbol{v}_{\text{phys}} = H\boldsymbol{x}_{\text{phys}}. \tag{1.2.10}$$

Now, identifying the physical velocity with the recession velocity v_r and physical distance with the seperation distance d, Hubble's law in the form of Eq. 1.1.2 follows immediately from the above equation.

1.3 Dynamics of the Universe

In the previous section the metric expansion of space is discussed and the corresponding metric $g_{\mu\nu}$ was derived using the Cosmological Principle. As shown, the metric depends only on one dynamical function of time: the scale factor $a(t)$. However, the scale factor is never presented as an explicit function of time: that will be the main purpose of this section. The scale factor depends upon the energy and matter content in the universe as described by the energy-momentum tensor $T_{\mu\nu}$ of the universe. Since the scale factor is part of the metric tensor $g_{\mu\nu}$, it is related to $T_{\mu\nu}$ via the Einstein Field Equations (EFE's)

$$G_{\mu\nu} = 8\pi G\, T_{\mu\nu}. \tag{1.3.1}$$

To be more specific, the EFE's are used to obtain two differential equations, the so-called Friedmann equations, which can be solved for the scale factor as function of time.

To find the Friedmann equations, first the energy-momentum tensor $T_{\mu\nu}$ (right-hand side of EFE's) will be introduced for a homogeneous and isotropic universe in section 1.3.2. Then, in section 1.3.2, conservation of energy and momentum will be used to examine the evolution of the energy density as function of the scale factor. The Einstein tensor $G_{\mu\nu}$ (left-hand side of EFE's) is discussed in section 1.3.4. Finally, in section 1.3.5, the results of the preceding sections are combined to obtain the Friedmann equations.

1.3.1 Energy-Momentum Tensor of the Universe

Just like the metric tensor $g_{\mu\nu}$, also the energy-momentum tensor $T_{\mu\nu}$ must satisfy the requirements of isotropy and homogeneity. In Appendix B.1, it is shown that the CP constrains the energy-momentum tensor to be:

$$\boxed{T_{\mu\nu} = (\rho + P)U_\mu U_\nu + P g_{\mu\nu} = \begin{bmatrix} \rho & 0 \\ 0 & P g_{ij} \end{bmatrix},} \tag{1.3.2}$$

which is known as the energy-momentum tensor of a *perfect fluid*, see also chapter 14.2 of [89].[8] The energy density and pressure of the fluid are given by ρ and P respectively and its four-velocity relative to the (comoving) observer is denoted by $U_\mu \equiv dx^\mu/d\eta$.

[8] We will shortly discuss the energy-momentum tensor for a non-perfect fluid in section 4.8.1.

1.3.2 Energy Conservation and the Continuity Equation

Now that the energy-momentum tensor for the universe is known, the next task is to describe the evolution of the energy density ρ in an expanding universe. To do that, conservation of energy and momentum will be considered first. Mathematically, energy and momentum conservation is represented by a vanishing covariant derivative of the energy-momentum tensor:

$$\nabla_\mu T^{\mu\nu} = 0, \tag{1.3.3}$$

where the covariant form of the energy-momentum tensor is obtained as follows:

$$T^{\mu\nu} = g^{\mu\rho} g^{\nu\sigma} T_{\rho\sigma} = \begin{bmatrix} \rho & 0 \\ 0 & P g^{ij} \end{bmatrix}. \tag{1.3.4}$$

Writing out the covariant derivative $\nabla_\mu T^{\mu\nu}$ yields:

$$\nabla_\mu T^{\mu\nu} = \partial_\mu T^{\mu\nu} + \Gamma^\mu_{\mu\alpha} T^{\alpha\nu} + \Gamma^\nu_{\mu\alpha} T^{\mu\alpha} = 0. \tag{1.3.5}$$

Evaluating the above expression for $\nu = 0$ gives the evolution equation for the energy density, which is known as the continuity equation:

$$\dot{\rho} + 3H(\rho + P) = 0. \tag{1.3.6}$$

This differential equation may be solved for ρ in terms of a. To do so, it is convenient to define the equation of state parameter $w = P/\rho$, which assumes that there is a linear relationship between the pressure and the energy density of the fluid. The solution $\rho = \rho(a)$ is:

$$\rho \propto a^{-3(1+w)}. \tag{1.3.7}$$

Lastly, for later analysis, it proves convenient to rewrite the continuity equation in the form $|d\ln\rho/d\ln a|$, reading:

$$\boxed{\left|\frac{d\ln\rho}{d\ln a}\right| = 3(1+w).} \tag{1.3.8}$$

1.3.3 Energy and Matter Content in the Universe

The universe went through different era's characterized by different dominating matter components (we will come back to this in section 1.4). For instance, at early times the universe was dominated by photons (i.e. radiation) and in course of time, the universe became matter dominated. Nowadays, the universe is dominated by so-called dark energy, since almost 70% of the energy and matter content in the universe is in the form of dark energy. These different components correspond to different relations between the pressure P and the density ρ and hence to different equations of state w.

Below, the known energy and matter components constituting the content in the universe will in classified based on their equation of state:

- ▷ *Matter.*—All components for which the pressure is negligibly small compared to the energy density, i.e. $P \ll \rho$, are referred to as matter. In the limit $P \ll \rho$, the equation of state vanishes $w \to 0$ and Eq. 1.3.7 gives:

$$\rho \propto a^{-3} \propto V^{-1}, \tag{1.3.9}$$

since $V \propto a^3$. Note that this inverse proportionality between ρ and V originates solely from the expansion of the universe. This is the case for a gas of non-relativistic particles, for which the energy density is dominated by the mass. The sources that behave in this way are dark matter and ordinary matter (nuclei and electrons). Usually, cosmologists refer to ordinary matter as just baryons, which is strictly speaking wrong, since electrons are leptons. However, most of the mass of ordinary matter is contained in baryons since as they are much heavier compared to electrons ($m_p/m_e = \mathcal{O}(10^3)$). Therefore, ordinary matter is usually referred to as baryons.

▷ *Radiation.*—For a gas of relativistic particles, the pressure is approximately one-third of the energy density:

$$P = \frac{1}{3}\rho, \tag{1.3.10}$$

and the equation of state is $w = 1/3$. Components satisfying $w = 1/3$ are referred to as radiation. Substitution of this value for w in Eq. 1.3.7 gives:

$$\rho \propto a^{-4}. \tag{1.3.11}$$

In this case the dilution of the energy density includes both the expansion of the universe, contributing a^{-3}, and the red-shifting of the energy, which contributes a^{-1}. Particle species that behave like radiation are photons, neutrinos and gravitons.

▷ *Dark Energy.*—Finally, a negative pressure component is needed to describe the observed universe (see section 1.4). This component is known as dark energy (DE) and satisfies:

$$P \approx -\rho, \tag{1.3.12}$$

with the corresponding equation of state $w_{\mathrm{DE}} \approx -1$. Remarkably, the energy density of this component does not dilute according to Eq. 1.3.7:

$$\rho \propto a^0 = 1. \tag{1.3.13}$$

Since the energy density remains constant with the expansion, additional dark energy must be created in course of time to counteract the effect of the expansion on the energy density. To very good approximation, dark energy behaves the same as vacuum energy or a cosmological constant Λ (see last part of section 1.4). This is the reason why the terms dark energy, cosmological constant and vacuum energy are often used interchangeably in literature as well as in this work, whereas they formally have different meanings.

1.3.4 Einstein Tensor of the FRW Universe

In this section, the Einstein tensor $G_{\mu\nu}$ of the FRW universe will be introduced. Together with the cosmological constant Λ, the Einstein tensor constitutes the left-hand side of the EFE's and it can be written as:

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}, \tag{1.3.14}$$

where $R_{\mu\nu}$ and R are the Ricci tensor and scalar, respectively. Since the FRW metric $g_{\mu\nu}$ is already known, computation of $R_{\mu\nu}$ and R using the FRW metric will yield the complete expression for $G_{\mu\nu}$.

The Ricci tensor $R_{\mu\nu}$ is defined as the contraction of the upper index with the middle lower index of the Riemann tensor and reads:

$$R_{\mu\nu} = R^{\sigma}_{\mu\sigma\nu} = \partial_\lambda \Gamma^\lambda_{\mu\nu} - \partial_\nu \Gamma^\lambda_{\mu\lambda} + \Gamma^\lambda_{\lambda\rho}\Gamma^\rho_{\mu\nu} - \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\rho}. \tag{1.3.15}$$

The only non-vanishing components of the Ricci tensor are R_{00} and R_{ij}, where the latter is proportional to the spatial metric g_{ij}:

$$\boxed{R_{00} = -3\frac{\ddot{a}}{a}, \qquad R_{ij} = \left[\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{K}{a^2}\right] g_{ij}.} \tag{1.3.16}$$

From the Ricci tensor, the Ricci scalar R is obtained by contraction with the inverse metric $g^{\mu\nu}$ as follows:

$$R = g^{\mu\nu} R_{\mu\nu}. \tag{1.3.17}$$

Consequently, the Ricci scalar R becomes:

$$\boxed{R = 6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2}\right].} \tag{1.3.18}$$

Using $R_{\mu\nu}$ and R, the Einstein tensor $G_{\mu\nu}$ takes the form:

$$G_{\mu\nu} = R_{\mu\nu} - 3\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2}\right]. \tag{1.3.19}$$

In Appendix B.3, the Ricci tensor and scalar are derived according to the definitions above and the Christoffel symbols of the FRW metric (Eq. A.2.3).

1.3.5 The Friedmann Equations

Now that both the Einstein tensor $G_{\mu\nu}$ (Eq. 1.3.19) and the energy-momentum tensor $T_{\mu\nu}$ (Eq. 1.3.2) are known for the FRW universe, the EFE's can be used to deduce the Friedmann equations. Using the obtained expressions for $G_{\mu\nu}$ and $T_{\mu\nu}$, the EFE's become:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = R_{\mu\nu} - 3\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2}\right] g_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{T_{\mu\nu}}{M_{\rm pl}^2}, \tag{1.3.20}$$

The first Friedmann equation is obtained by considering the above equation for $\mu = \nu = 0$ and reads:

$$\boxed{\left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2} = \frac{1}{3M_{\rm pl}^2}\left(\rho + \rho_\Lambda\right),} \tag{1.3.21}$$

where $\dot{a}/a = H$ is the Hubble parameter and $\rho_\Lambda = \Lambda/8\pi G = M_{\rm pl}^2\Lambda$ is the energy density due to the cosmological constant Λ. The second Friedmann equation, also known as the acceleration equation, is obtained by considering $\mu = i$ and $\nu = j$ in the EFE's:

$$\boxed{\frac{\ddot{a}}{a} = -\frac{1}{6M_{\rm pl}^2}(3P + \rho - 2\rho_\Lambda).} \tag{1.3.22}$$

In Appendix B.4, the first and second Friedmann equations are explicitly derived from the EFE's (Eq. 1.3.20).

Matter Content	w	$\rho(a)$	$a(t)$
Radiation Dominated (RD)	$1/3$	a^{-4}	$t^{1/2}$
Matter Dominated (MD)	0	a^{-3}	$t^{2/3}$
Vacuum Energy Dominated (ΛD)	-1	a^{0}	e^{Ht}

Table 1.1: Solutions for the energy density and scale factor for universes dominated by different matter components (corresponding to different equations of state w).

1.3.6 Scale Factor Evolution in a Single Component Universe

Combining the first Friedmann equation (Eq. 1.3.21) with the solution for the energy density $\rho = \rho(a)$ (Eq. 1.3.7), the time dependence of the scale factor can be obtained for single-component universes. Single component universes are filled with only one form of matter and are good approximations to the real universe during times where the considered matter component dominates all the others.

In order to simplify the derivation of the scale factor for such universes, the curvature parameter is set to zero[9] $K \equiv 0$. Furthermore, since only single-component universes are considered here, there is no need to distinguish between energy density ρ and the contribution due to the cosmological constant ρ_Λ. Therefore, the first Friedmann equation is rewritten as:

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{\rho}{3M_{\rm pl}^2}, \tag{1.3.23}$$

since $\rho + \rho_\Lambda \to \rho$ and spatial curvature is neglected, $K \equiv 0$. Using the solution $\rho = \rho(a)$ as given by Eq. 1.3.7, the previous equation becomes:

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 \propto \frac{1}{3M_{\rm pl}^2}\, a^{-3(1+w)}. \tag{1.3.24}$$

Omitting the irrelevant factor $1/3M_{\rm pl}^2$ and rewriting in terms of $\dot{a}$ yields:

$$\dot{a} \propto a^{-3(1/3+w)/2}. \tag{1.3.25}$$

Integration gives the following form for the time dependence of a for an arbitrary equation of state w (except for the special case $w = -1$, see below):

$$a(t) \propto t^{2/3(1+w)}. \tag{1.3.26}$$

In the case of matter domination (MD) and radiation domination (RD) the equations of state are $w = 0$ and $w = 1/3$ respectively and the scale factor evolves as:

$$a(t) \propto t^{2/3} \quad \text{(MD)}, \qquad a(t) \propto t^{1/2} \quad \text{(RD)}. \tag{1.3.27}$$

For the case $w = -1$, corresponding to a cosmological constant (vacuum energy) and to good approximation to dark energy, the time evolution of the scale factor assumes a different

[9] In the next section it will be shown that the assumption $K = 0$ is indeed favored by observations made by e.g. the Planck satellite [3].

form. According to Eq. 1.3.7, the energy density of a matter component satisfying $w = -1$ does not dilute with the expansion and $\rho \propto a^0 = 1$. It follows immediately that the Hubble parameter is constant in that case, since $H^2 \propto \rho$ according to the first Friedmann equation. The differential equation for the scale factor and its solution thus read:

$$\dot{a} = Ha, \qquad a \propto e^{Ht} \ \ (\Lambda\text{D}), \tag{1.3.28}$$

where $H \neq H(t)$. The solutions for the three different single component universes discussed above (RD, MD and ΛD) are summarized in Table 1.1.

De Sitter Universe

A universe containing solely vacuum energy is often referred to as a *De Sitter* (dS) universe, named after W. De Sitter, a Dutch physicist who first proposed this type of single-component universe. The space-time of a dS universe is given by:

$$ds^2 = -dt^2 + a^2(t)\, d\boldsymbol{x}^2 = -dt^2 + e^{2Ht} d\boldsymbol{x}^2. \tag{1.3.29}$$

Since the scale factor is given by $a(t) = e^{Ht}$ and H is constant in time, a dS universe expands exponentially with time. In other words, a dS universe undergoes acceleration expansion.

Notice that a perfect dS universe never stops expanding exponentially since $H \neq H(t)$ and a dS universe can therefore never transit into an era with a different scale factor (such as the radiation dominated universe). For that reason, a dS space-time may only serve as a good approximation to cosmological era's in which the expansion is almost exponential. This is the case for inflation, where the Hubble parameter only depends mildly on time and the scale factor can thus be approximated by $a \simeq e^{Ht}$. As a consequence, the inflationary era is often referred to as a quasi-De Sitter period: this will come back in the next chapter.

1.4 The ΛCDM Model and Observations

Cosmological observations made by for instance the Planck satellite [3] show that the current universe is filled with radiation, matter and dark energy (or vacuum energy). Since the vast majority of the energy content is in the form of dark energy Λ (70%) and cold dark matter (25%), the model that corresponds best to these observations is called the ΛCDM model. The energy density contribution due to radiation, matter and dark energy are denoted by ρ_r, ρ_m and ρ_Λ, respectively. In addition, the matter and radiation densities are usually divided into contributions of different matter and radiation species. For matter, ρ_m is the sum of baryons ρ_b and cold dark matter ρ_{CDM} (CDM):

$$\rho_m = \rho_b + \rho_{\text{CDM}}. \tag{1.4.1}$$

Similarly, the radiation density is given by the sum of the contributions due to photons ρ_γ and neutrinos ρ_ν:

$$\rho_r = \rho_\gamma + \rho_\nu. \tag{1.4.2}$$

Usually, the density of the different contributions is represented in units of the so-called critical density ρ_c. This critical density is defined as the density for which the first Friedmann

equation yields zero curvature ($K = 0$). To find an expression for ρ_c, the first Friedmann equation is written as:

$$\frac{K}{a^2} = \frac{\rho}{3M_{\rm pl}^2} - H^2. \tag{1.4.3}$$

For K to be zero, it follows from the above equation that ρ must satisfy:

$$\rho_c = 3M_{\rm pl}^2 H^2 = \frac{3H^2}{8\pi G} = 1.87847(23) \cdot h^2\, 10^{-29}\,\mathrm{g\,cm^{-3}}, \tag{1.4.4}$$

where the numerical value is taken from Ref. [71] and $h = 0.673(12)$. The critical density can now be used to define the dimensionless density parameter for matter component i as:

$$\boxed{\Omega_i \equiv \frac{\rho_i}{\rho_c}.} \tag{1.4.5}$$

The sum of all contributions is indicated by the total density paramter Ω (without a subscript):

$$\Omega = \sum_i \Omega_i. \tag{1.4.6}$$

For a flat universe, $K = 0$, the total density parameter is equal to one: $\Omega = 1$.

As for the scale factor, which is often set to unity today, present-day values for the dimensionless density parameters will be indicated using a zero in the subscript. To illustrate this convention, $\Omega_{r,0}$ is today's density radiation density in units of the critical density $\rho_{c,0}$ today, which corresponds to the current value of the Hubble parameter H_0:

$$\rho_{c,0} \equiv \frac{3H_0^2}{8\pi G}. \tag{1.4.7}$$

Furthermore, this convention allows to write the density ρ_i corresponding to a generic value for the scale factor a in terms of the present-day density $\rho_{i,0}$ and scale factor a_0. For instance, recall that radiation density scales as $\rho_r \propto a^{-4}$ and therefore $\rho_r(a)$ can be written as:

$$\rho_r = \rho_{r,0}\left(\frac{a_0}{a}\right)^4. \tag{1.4.8}$$

Assuming the energy density splits into contributions due to dark energy (Λ), radiation and matter, it can be written as:

$$\rho = \rho_\Lambda + \rho_m + \rho_r = \rho_{\Lambda,0} + \rho_{m,0}\left(\frac{a_0}{a}\right)^3 + \rho_{r,0}\left(\frac{a_0}{a}\right)^4. \tag{1.4.9}$$

Notice that for dark energy the density $\rho_\Lambda = \rho_{\Lambda,0}$, since it does not dilute with the expansion.

At early times, the content in the universe was dominated by radiation. This is expected since the radiation density decays fastest with time of all components since $\rho(a) \propto a^{-4}$ (see Tab. 1.1). At late times, the universe is dominated by the dark energy Λ since ρ_Λ does not dilute with time and both the matter and radiation density decay away with time (again, see 1.1).

Invoking the above notation for the total energy density, the first Friedmann equation (Eq. 1.3.21 with ρ_Λ absorbed in ρ) may be divided by the current critical energy density $\rho_{c,0}$ to obtain:

$$H^2(a) = H_0^2\left[\Omega_{\Lambda,0} + \Omega_{m,0}\left(\frac{a_0}{a}\right)^3 + \Omega_{r,0}\left(\frac{a_0}{a}\right)^4 + \Omega_{K,0}\left(\frac{a_0}{a}\right)^2\right], \tag{1.4.10}$$

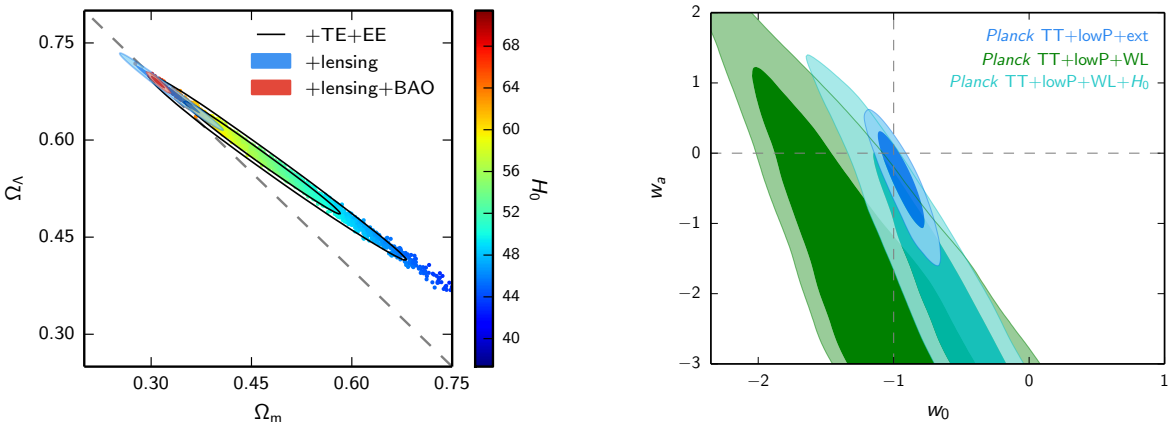


Figure 1.3: Left: confidence contours in the Ω_m-Ω_Λ plane, note that combined data (Planck + lensing + BAO) simultaneously constrains the possible values Ω_m and Ω_Λ to a narrow range in which their sum is approximately one. This supports a flat spatial geometry of the universe, i.e. $\Omega_k = 0$. Right: constraints on the equation of state of dark energy. The time dependence of the equation of state is governed by w_a: no time dependence corresponds to $w_a = 0$, in which case $w_{\rm DE} = w_0$. Observe that combined data (Planck + ext) suggests that dark energy has a time independent equation of state close to that of a perfect cosmological constant: $w_a = 0$ and $w_0 = -1$. Figures taken from [3].

where the density parameter for curvature is defined as $\Omega_{k,0} \equiv -k/(a_0 H_0)^2$. Finally, assigning the value one to the current scale factor yields:

$$\frac{H^2}{H_0^2} = \Omega_\Lambda + \Omega_m a^{-3} + \Omega_r a^{-4} + \Omega_K a^{-2}. \tag{1.4.11}$$

In the final expression for the Friedmann equation, the zeros in the subscripts denoting the present-day values are omitted again to be in line with usual convention in literature. That is, in literature Ω_i refers to the density of the matter component i *today*, measured in units of the critical density *today*.

Observations

Constraints on Density Parameters.—Observational bounds on cosmological parameters such as the density parameters are deduced from the measurements of the Planck satellite on the Cosmic Microwave Background (CMB) [3]. To obtain even stronger constraints on these parameters, Planck data (*Planck*) is combined with gravitational lensing reconstruction (*lensing*) and external data (*ext*) from e.g. Baryon Acoustic Oscillations (BAO) and measurements on the Hubble parameter.

At the confidence level of 1σ (68%), the values for Ω_Λ and Ω_m are:[10]

$$\Omega_\Lambda = 0.6911 \pm 0.0062, \qquad (68\%,\ \text{Planck + lensing + ext}), \tag{1.4.12}$$

$$\Omega_m = 0.3089 \pm 0.0062, \qquad (68\%,\ \text{Planck + lensing + ext}), \tag{1.4.13}$$

[10] Values from last column in Table 4 of Ref. [3].

when the various datasets are combined. The confidence contours of the results are also shown in the Ω_m-Ω_Λ plane (left panel of Fig. 1.3) Since the matter density splits up into baryonic and CDM contributions, the density parameter splits as well: $\Omega_m = \Omega_b + \Omega_{\rm CDM}$. The bounds on both Ω_b and $\Omega_{\rm CDM}$ are:

$$\Omega_b\, h^2 = 0.02230 \pm 0.00014, \qquad (68\%,\ \text{Planck + lensing + ext}), \tag{1.4.14}$$

$$\Omega_{\rm CDM}\, h^2 = 0.1188 \pm 0.0010, \qquad (68\%,\ \text{Planck + lensing + ext}), \tag{1.4.15}$$

where $h = 0.673(12)$. Using the value for h, the values for Ω_b and $\Omega_{\rm CDM}$ are approximately 0.049 and 0.26, respectively. Note that (only) 5% of the cosmic inventory is ordinary matter (baryons). Radiation contributes a minuscule relative amount of $\Omega_r \simeq 8.24 \cdot 10^{-5}$ to the total energy content. It should be mentioned that the relative minority of radiation in the current universe does not mean their physics is irrelevant. In fact, photons arguably play the most crucial role in the exploration of the early universe via the fingerprint they left in CMB: more on this in later chapters. Finally, the bound on the curvature density parameter Ω_K strongly suggests that universe obeys a flat geometry (i.e. $K = 0$), since the obtained value is:

$$\Omega_K = 0.000 \pm 0.005, \qquad (95\%,\ \text{Planck + lensing + BAO}). \tag{1.4.16}$$

Notice that the flat geometry of the universe also implicitly follows from the left panel of Fig. 1.3, since the best constraint values (red confidence contours) for Ω_m and Ω_Λ sum to one to good approximation.

Equation of State of Dark Energy.—In the final part of section 1.3.2, it is argued that dark energy (DE) has an equation of state very similar to that of a perfect cosmological constant, i.e. $w_{\rm DE} \approx -1$. This is indeed verified by experiment. In particular, the best constraint on the DE equation of state currently available is given by:

$$w_{\rm DE} = -1.019^{+0.075}_{-0.080}, \qquad (95\%,\ \text{Planck + lensing + ext}), \tag{1.4.17}$$

corresponding to Eq. 52c in Ref. [3]. The observational constraints on $w_{\rm DE}$ indeed show that DE behaves very similar to a true cosmological constant Λ, for which $w_\Lambda = -1$.

However, if w differs from -1, which is still possible within the current bounds, it is very likely that w changes with time. To account for this possibility, the equation of state can be expanded to first order in the scale factor, which then automatically induces time dependence since $a = a(t)$. This gives:

$$w_{\rm DE} = w_0 + (1 - a) w_a, \tag{1.4.18}$$

where w_0 and w_a a constants independent of time – all time dependence is captured in the scale factor a. Note that by construction of the above expansion, $w_a = 0$ would imply a time-independent equation of state $w = w_0$. Constraints on both w_0 and w_a in terms of confidence contours are given in right panel of Fig. 1.3. Observe that the strongest bounds indeed indicate that DE has a time-independent equation of state very comparable to that of perfect cosmogical constant, since w_0 and w_a are approximately -1 and 0, respectively.

Chapter 2

Big Bang Puzzles and Inflation

"We cannot solve our problems with the same thinking we used when we created them."

— Albert Einstein

The Friedmann-Robertson-Walker (FRW) cosmology described in the last chapter can be brought in good agreement with certain observations, as discussed in section 1.4. However, conventional Big Bang theory also suffers from central problems – or puzzles – caused by cosmological observations that can *not* be explained by the theory. In this chapter, some of the so-called *Big Bang Puzzles* will be discussed. In particular, in this chapter two questions that cannot be explained by conventional Big Bang theory will be introduced.

The first question or puzzle is concerned with the flatness of the geometry of the universe. It follows from observations that the present universe possesses an almost completely flat geometry, corresponding to $K = 0$ and $\Omega_K = 0$ (see section 1.4). However, as will be shown in this chapter, the observed flatness of the universe today requires an enormous amount of fine-tuning in the early universe. This Big Bang puzzle is usually referred to as the *flatness problem*.

The second puzzle is related to the homogeneity and isotropy of the Cosmic Microwave Background (CMB). Despite the fact that the observed uniformity of the CMB indeed supports the Cosmological Principle, it also raises a big puzzle concerned with causality. Namely, according to conventional Big Bang theory, the CMB should consist of about 10^4 causally disconnected patches. However, all these patches appear are observed to have the same temperature to very high degree. For the different patches to have the same temperature, i.e. attain a state of thermal equilibrium, they have to interact with each other, which is not possible since the patches are causally disconnected. This puzzle is known as the *horizon problem*.

In order to resolve these apparent puzzles, an extension to the conventional Big Bang theory will be introduced: *inflation*. This extension, which essentially adds an epoch of accelerated expansion to the early universe, was first proposed by A. Guth in 1981 [47]. In this chapter, it will be described how inflation dynamically solves the addressed puzzles of conventional Big Bang theory. Finally, different conditions for inflation will be introduced and their equivalence will be demonstrated. Eventually, this analysis will lead to the conclusion that the fluid driving inflation possesses properties very different from ordinary matter fluids such as baryons or radiation.

2.1 Conformal Time, Horizons and the Growing Hubble Sphere

In order to start a quantitative discussion of the Big Bang puzzles, it turns out to be convenient to introduce some preliminary concepts first. To be specific, the notion of conformal time and horizons will be of particular interest in describing – and solving – the puzzles.

2.1.1 Conformal Time and Light Propagation

In studying the propagation of light and hence causal structure in FRW space-time, it is convenient to redefine the time coordinate as:

$$d\tau \equiv \frac{dt}{a(t)}, \tag{2.1.1}$$

where τ is called conformal time. In literature, conformal time may also be denoted by the variable η, i.e. $d\eta \equiv dt/a(t)$, but in this work η will be reserved for a so-called Hubble flow parameter to be introduced later. Under the redefinition of the temporal coordinate, the metric for a flat universe in spherical coordinates becomes:

$$ds^2 = a^2(\tau)\left[-d\tau^2 + dr^2 + r^2 d\Omega^2\right] \equiv a^2(\tau)\, \eta_{\mu\nu}\, dx^\mu dx^\nu, \tag{2.1.2}$$

where $d\Omega$ is the differential solid angle. Notice that the FRW metric is factorized into a time-dependent conformal factor $a(\tau)$ and a static Minkowski metric $\eta_{\mu\nu}$.

To study the propagation of photons, $ds^2 \equiv 0$ since they move along null-geodesics. Furthermore, only radial propagation will be considered, that is $\theta = \phi = 0$. In that case, their path is defined as:

$$\Delta r(\tau) = \pm\Delta\tau + C, \tag{2.1.3}$$

where C is a constant. The above result reveals the main advantage of the usage of conformal time over regular coordinate time t, since in the r-τ plane light propagates at angles of 45 degrees.

2.1.2 Horizons

Using the definition of conformal time τ, two horizons can now be defined in a convenient way: the particle and event horizon. The former is related to the possible histories of a photon whereas the latter constrains the future of a photon. The notion of particle and event horizons is discussed here solely in relation to photons, but the discussion here also holds for any other particle species, since no particle can propagate faster than the speed of light.

Particle Horizon.—It follows from Eq. 2.1.3 that the maximal comoving distance a photon can travel between two times τ_1 and $\tau_2 > \tau_1$ is simply given by:

$$\Delta\tau = \tau_2 - \tau_1, \tag{2.1.4}$$

since $c \equiv 1$. Therefore, assuming the Big Bang 'began' with a singularity at initial time $t_i \equiv 0$, the largest comoving distance from which an observer at any later time t can receive signals (light) reads:

$$r_{\rm PH}(\tau) = \tau - \tau_i = \int_{t_i}^{t} \frac{dt}{a(t)}, \tag{2.1.5}$$

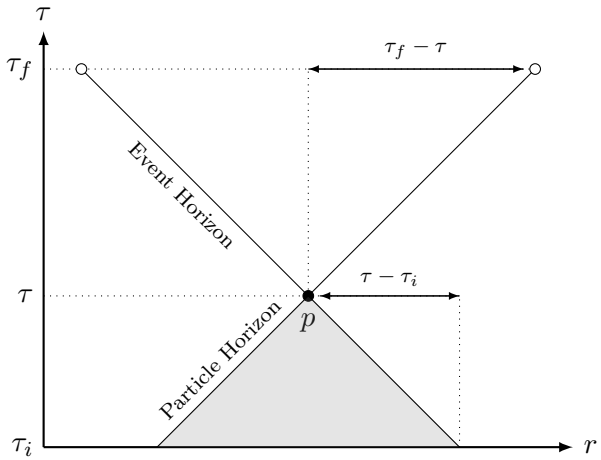


Figure 2.1: Visualization of the particle and event horizons in the r-τ plane for an observer located at p.

this is called the (comoving) particle horizon (PH).

Event Horizon.—Just as the there are restrictions on the possible distance from which an observer can retrieve signals, there are also restrictions on the maximal comoving distance that signals can reach within an amount of time. The maximum distance a signal can travel between emission time τ and a later time $\tau_f > \tau$ is given by:

$$r_{\mathrm{EH}}(\tau) = \tau_f - \tau = \int_t^{t_f} \frac{dt}{a(t)}. \tag{2.1.6}$$

This quantity is called the (comoving) event horizon, where the final time t_f may be taken as $+\infty$. However, note that infinite final time t_f does not necessarily correspond to infinite final conformal time τ_f: this depends on the form of the scale factor $a(t)$. For instance, for a universe dominated by dark energy, the scale factor goes approximately like $a(t) = e^{Ht}$, where H is only slowly varying with time. Neglecting the mild time dependence in H, the event horizon from t_0 to $t_f \to +\infty$ becomes:

$$r_{\mathrm{EH}} = \int_{t_0}^{+\infty} e^{-Ht}\, dt \longrightarrow H(e^{-Ht_0} - 1), \tag{2.1.7}$$

which is clearly a finite result. The notion of the particle and event horizon is also visualized in the conformal diagram in Fig. 2.1.

2.1.3 The Growing Hubble Sphere

The particle horizon r_{PH}, which will be relevant in describing the causal structure of the FRW universe and the discussion of the horizon problem, can be rewritten in a way that relates r_{PH} to the so-called comoving Hubble radius $(aH)^{-1}$. To relate r_{PH} to $(aH)^{-1}$, note that using $dt = da/\dot{a}$ the particle horizon changes to:

$$r_{\mathrm{PH}} = \int_{t_i}^{t} \frac{dt}{a} = \int_{a_i}^{a} \frac{da}{a\dot{a}}. \tag{2.1.8}$$

Now using $d\ln a = da/a$ and $H = \dot{a}/a$, the particle horizon in terms of the comoving Hubble radius becomes:

$$r_{\rm PH} = \int_{\ln a_i}^{\ln a} \frac{1}{aH}\, d\ln a. \tag{2.1.9}$$

The above result relates the causal structure of the FRW space-time to the evolution of the comoving Hubble radius $(aH)^{-1}$.

Assuming the universe is dominated by a fluid with a constant equation of state $w \equiv P/\rho$, Eq. 1.3.25 can be used to relate the comoving Hubble radius to the scale factor and the equation of state:

$$(aH)^{-1} = \dot{a}^{-1} = H_0^{-1} a^{(1+3w)/2}, \tag{2.1.10}$$

where the proportionality factor H_0^{-1} is used to scale $(aH)^{-1}$ according to its present value (with $a_0 \equiv 1$). Since $a(t)$ is an increasing function with time, it follows from the above relation that in case:

$$1 + 3w > 0, \tag{2.1.11}$$

the comoving Hubble radius $(aH)^{-1}$ increases as the universe expands in time. The condition $1 + 3w > 0$ is known as the *Strong Energy Condition* (SEC) and is satisfied by all common fluids, e.g. matter and radiation.[1] In conventional Big Bang theory, the increasing behavior of the comoving Hubble radius $(aH)^{-1}$ is referred to as the increasing Hubble sphere.

The particle horizon $r_{\rm PH}$ can be evaluated explicitly conform Eq. 2.1.5 for a fluid with equation of state $w > -1/3$ as:

$$r_{\rm PH} \equiv \tau - \tau_i = H_0^{-1} \int_{\ln a_i}^{\ln a} e^{(\ln a)(1+3w)/2}\, d\ln a = \frac{2H_0^{-1}}{1+3w}\left[a^{(1+3w)/2} - a_i^{(1+3w)/2}\right]. \tag{2.1.12}$$

Hence, it follow that, for $w > -1/3$, the initial conformal time τ_i is given by:

$$\tau_i = \frac{2H_0^{-1}}{1+3w}\, a_i^{(1+3w)/2}. \tag{2.1.13}$$

Since for all fluids satisfying $w > -1/3$ the Hubble sphere $(aH)^{-1}$ increases, the particle horizon receives the largest contribution from the upper integration limit and the contribution of early times is small. In fact, assuming (a) that the fluid satisfies $w > -1/3$ and (b) that the Big Bang singularity is described by the limit $a_i \to 0$, the initial conformal time value τ_i vanishes:

$$\tau_i = \frac{2H_0^{-1}}{1+3w}\, a_i^{(1+3w)/2} \xrightarrow{a_i \to 0,\ w>-1/3} 0. \tag{2.1.14}$$

In other words, conventional Big Bang evolution starts at $\tau_i = 0$, this (apparent) result will be of significant importance in describing and solving the horizon problem.

2.2 Puzzle 1: The Flatness Problem

The first puzzle of conventional Big Bang theory to address is the so-called *flatness problem.* According to the observations described in section 1.4, the present universe obeys a flat geometry to very high degree, corresponding to $K = \Omega_K = 0$ and $\Omega(a_0) = 1$. However,

[1] Note that the SEC is not satisfied by Dark Energy or a Cosmological Constant.

the flatness of the universe today requires a tremendous amount of fine-tuning in the early universe.

We will describe the flatness problem quantitatively by rewriting the first Friedmann in terms of the density parameter Ω. To obtain this form, the first Friedmann equation:

$$\frac{K}{a^2} = \frac{\rho}{3M_{\rm pl}^2} - H^2, \tag{2.2.1}$$

is rewritten in the form:

$$1 - \frac{\rho}{3M_{\rm pl}^2 H^2} = 1 - \Omega(a) = -\frac{K}{(aH)^2}. \tag{2.2.2}$$

Note that now the density parameter $\Omega = \Omega(a)$ depends on the scale factor and hence on time. Finally, the first Friedmann equation becomes:

$$\boxed{\Omega(a) - 1 = \frac{K}{(aH)^2}.} \tag{2.2.3}$$

Notice that the right hand side of the last equation is dependent on time. In particular, the comoving Hubble radius $(aH)^{-1}$ increases with time in conventional Big Bang theory, that is:

$$(aH)^{-1} \propto t^{\xi}, \tag{2.2.4}$$

where ξ is a constant defined as $\xi \equiv (1+3w)/3(1+w)$. This quantity ξ is greater than zero when the SEC is satisfied ($w > -1/3$).[2] Therefore, the time-dependence of $\Omega(a) - 1$ can be expressed as:

$$\Omega(a) - 1 \propto K t^{2\xi}, \tag{2.2.5}$$

which shows that for $\xi > 0$ the difference between Ω and one increases over time.

As mentioned, current observations show that the present values of Ω_K and K are very close to zero, such that $\Omega(a_0) = 1$ by Eq. 2.2.3. But by the time evolution shown in the previous equation, the present observational fact $\Omega(a_0) = 1$ implies that K must have been very close to zero at early times, since in case there was some curvature in the early universe, it would have grown rapidly with time according to Eq. 2.2.5. Stated otherwise, achieving today's observed value $\Omega(a_0) = 1$ requires K to be extremely fine-tuned at early times. To meet the present day value of Ω, the deviation from flatness $|\Omega(a) - 1|$ must satisfy the following conditions during Big Bang Nucleasynthesis (BBN), the GUT era and the Planck era [14, 15]:

$$\begin{aligned} |\Omega(a_{\rm BBN}) - 1| &\leq \mathcal{O}(10^{-16}), \\ |\Omega(a_{\rm GUT}) - 1| &\leq \mathcal{O}(10^{-55}), \\ |\Omega(a_{\rm pl}) - 1| &\leq \mathcal{O}(10^{-61}), \end{aligned} \tag{2.2.6}$$

The above results show that approaching the Big Bang singularity $a \to 0$, requires more and more fine-tuning in Ω (and hence K). The tremendous amount of fine-tuning needed to agree with present-day observations is one approach to describe the flatness problem. The Flatness problem is also visualized in Fig. 2.2.

[2] The quantity ξ can be derived by using $(aH)^{-1} = \dot{a}^{-1}$ and combining the expressions for $\dot{a}$ (Eq. 1.3.25) and a (Eq. 1.3.26).

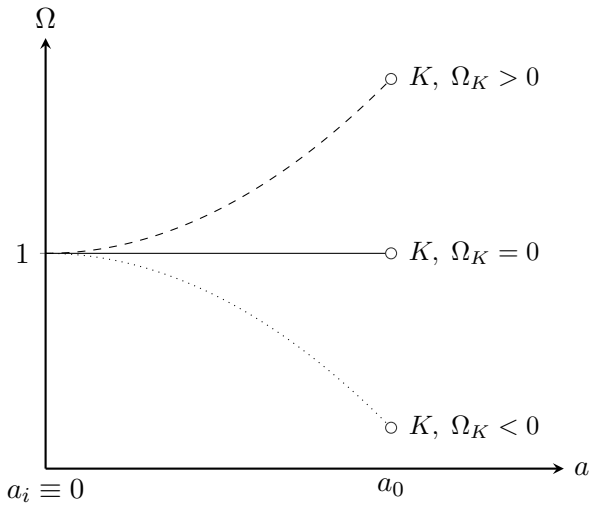


Figure 2.2: Visualization of the flatness problem in conventional Big Bang theory. If the universe started perfectly flat at early times, corresponding to $K = \Omega_K = 0$ and $\Omega = 1$, Ω would not evolve with time or the scale factor and the observed value $\Omega(a_0) = 1$ would be a logical consequence. However, if K would have deviated only a tiny amount from zero at early times, Eq. 2.2.5 and Eq. **??** show that Ω would have been driven away from unity very rapidly. Therefore, unless one excepts the precise initial conditions $K = 0$ and $\Omega = 1$, the present value $\Omega(a_0) = 1$ arises a fine-tuning issue: the flatness problem.

2.3 Puzzle 2: The Horizon Problem

The second puzzle of conventional Big Bang theory is concerned with the uniformity of the Cosmic Microwave Background (CMB). Approximately 380 000 years after the Big Bang, the opaque plasma of electrons, protons and photons cooled enough for electrons and protons to form neutral hydrogen in a process called *recombination* (at redshift $z_{\rm rec} \simeq 1090$ [3][3]). From that moment on, photons were able to move freely without scattering off the electrons and protons. In other words, the universe became transparant to the photons and they are today observed as background radiation: the CMB. The temperature of this background radiation is measured to be $T_{\rm CMB} = 2.725 \pm 0.002\,{\rm K}$ [49] in every direction along the sky, with only tiny temperature fluctuations ΔT of order:

$$\frac{\Delta T}{\bar{T}_{\rm CMB}} = \mathcal{O}(10^{-5}). \tag{2.3.1}$$

However, the observed uniformity of the CMB poses a fundamental problem concerned with causality in conventional Big Bang theory. Describing this puzzle, known as the horizon problem, in a quantitative way is the main goal of this section.

Between the Big Bang singularity and recombination, only a finite conformal (and physical) time interval elapsed, namely:

$$\Delta\tau = \tau_{\rm rec} - \tau_i = \tau_{\rm rec}. \tag{2.3.2}$$

[3]In Ref. [3], the redshift value at recombination is denoted as z_* instead of $z_{\rm rec}$ and its value is listed in Table 4.

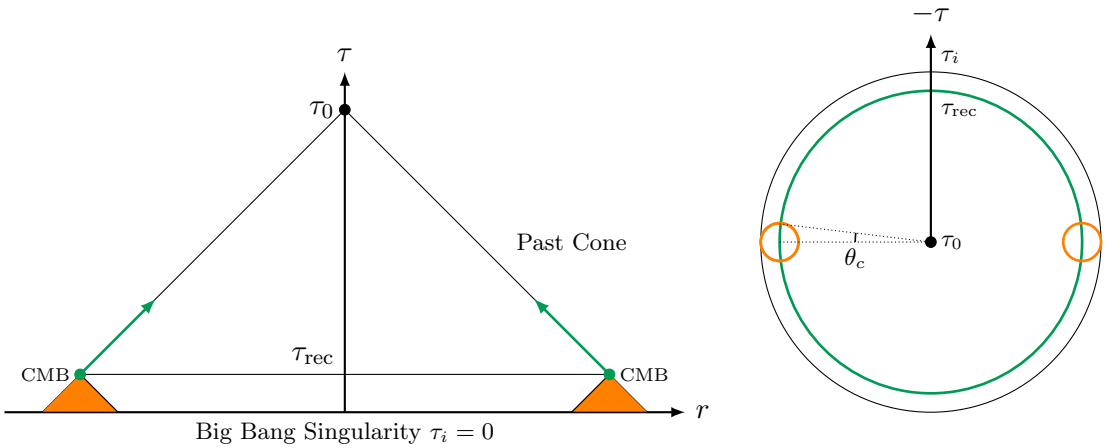


Figure 2.3: Diagrammatic visualization of the horizon problem. Left: conformal diagram for the CMB in conventional Big Bang theory. For the two considered points (green dots), the past light cones (orange) do not intersect since they are terminated by the Big Bang singularity at $\tau_i = 0$. Why have both points on the CMB the same temperature to very high degree, while they apparently have never been in causal contact? Right: alternative representation of the horizon problem used in the quantitative analysis of the horizon problem.

The finiteness of this time interval $\Delta\tau$ yields a serious problem concerned with causality: it implies that most spots of the CMB have non-overlapping past light cones. This is illustrated with the conformal diagram in Fig. 2.3 (left). The two green dots represent two points on the CMB, which may be considered as opposite points on the sky. As shown using the orange triangles, the past light cones of these two points do not overlap and hence have never been in causal contact (by definition of the light cone). Put differently, the past light cones are terminated by the Big Bang singularity at $\tau_i = 0$, which prevents any causal correlation between the two points.

Now, the problem is the aforementioned very uniform temperature of the CMB. For the different points in Fig. 2.3 (left) to obtain the same temperature, they must have been in causal contact, which is forbidden. Therefore, the fact that the observed uniform temperature of the CMB cannot be explained by conventional Big Bang theory is a fundamental drawback of the theory and is usually referred to as the *horizon* problem.

Quantitatively, analysis of the horizon problem in conventional Big Bang theory implies that there should be vanishing (temperature) correlations in the CMB for angles:

$$\theta \equiv 2\theta_c \gtrsim 2^\circ, \tag{2.3.3}$$

equivalent to approximately $\theta = 0.034$ radians. Hence causal patches in the CMB cover 0.0012 steradians (sr) according to conventional Big Bang theory. Since a unit sphere subtends 4π steradians, the CMB should therefore consist of $4\pi/0.0012 \simeq 10^4$ causally disconnected patches. This statement, which will be derived explicitly in Appendix C.1, is in violation with the observed uniform temperature of the CMB.

2.4 Inflation: a Decreasing Hubble Sphere

In the preceding two sections, two fundamental puzzles of conventional Big Bang theory are introduced and explained: the flatness and horizon problem. In both cases, the core of the problem relies on the fact that the comoving Hubble radius increases over time:

$$\frac{d(aH)^{-1}}{dt} > 0, \tag{2.4.1}$$

in case the strong energy condition (SEC) $w > -1/3$ is satisfied. The idea of cosmological inflation solves these problems by proposing an era in the early universe in which $(aH)^{-1}$ decreases over time:

$$\boxed{\frac{d(aH)^{-1}}{dt} < 0.} \tag{2.4.2}$$

That is, one may *define* inflation as an cosmological era in which the above equation holds. In the remaining part of this section, it will be shown how a decreasing Hubble sphere solves the flatness and horizon problem. The next section will relate the shrinking Hubble sphere as a definition of inflation (Eq. 2.4.2) to other popular definitions of inflation, such as accelerated expansion.

2.4.1 Solution to the Flatness Problem

Inflation solves the flatness problem in a fairly simple though elegant way. In conventional Big Bang theory, the increasing Hubble sphere drives the universe away from flatness ($\Omega = 1$ and $K = 0$). This was explained in detail using Eq. 2.2.3:

$$\Omega(a) - 1 = \frac{K}{(aH)^2}, \tag{2.4.3}$$

because $(aH)^{-1}$ increases, any minor deviation from spatial flatness ($\Omega = 1$) at early times is amplified in course of time. However, during inflation the opposite occurs. Since $(aH)^{-1}$ decreases over time during inflation, the difference between Ω and unity becomes smaller and smaller as time evolves. If inflation lasts long enough, it can drive even a highly curved early universe towards flatness.

2.4.2 Solution to the Horizon Problem

To see how inflation solves the horizon problem, recall the relation between the Hubble sphere $(aH)^{-1}$ and the scale factor (Eq. 2.1.10):

$$(aH)^{-1} = H_0^{-1} a^{(1+3w)/2}. \tag{2.4.4}$$

Because the Hubble sphere shrinks during inflation and the scale factor is an increasing function of time, the power of the scale factor $(1+3w)/2$ must be smaller than zero. In other words, inflation requires a violation of the SEC ($w > -1/3$):

$$\frac{d(aH)^{-1}}{dt} < 0 \qquad \longrightarrow \qquad w < -1/3. \tag{2.4.5}$$

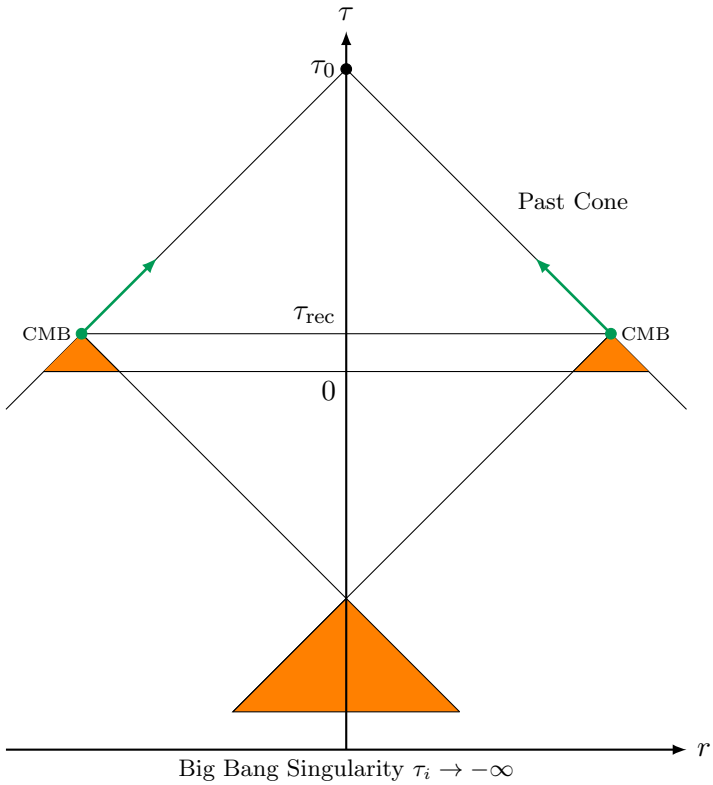


Figure 2.4: Resolution of the horizon problem by inflation. Due to the shrinking Hubble sphere, the Big Bang singularity is pushed towards $\tau_i \to -\infty$. Hence, there is enough conformal time between the singularity and recombination and the apparently disconnected CMB-patches have been in causal contact: their equal temperatures are no longer a puzzle but a logical consequence of inflation.

Therefore, the conformal time value τ_i of the singularity $a_i \to 0$ is not zero as in Eq. 2.1.14, but instead:

$$\tau_i = \frac{2H_0^{-1}}{1+3w}\, a_i^{(1+3w)/2} \xrightarrow{a_i \to 0,\ w<-1/3} -\infty. \tag{2.4.6}$$

That is, due to the shrinking Hubble sphere the conformal time value of the singularity τ_i is pushed towards negative infinity. In other words, when introducing inflation, there is *more* conformal time between the Big Bang singularity and recombination. In fact, the time interval $\Delta\tau$ between these two events is:

$$\Delta\tau = \tau_{\rm rec} - \tau_i \longrightarrow +\infty. \tag{2.4.7}$$

Because of the infinite 'sea' of conformal time between the singularity and recombination, the light cones that were terminated in Fig. 2.3 (left) by the singularity at $\tau_i = 0$ can now easily overlap. Hence, the 10^4 apparently disconnected patches of the CMB actually have been in causal contact and the fact that they have the same temperature is no surprise anymore: the horizon problem is resolved. The resolution of the horizon problem is also shown in the conformal diagram of Fig. 2.4.

A natural question to ask is the amount of inflation that is required to solve the horizon and flatness problems. The amount of inflation occured between time t_1 and t_2 can be

measured by the number of e-foldings of inflationary expansion, which is defined as:

$$N(t_1, t_2) = \int_{t_1}^{t_2} dt\, H(t). \tag{2.4.8}$$

Assuming H to be almost constant and the scale factor to scale as $a = e^{Ht}$, we can compute N as:

$$N \equiv \ln\left[\frac{a(t_2)}{a(t_1)}\right]. \tag{2.4.9}$$

In order to solve the horizon and flatness problems, it is estimated that at least 50 to 60 e-foldings of inflationary expansion are required [14, 15].

2.4.3 Other Definitions of Inflation

In the above, inflation is defined as period in the early universe characterized by a shrinking Hubble sphere $(aH)^{-1}$ (Eq. 2.4.2). This definition of inflation is presented since it relates most easily to the horizon and flatness problems. Nevertheless, there are other, equivalent, ways to define a period of inflation. Perhaps the most well-known definition of inflation is a period of accelerated expansion ($\ddot{a} > 0$), as quantified by an almost exponentially increasing scale factor:

$$a(t) \propto e^{Ht}, \tag{2.4.10}$$

with the Hubble parameter only varying slowly with time: $\dot{H} \simeq 0$. The required mild time-dependence of H can be quantified by a quantity called the first Hubble flow parameter:

$$\varepsilon \equiv -\frac{\dot{H}}{H^2} \ll 1, \tag{2.4.11}$$

which must be much smaller than unity in order for H to be almost constant over time.

However, in order for inflation to solve the horizon and flatness problem it should persist for a sufficiently long period of time. Mathematically, this implies that ε changes only slowly with time. In terms of e-folds, this additional condition is captured by the requirement that the fractional change $d\ln\varepsilon$ changes only slowly with the number of e-folds dH. The parameter η is used to formalise this second condition and is defined as:

$$\eta \equiv \frac{\dot{\varepsilon}}{H\varepsilon}. \tag{2.4.12}$$

The inequality $|\eta| < 1$ then assures that the fractional change of ε with N is small. In conclusion, the conditions for inflation are characterised by:

$$\varepsilon, |\eta| < 1, \tag{2.4.13}$$

where the first and second parameter assure inflation occurs and lasts sufficiently long, respectively. These parameters ε and $|\eta|$ are often referred to as the *Hubble flow parameters*.

As a De-Sitter space (section 1.3.6) is quantified by a constant Hubble parameter (limit $\varepsilon \to 0$), the space-time during inflation may also be regarded as a quasi De-Sitter space-time. In Appendix C.2, we will show the equivalence between various definitions of inflation.

2.4.4 Sourcing Inflation

As described in Appendix C.2, a less trivial (but equally valid) definition of inflation is that the fluid driving inflation must violate the strong energy condition (SEC):

$$1 + 3w > 0. \tag{2.4.14}$$

That is, inflation requires a fluid with equation of state $w < -1/3$ and hence a negative pressure.

A fluid that violates the SEC and that can attain a state of negative pressure is a scalar field, usually denoted by ϕ. Therefore, inflation is often modelled using a scalar field, called the inflaton, which is characterized by a potential function $V(\phi)$. To be more precise, inflationary models using only one scalar field are referred to as *single field inflation*.[4] It is assumed that, during inflation, the energy density in the universe is dominated by the inflaton contribution ρ_ϕ and hence that the effect of other possible energy sources is completely negligible. This is essentially a manifestation of the single-component universe condition (see section 1.3.6); the component being the inflaton in this case.

The central system that will be studied is a single scalar field – the inflaton – minimally coupled to Einstein gravity. The action governing this system consists of two contributions: the Einstein-Hilbert term S_{EH}:

$$S_{\text{EH}} = \frac{M_{\text{pl}}^2}{2} \int d^4x \sqrt{-g} R. \tag{2.4.15}$$

and the scalar field action S_ϕ, corresponding to the scalar field Lagrangian $\mathcal{L}_\phi$ which is given by:

$$\mathcal{L}_\phi = -\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi). \tag{2.4.16}$$

Therefore, the action describing a scalar field weakly coupled to gravity reads:

$$\boxed{S = S_{\text{EH}} + S_\phi = \int d^4x \sqrt{-g} \left[\frac{1}{2} M_{\text{pl}}^2 R - \frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) \right].} \tag{2.4.17}$$

The fact that this action describes the inflaton *weakly* coupled to gravity refers to the fact that there is no direct coupling between the Ricci scalar R and ϕ: the only coupling between the inflaton and gravity is via the metric determinant $\sqrt{-g}$. To describe this system within a flat FRW space-time, one simply uses the (flat) FRW metric for $g_{\mu\nu}$, as provided by Eq. 1.2.5.

In the remaining part of this chapter, the mechanism by which a scalar field can generate a period of inflation will discussed in more detail. To this end, the approach will be as follows. First, we formally introduce the system that will be studied by means of the corresponding action below. Subsequently, this action will be used in section 2.5 to derive the equation of motion of the scalar field, called the Klein-Gordon equation, within a FRW space-time. In the last section, the so-called slow roll approximation will be introduced, which simplifies the equations of motion.

[4] Inflation may also be driven by multiple scalar fields (dubbed *multi-field inflation*); this possibility will be considered later on in this work.

2.5 Klein-Gordon Equation

To determine the time evolution of the inflaton, the corresponding equation of motion for ϕ should be derived. Since the action of the system is known (Eq. 2.4.17) and the inflaton is assumed to couple minimally to gravity, the derivation of the equation of motion using the principle of least action is relatively straightforward.

The action corresponding to the scalar field S_ϕ reads:

$$S_\phi = \int d^4x\sqrt{-g}\mathcal{L}_\phi = \int d^4x\sqrt{-g}\left[-\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right]. \tag{2.5.1}$$

Applying the principle of least action $\delta S_\phi \equiv 0$ yields the generic Klein-Gordon equation governing the dynamics of the scalar field on a given metric $g_{\mu\nu}$:

$$\Box\phi \equiv V_\phi(\phi), \tag{2.5.2}$$

where $V_\phi(\phi) = dV/d\phi$. The term $\Box\phi$ denotes the d'Alembertian operator acting on the scalar field, it is defined in terms of $g_{\mu\nu}$ and ϕ as:

$$\Box\phi \equiv \frac{1}{\sqrt{-g}}\partial_\nu\left(\sqrt{-g}g^{\mu\nu}\partial_\mu\phi\right). \tag{2.5.3}$$

Klein-Gordon Equation in FRW Space-Time

Now the generic Klein-Gordon equation can be applied to a flat FRW space-time by specifying $g_{\mu\nu}$ to Eq. 1.2.5:

$$g_{\mu\nu} = \begin{bmatrix} -1 & 0 \\ 0 & a^2\delta_{ij} \end{bmatrix}. \tag{2.5.4}$$

Consider the $\mu = \nu = 0$ component of the generic KG equation (Eq. 2.5.2), which gives:

$$\frac{1}{\sqrt{-g}}\partial_0\left(\sqrt{-g}g^{00}\partial_0\phi\right) = V_\phi(\phi) \tag{2.5.5}$$

Using $g^{00} = -1$ and $\sqrt{-g} = a^3$, where $g = \det(g_{\mu\nu})$, the equation becomes:

$$-\frac{1}{a^3}\partial_0(a^3\dot{\phi}) = V_\phi(\phi). \tag{2.5.6}$$

Performing the time derivative and moving all terms to the left hand side yields the Klein-Gordon equation in a FRW space-time:

$$\boxed{\ddot{\phi} + 3H\phi + V_\phi(\phi) = 0.} \tag{2.5.7}$$

Note how the expansion of the FRW space-time yields a friction term $3H\dot{\phi}$ in the equation of motion.

2.6 Friedmann Equations during Inflation

To determine the evolution of the FRW universe under the presence of the scalar field, the Friedmann equations will constructed for a universe in which the energy density and pressure are dominated by the scalar field. In other words, the Friedmann equations for a single-component universe – the component being the inflaton – will be derived. To achieve this, first the pressure P_ϕ and energy density ρ_ϕ of the scalar field inferred from the scalar field energy-momentum tensor $T^{(\phi)}_{\mu\nu}$, which is given by:[5]

$$T^{(\phi)}_{\mu\nu} = \partial_\mu\phi\partial_\nu\phi - g_{\mu\nu}\left[\frac{1}{2}g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi + V(\phi)\right]. \tag{2.6.2}$$

From the above result, the energy density and pressure can inferred as the purely temporal components and the diagonal spatial components. To derive the explicit expression for ρ_ϕ and P_ϕ, we use the fact that $g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi = -\dot{\phi}^2$ since $\partial_i\phi = 0$ as required by homogeneity and isotropy. Hence, the energy-momentum tensor for the scalar field may also be written as:

$$T^{(\phi)}_{\mu\nu} = \partial_\mu\phi\partial_\nu\phi - g_{\mu\nu}\left[-\dot{\phi}^2/2 + V(\phi)\right]. \tag{2.6.3}$$

Using the above expression, one can easily confirm that ρ_ϕ and P_ϕ are given by:

$$\rho_\phi = \frac{\dot{\phi}^2}{2} + V(\phi), \qquad P_\phi = \frac{\dot{\phi}^2}{2} - V(\phi), \tag{2.6.4}$$

these equations relate the microscopic structure of the system (the inflaton) to the macroscopic quantities, i.e. the energy density and pressure. The equation of state $w_\phi = P_\phi/\rho_\phi$ is evidently:

$$\boxed{w_\phi = \frac{\dot{\phi}^2/2 - V(\phi)}{\dot{\phi}^2/2 + V(\phi)}.} \tag{2.6.5}$$

Now that ρ_ϕ and P_ϕ are known, it is relatively straightforward to construct the Friedmann equations for a universe whose energy density is dominated by the scalar field:

$$H^2(\phi) = \frac{\rho_\phi}{3M_{\rm pl}^2} = \frac{1}{3M_{\rm pl}^2}\left[\dot{\phi}^2/2 + V(\phi)\right]. \tag{2.6.6}$$

Together with the Klein-Gordon equation (Eq. 2.5.7) this equation forms the first principles starting point for studying single-field inflation. The first Hubble flow parameter ε can be written explicitly as:

$$\boxed{\varepsilon = \frac{\dot{\phi}^2}{2M_{\rm pl}^2 H^2},} \tag{2.6.7}$$

by using the time derivative of the Friedmann equation, the continuity equation (Eq. 1.3.6) and the fact that $\rho_\phi + P_\phi = \dot{\phi}^2$.

[5] Given the scalar field Lagrangian $\mathcal{L}_\phi$ (Eq. 2.4.16), the corresponding energy-momentum tensor can be computed using the relation between $T_{\mu\nu}$ and the matter Lagrangian $\mathcal{L}_{\rm M}$, which is given by:

$$T_{\mu\nu} = -2\frac{\partial\mathcal{L}_{\rm M}}{\partial g^{\mu\nu}} + g_{\mu\nu}\mathcal{L}_{\rm M}. \tag{2.6.1}$$

Setting $\mathcal{L}_{\rm M} \equiv \mathcal{L}_\phi$ yields directly the energy-momentum tensor for the inflaton scalar field.

Since $\varepsilon < 1$ for inflation to occur, the kinetic energy term $\dot{\phi}^2/2$ may only contribute to small extent to the total energy density $\rho_\phi = 3M_{\rm pl}^2 H^2$. Furthermore, for inflation to last long enough, it is required that ε changes slowly with time, as captured by the second parameter η. To derive η as function of the field, it is convenient to introduce a dimensionless field acceleration parameter δ, as measured per Hubble time:

$$\delta \equiv -\frac{\ddot{\phi}}{H\dot{\phi}}. \tag{2.6.8}$$

Recall that the second parameter η is defined as (Eq. 2.4.12):

$$\eta \equiv \frac{\dot{\varepsilon}}{H\varepsilon}. \tag{2.6.9}$$

Using the above expressions for δ and ε in terms of the inflaton field and its first and second time derivative, the paramter η can be written as:

$$\eta = 2(\varepsilon - \delta). \tag{2.6.10}$$

This expression can be obtained by taking the first time derivative of ε as an explicit function of time, and deviding the resulting expression for $\dot{\varepsilon}$ by $H\varepsilon$. Note that if both ε and $|\delta|$ are small compared to unity, then so is $|\eta|$ as well on account of the above relation, that is:

$$\{\varepsilon, |\delta|\} \ll 1 \quad \longrightarrow \quad \{\varepsilon, |\eta|\} \ll 1. \tag{2.6.11}$$

2.7 Slow-Roll Approximation

The above analysis is exact: no approximations are made to derive the result that in the parameter regime $\{\varepsilon, |\delta|\} \ll 1$ inflation will occur and persists. The conditions on ε and δ can be used to simplify the inflationary equation of motion in an approach which is called the Slow-Roll approximation. Using this approximation, the conditions on ε and δ can also be related directly to the properties that the inflaton potential $V(\phi)$ should possess for causing a successive period of inflation.

The condition on the first parameter ε relates to the contribution of the kinetic energy of the scalar field $\dot{\phi}^2/2$ to the total energy density $\rho_\phi = 3M_{\rm pl}^2 H^2$:

$$\varepsilon = \frac{\dot{\phi}^2}{2M_{\rm pl}^2 H^2} = \frac{3\dot{\phi}^2}{2\rho_\phi} \ll 1. \tag{2.7.1}$$

This translates to $\dot{\phi}^2/2 \ll V(\phi)$ (hence the terminology *Slow Roll*), so that the energy density can be approximated as $\rho_\phi = V(\phi)$ and the first Friedmann equations reads:

$$H^2 = \frac{V(\phi)}{3M_{\rm pl}^2}. \tag{2.7.2}$$

Note that in the SR approximation the equation of state for the scalar field w_ϕ indeed satisfies the constraint $w_\phi < -1/3$. In fact, the inflationary equation of state approaches that of a true cosmological constant w_Λ under the approximation $\dot{\phi}^2/2 \ll V(\phi)$:

$$\boxed{w_\phi = \frac{\dot{\phi}^2/2 - V(\phi)}{\dot{\phi}^2/2 + V(\phi)} \quad \xrightarrow{\dot{\phi}^2/2 \ll V(\phi)} \quad -1 = w_\Lambda.} \tag{2.7.3}$$

So in the SR approximation, the fact that the fluid driving inflation resembles a cosmological constant manifests explicitly.

The condition $|\delta| \ll 1$ allows to simplify the Klein Gordon equation and neglect the acceleration term $\ddot{\phi}$ of the field:

$$\ddot{\phi} + 3H\dot{\phi} + V_\phi \simeq 3H\dot{\phi} + V_\phi = 0. \tag{2.7.4}$$

Using these forms of the Friedmann and Klein Gordon equation in the Slow Roll approximation, the parameter ε can be written as:

$$\boxed{\varepsilon = \frac{\dot{\phi}^2}{2M_{\rm pl}^2 H^2} \simeq \frac{M_{\rm pl}^2}{2}\left(\frac{V_\phi}{V}\right)^2 \equiv \epsilon_{\rm v}.} \tag{2.7.5}$$

The time derivative of the simplified Klein Gordon equation provides an expression for the field acceleration $\ddot{\phi}$:

$$3\dot{H}\dot{\phi} + 3H\ddot{\phi} + V_{\phi\phi}\dot{\phi} = 0, \tag{2.7.6}$$

$$\ddot{\phi} = -\frac{\dot{H}}{H}\dot{\phi} - \frac{V_{\phi\phi}\dot{\phi}}{3H}. \tag{2.7.7}$$

The dimensionless acceleration parameter δ then reads:

$$\delta = -\frac{\ddot{\phi}}{H\dot{\phi}} = \frac{\dot{H}}{H^2} + \frac{V_{\phi\phi}}{3H^2} = -\frac{\dot{\phi}^2}{2M_{\rm pl}^2 H^2} + M_{\rm pl}^2\frac{V_{\phi\phi}}{V} = M_{\rm pl}^2\frac{V_{\phi\phi}}{V} - \varepsilon \tag{2.7.8}$$

Now, the sum of $\delta + \varepsilon$ in the Slow Roll approximation is denoted by $\eta_{\rm v}$ and reads:

$$\boxed{\eta_{\rm v} \equiv \delta + \varepsilon \simeq M_{\rm pl}^2\frac{V_{\phi\phi}}{V}.} \tag{2.7.9}$$

In contrast to the Hubble flow parameters ε and η, the SR parameters $\epsilon_{\rm v}$ and $\eta_{\rm v}$ are called potential slow roll parameters. The two sets of parameters are related during SR inflation via:

$$\varepsilon \simeq \epsilon_{\rm v}, \qquad \eta_{\rm v} \simeq 2\varepsilon - \frac{1}{2}\eta. \tag{2.7.10}$$

Based on these relations above, the inflation conditions $\{\varepsilon, |\eta|\} \ll 1$ are equivalent to $\{\epsilon_{\rm v}, |\eta_{\rm v}|\} \ll 1$ in the SR approximation. In terms of the shape of the inflaton potential, the conditions on $\epsilon_{\rm v}$ and $\eta_{\rm v}$ imply that the potential curve should be very flat, as required by $\epsilon_{\rm v} \ll 1$, for a sufficiently large field range, as enforced by $\eta_{\rm v} \ll 1$.

Figure 2.5 shows a typical inflaton potential $V(\phi)$, possessing the required features to induce a period of inflation. At small field values ϕ, the potential is very flat for a sufficiently large ϕ-range and the field, denoted by the orange dot on the potential curve, can slowly roll down the potential. Hence, the grey regime in the plot satisfies the slow roll condition are satisfied. The field value at the end of inflation $\phi_{\rm end}$ is defined via the exact condition $\varepsilon(\phi_{\rm end}) \equiv 1$, or in the SR approximation by $\epsilon_{\rm v} \equiv 1$. As the field departures from the inflation regime it starts to explore parts of the potential where $\varepsilon > 1$. In particular, the field accelerates towards the bottom of the potential, where it will start to oscillate and it releases all of its potential energy. From this excess energy, the particles of the standard model (SM) are believed to be formed in a process called *reheating*. That is, the inflaton field decays into the degrees of freedom of the SM: this defines the onset of conventional Big Bang theory.

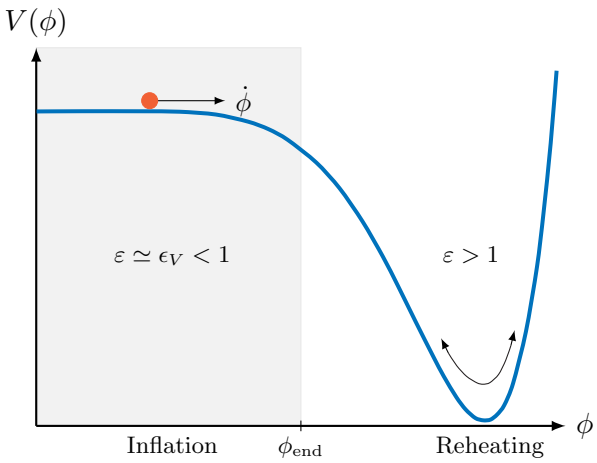


Figure 2.5: Plot of a typical inflaton potential $V(\phi)$. The regimes corresponding to inflation and reheating are indicated as well. See the main text for a detailed description.

Part II

Quantum Origin of Structure and Cosmological Perturbations

Chapter 3

From Quantum Fluctuations to Large Scale Structure and CMB Anisotropies

"It was incredibly brave in the early 1980's, when the initial formulation of inflation was made, to apply the most advanced theories from particle physics to the early universe. The results, as you will see, are beautiful."

— Leonardo Senatore [79]

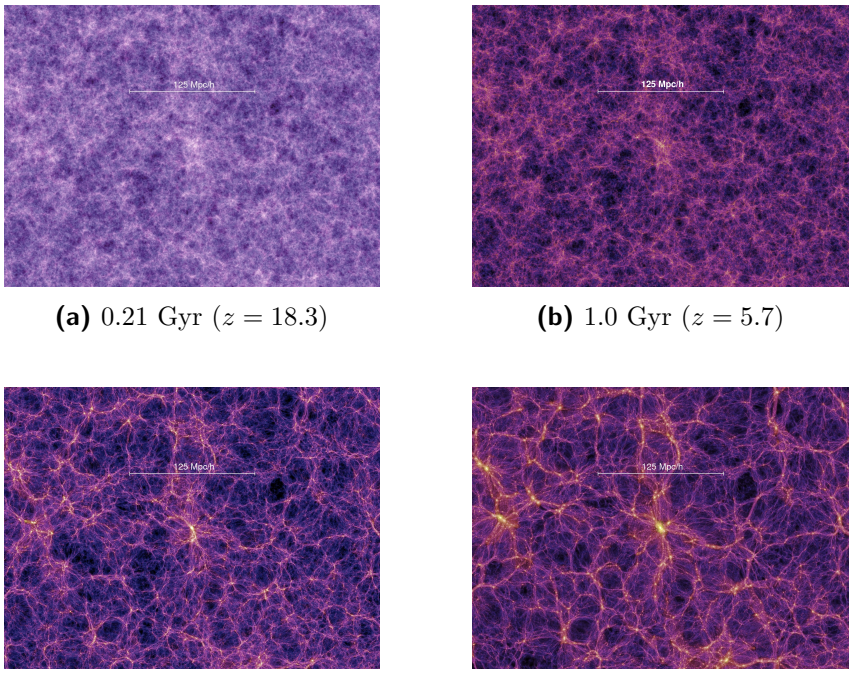
In the preceding chapter, we concluded that inflation drives the universe into a state of very high homogeneity and isotropy. Yet, the universe is filled with structures such as clusters and superclusters of galaxies – the very fact that we are here proves that there must have been a mechanism at play that generated the structure in the universe. If inflation drove the universe to homogeneous and isotropic state, how are the structures we observe today generated?

The current understanding of the mechanism that is responsible for the generation of structure in the universe relies on *gravitational instability*. According to gravitational instability, small density inhomogeneities in the early universe grew under the influence of gravity into the structures observed today. Due to the presence of fluctuations in the density field in the early universe, some regions in the universe contained more energy and are *overdense*, whereas other regions contain less energy and are hence *underdense*. Since the gravitational field is sourced by the present energy and matter, it will be stronger in overdense regions and weaker in underdense regions. Therefore, overdense regions will attract more matter under the influence of the enhanced gravitational field and matter will eventually accumulate into structures. According to this simply but yet powerful mechanism, all observed structures in the universe are believed to be formed. Fig. 3.1 shows the generation of structure over time under the influence of gravitational instability in a simulated universe.

Although the theory of structure formation by means of gravitational instability is intuitive and powerful, it leaves one fundamental question unanswered: what caused the primordial density fluctuations in the universe that acted as the seeds for structure formation? Before the theory of inflation was formulated, there was no explanation available about the origin of primordial density perturbations in the early universe. Shortly after the theory of inflation was invented, however, it was discovered that the primordial seeds required for structure formation could be produced during the inflationary stage.

We will now discuss qualitatively how inflation can generate the primordial seeds for structure formation.[1] In the preceding chapters, the inflaton field is studied in the classical

[1] For convenience, we will restrict ourselves here, and in the next part, to the single-field scenario. However,



(a) 0.21 Gyr ($z = 18.3$) **(b)** 1.0 Gyr ($z = 5.7$)

(c) 4.7 Gyr ($z = 1.4$) **(d)** 13.6 Gyr ($z = 0$)

Figure 3.1: Evolution of structure in a simulated universe over time (or redshift) by means of gravitational instability. Notice that as time evolves, the matter accumelates into clear structures and the universe departs from the homogeneous state at early times. The simulation is performed by the VIRGO Consortium as part of the *Millenium Run Project*.

framework of General Relativity. However, studying inflation in the framework of quantum field theory (QFT) will show that the inflaton field is subject to so-called quantum fluctuations. These fluctuations are small temporary changes in the energy of the field at different locations. They are a direct consequence of the Heisenberg uncertainty principle. Mathematically, the fluctuations can be described as position-dependent first order perturbations around the background field, so that the inflaton field can be expressed as:

$$\phi(\boldsymbol{x}, t) = \bar{\phi}(t) + \delta\phi(\boldsymbol{x}, t), \tag{3.0.1}$$

where the background and fluctuation are written as $\bar{\phi}(t)$ and $\delta\phi(\boldsymbol{x}, t)$, respectively. Notice that the background field adheres the cosmological principle and therefore has only temporal dependence, whereas the fluctuation is position dependent.

Quantum fluctuations are not solely present in the inflaton field. Instead, all fields (e.g. the Higgs field) exhibit these fluctuations. However, the consequences of these fluctuations are rather different in the case of the inflaton compared to a field in a static (Minkowski) background. Essentially, the inflaton can be regarded as a local *clock* measuring the amount of inflationary expansion still to come in a specific patch of the universe. More specifically, at an arbitrary time t, the field space seperation:

$$\Delta\phi \equiv |\phi(t) - \phi_{\text{end}}|, \tag{3.0.2}$$

the primordial seeds could be generated equally well in the multi-field scenario, which will be discussed later on in this work.

between the field value $\phi(t)$ and the value $\phi_{\rm end}$ for which $\varepsilon \equiv 1$, corresponding to the end of inflation, measures the amount of inflation still to occur (see also Fig. 2.5). However, due to the position-dependent quantum fluctuations $\delta\phi(\boldsymbol{x}, t)$, the inflaton field will be at slightly different locations on the potential in different patches of the universe at an arbitrary time t. As a result, the amount of inflation still to occur in different patches of the early universe different as well. In other words, some patches of the universe will inflate for a slightly longer time than others.[2] These subtle difference in the expansion histories of different patches causes density fluctuations after inflation. At later times, specifically during the radiation and matter dominated era's, the density perturbations generated during inflation precisely act as the primordial seeds needed for structure formation. [3]

In addition to the generation of primordial seeds, the quantum fluctuations during inflation also have an important effect the cosmic microwave background (CMB). The generated density fluctuations cause minute temperature fluctuations in the CMB of the order:

$$\frac{\delta T}{\bar{T}} = \mathcal{O}(10^{-5}), \tag{3.0.3}$$

around the background temperature $\bar{T}$. These temperature fluctuations are still observed today and open an observational window to the early universe. In particular, observations of the CMB temperature anisotropies can be used to constrain inflationary models.

In this chapter, it will be explained globally how quantum fluctuations during inflation can generate the primordial seeds for structure formation and CMB anisotropies, leaving the detailed computations for the subsequent chapters. Furthermore, we will develop the statistical tools needed to connect theoretical predictions to cosmological observations. Lastly, the concepts introduced in this chapter will be illustrated in a toy model.

3.1 The Big Picture

Before going into mathematical details in the remaining part of this chapter and the subsequent chapters, the main mechanism or big picture will be described here in a more quantitative way than in the introduction. The description of the big picture is based on the diagram shown in Fig. 3.2. First, we will discuss the physics behind quantum fluctuations during inflation. Subsequently, it will be explained how these quantum fluctuations can cause both the temperature fluctuations in the CMB and the primordial seeds needed for large-scale structure formation.

3.1.1 Quantum Fluctuations during Inflation

For reasons explained extensively in section 4.2, the quantum fluctuation in the inflaton $\delta\phi(\boldsymbol{x}, t)$ is written in terms of Fourier modes:

$$\boxed{\delta\phi(\boldsymbol{x}, t) = \int \frac{d^3k}{(2\pi)^3} \delta\phi_{\boldsymbol{k}}(t)\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}},} \tag{3.1.1}$$

[2] In other words, relative to the background evolution, some patches of the universe are slightly ahead and others will be slightly behind, as induced by a local shift in time $\delta t(t, \boldsymbol{x})$. As shown in section 4.6, the ability to express perturbations in terms of common time shift is closely related to the fact that single-field inflation produces adiabatic perturbations.

[3] Later on, it was shown that quantum fluctuations in an additional field, called the *curvaton*, could also cause the primordial seeds. This possibility is called the *curvaton scenario* [63], but it will not be considered in this work.

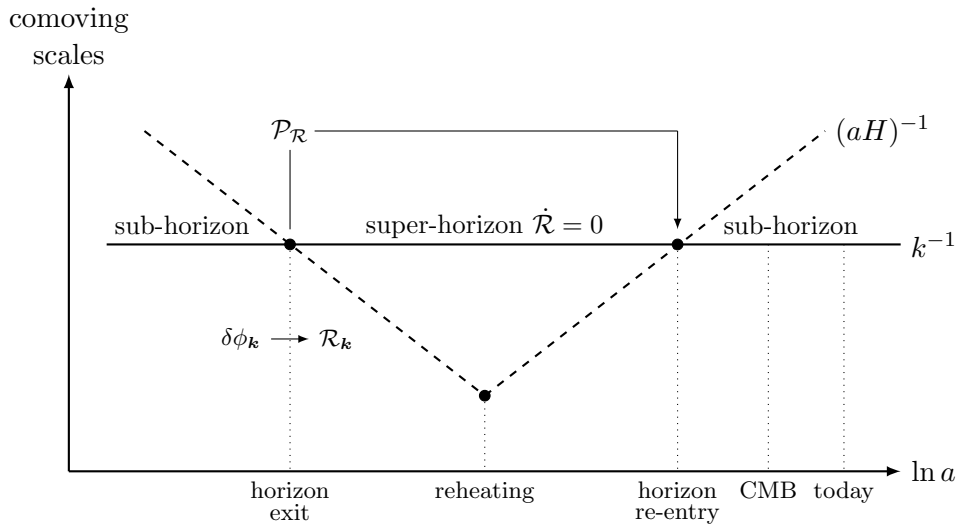


Figure 3.2: Schematic overview of the dynamics of fluctuation modes during inflation and beyond. At very early times, all modes of cosmological interest are far inside the horizon and are hence called *sub-horizon.* Since during inflation the comoving Hubble sphere $(aH)^{-1}$ schrinks the modes eventually becomes super-horizon, at which its evolution freezes. During the subsequent stages of conventional Big Bang evolution, the Hubble sphere increases agian and the modes re-enter the horizon again.

where $\delta\phi_{\boldsymbol{k}}(t)$ is the Fourier mode of the inflaton fluctuation: $\boldsymbol{k}$ and $k = |\boldsymbol{k}|$ are the wave vector and wavenumber of the considered mode, respectively. Note that since we work in natural units $\boldsymbol{k}$ is equal to 3-momentum as well. The wavenumber k is a comoving quantity the corresponding characteristic comoving length scale is given by k^{-1}.

The average of the quantum fluctuations vanishes due to their random nature: $\langle \delta\phi_{\boldsymbol{k}} \rangle = 0$. However, the variance of the fluctuations does not vanish:

$$\langle |\delta\phi_{\boldsymbol{k}}|^2 \rangle \equiv \langle 0 | |\delta\phi_{\boldsymbol{k}}|^2 | 0 \rangle \neq 0, \tag{3.1.2}$$

where $|0\rangle$ refers to the quantum vacuum state and hence $\langle 0 | |\delta\phi_{\boldsymbol{k}}|^2 | 0 \rangle$ may also be interpreted as quantum mechanical expectation value.[4] Therefore, the variance can be used to describe the quantum fluctuations in a statistical way. Subsequently, the statistical description of the fluctuations can be used to compare theory with observations: this is explained in section 3.2. As described in that section, the essential link between theory and observation is provided by the power spectrum, which is defined in Fourier space as:

$$\boxed{\mathcal{P}_{\delta\phi} = \frac{k^3}{2\pi^2} |\delta\phi_{\boldsymbol{k}}|^2.} \tag{3.1.3}$$

In this case, the power spectrum corresponds to the inflaton fluctuation $\delta\phi$. Nevertheless, it should be mentioned that the power spectrum does not solely apply to $\delta\phi$ but can be constructed for any perturbation in a similar way. For instance, later in this section, the power spectrum of the so-called comoving curvature perturbation $\mathcal{R}$ will be considered. Upon the replacement of $\delta\phi$ by $\mathcal{R}$, the power spectrum for $\mathcal{R}$ is equivalent to the one given above.

[4] More formally, $\langle \delta\phi_{\boldsymbol{k}} \rangle$ and $\langle |\delta\phi_{\boldsymbol{k}}|^2 \rangle$ are called the one-point and two-point correlation function, respectively.

As mentioned before, the quantum fluctuations are not a special property of the inflaton field. Instead, all fields are subject to these fluctuations. However, the effect of these fluctuations is rather different during inflation compared to for instance the fluctuations of a field in a static Minkowski background. Due to the rapid expansion during inflation, the quantum fluctuations in the inflaton $\delta\phi$ are stretched to cosmological scales. This is not the case for the quantum fluctuations of a scalar field in a static Minkowski background, as the background does not expand.

More quantitively, the fluctuations $\delta\phi$ are stretched to scales greater than the horizon at that time, as quantified by H^{-1}. In comoving coordinates, the Fourier modes of the fluctuations have constant wavelengths k^{-1}, as shown for one Fourier mode in Fig. 3.2. Since the comoving Hubble sphere $(aH)^{-1}$ decreases during inflation, the comoving wavelength k^{-1} eventually become greater than the comoving Hubble radius: this is equivalent to saying that $\delta\phi_{\boldsymbol{k}}$ is stretched to cosmological scales during inflation in physical coordinates.

To distinguish between different stages in the evolution of the modes, it will be convenient to compare the comoving wavelength k^{-1} with the Hubble radius $(aH)^{-1}$. At sufficiently early times, defined by small $\ln a$, all modes of interest were smaller than the comoving Hubble sphere. The modes are said to be sub-horizon since the comoving wavelengths of the modes satisfy $k^{-1} \ll (aH)^{-1}$. In other words, k-modes on sub-horizon scales are defined by:

$$\frac{k}{aH} = \frac{k}{\mathcal{H}} \gg 1. \qquad \text{(Sub-Horizon)} \tag{3.1.4}$$

Following the same logic, modes on super-horizon scales satisfy:

$$\frac{k}{aH} = \frac{k}{\mathcal{H}} \ll 1. \qquad \text{(Super-Horizon)} \tag{3.1.5}$$

Using the fact that during inflation conformal time can be written as $\tau = -1/aH = -1/\mathcal{H}$, we can also express the sub-horizon and super-horizon limits as $|k\tau| \gg 1$ and $|k\tau| \ll 1$, respectively.

The transition from sub- to super-horizon scales occurs when the inverse wavenumber and the comoving Hubble sphere are equal to each other, i.e. when $k = aH$. This event is referred to as horizon exit in Fig. 3.2. At horizon crossing, the inflaton fluctuation $\delta\phi_{\boldsymbol{k}}$ loses its quantum nature and the quantum expectation value $\langle 0|\delta\phi_{\boldsymbol{k}}|^2|0\rangle$ can be identified with the ensemble average of a classical stochastic field. That is, on super-horizon scales $\delta\phi$ may be regarded as a classical perturbation. The change of the fluctuations from quantum to classical nature is known as the quantum-to-classical transition. The modes then evolve on super-horizon scales before they re-enter the sub-horizon during the matter and radiation dominated era's, which are described by conventional Big Bang theory.

However, there is a serious issue concerned with super-horizon scales. On these scales, the physics is very uncertain: even the equations of motion governing the perturbations are not well-known. The uncertainty in this regime is mainly caused by the lack of precise knowledge on reheating, i.e. the transition from inflation to conventional Big Bang evolution. For most k-modes, reheating occurs when they are on super-horizon scales. Therefore, to relate the quantum fluctuations during inflation to late time observations such as the anisotropies in the CMB or fluctuations in energy density $\delta \equiv \delta\rho/\bar{\rho}$ at later times, the quantum fluctuations must be embedded in a quantity that is constant on super-horizon scales.

Fortunately, the so-called comoving curvature perturbation $\mathcal{R}$ possesses the feature of constancy on super-horizon scales. In Chapter 6 it will be proved that $\mathcal{R}$ is constant on

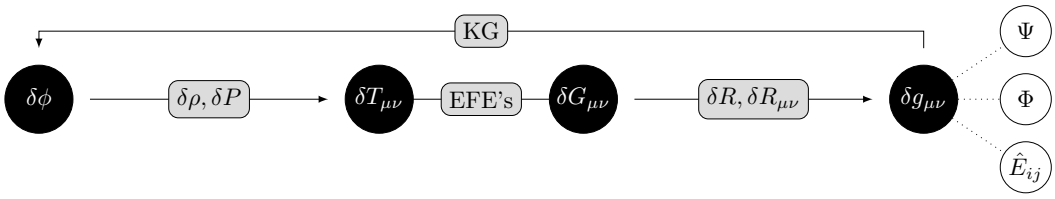


Figure 3.3: Schematic representation of the mechanism by which quantum fluctuations in the inflation $\delta\phi$ induce perturbations in the metric $\delta g_{\mu\nu}$. The backreaction of the metric perturbations on the evolution of the inflaton field through the Klein-Gorden (KG) equation is also shown.

super-horizon scales, independent of the possible unknown physics and equations governing the perturbations on these scales. At horizon exit, the switch from $\delta\phi$ to $\mathcal{R}$ will be made, as indicated in Fig. 3.2. In the Newtonian gauge[5] the comoving curvature perturbation $\mathcal{R}_{\boldsymbol{k}}$ is related to the inflaton fluctuation $\delta\phi_{\boldsymbol{k}}$ via the equation:

$$\boxed{\mathcal{R}_{\boldsymbol{k}} = H\frac{\delta\phi_{\boldsymbol{k}}}{\dot{\phi}},} \tag{3.1.6}$$

assuming single field inflation in the slow-roll approximation. On account of the above relation between $\mathcal{R}$ and $\delta\phi$ in the Newtonian gauge, the power spectra $\mathcal{P}_{\delta\phi}$ and $\mathcal{P}_{\mathcal{R}}$ are related as:

$$\boxed{\mathcal{P}_{\mathcal{R}} = \left(\frac{H}{\dot{\phi}}\right)^2 \mathcal{P}_{\delta\phi}.} \tag{3.1.7}$$

Since $\dot{\mathcal{R}} = 0$ on super-horizon scales, the power spectrum of the comoving curvature perturbation $\mathcal{P}_{\mathcal{R}}$ can be computed at horizon crossing $k = aH$ and the result can be taken directly to horizon re-entry (see Fig. 3.2). Subsequently, the power spectrum will evolve according to known laws of physics on sub-horizon scales and can be related to observables such as the temperature anisotropies in the CMB. Hence, via the comoving curvature perturbation $\mathcal{R}$, predictions made at horizon crossing, corresponding to high energies and early times, can be related to late time observables at low energies such as the CMB.

3.1.2 Metric Coupling, CMB Anisotropies and Primordial Gravitational Waves

As briefly mentioned in the introduction to this chapter and the previous section, quantum fluctuations during inflation induce small temperature anisotropies in the CMB and cause density fluctuations in the matter at later times, which act as primordial seeds needed for structure formation. In this section, we discuss the mechanism behind the generation of temperature fluctuations and primordial seeds a quantitative way.

We know that during single field slow-roll inflation, the energy density ρ and pressure P are dominated by the inflaton scalar field and for simplicity we assume that the scalar field is the only contribution to the energy content. The energy density and pressure due to the inflaton can be written as:

$$\rho_\phi = \dot{\phi}^2/2 + V(\phi), \qquad P_\phi = \dot{\phi}^2/2 - V(\phi). \tag{3.1.8}$$

[5] A gauge is a specific choice of coordinates: see section 4.3.1 for the details.

Quantum fluctuations in the inflaton $\delta\phi$ therefore induce small perturbations in energy density and pressure. More generally, the quantum fluctuations cause a perturbation in the inflaton energy-momentum tensor $T^{(\phi)}_{\mu\nu}$:

$$\delta\phi \longrightarrow \delta T^{(\phi)}_{\mu\nu}. \tag{3.1.9}$$

Then, via the EFE's, the perturbation in the energy-momentum tensor results in perturbations in the Einstein tensor $G_{\mu\nu}$:

$$\delta T_{\mu\nu} \longrightarrow \delta G_{\mu\nu} = 8\pi G\, \delta T_{\mu\nu}. \tag{3.1.10}$$

The perturbation in the Einstein tensor then finally induces a fluctuation in the metric $\delta g_{\mu\nu}$:

$$\delta G_{\mu\nu} \longrightarrow \delta g_{\mu\nu}. \tag{3.1.11}$$

Finally, the metric perturbations will cause a so-called backreaction on the evolution of the inflaton field via the Klein-Gordon equation:

$$\delta(\Box\phi) = V_\phi(\phi)\delta\phi \longrightarrow \delta\phi. \tag{3.1.12}$$

Based on the logic chain outlined above, we conclude that perturbations in the inflaton field are tightly coupled to metric perturbations and therefore they should be studied together. Figure 3.3 shows the chain explained in the text schematically. The framework we will use to study the evolution of the various perturbations is called cosmological perturbation theory (CPT) and will be introduced in detail in the next chapter.

As will be shown in section 4.2, a generic tensor (perturbation) can be decomposed into scalar, vector and tensor components. Hence, the perturbed metric can be written in terms of scalar (Φ and Ψ) and tensor ($\hat{E}_{ij}$) components. Strictly speaking, vector perturbations in the metric should be considered as well. However, vector perturbations are not always produced during inflation and even if they are, they would decay away with the expansion of the universe sufficiently fast to ignore them completely [14, 60]. The decaying behavior of vectorial perturbations is discussed more quantitavely in Appendix D.2.

The purely tensorial perturbations to the metric, denoted as $\hat{E}_{ij}$, correspond to primordial gravitational waves. In the absence of anisotropic stress, the perturbed field equations governing the evolution of the gravitational waves are unsourced.[6] In that case, the evolution equation of $\hat{E}_{ij}$ is a damped wave equation, the wave-solutions to this equation describe gravitational waves. Since these gravitational waves are believed to be generated during the inflationary state, they are called primordial gravitational waves. Over the last few years, attempts have been made to detect these primordial gravitational waves, but no direct observation of primordial waves is made. However, the cross-correlation of the gravitational waves (tensor) power spectrum with the scalar power spectrum is measured by Planck [3], giving rise to an important parameter called the tensor-to-scalar ratio r, see section 3.2.

The scalar perturbations in the metric, Φ and Ψ, set up during inflation act as a gravitational potential that couples to the energy density and matter distribution at later times,

[6] *Unsourced* means that the energy-momentum side of the perturbed EFE's is zero. This assumption is satisfied if anisotropic stress is negligible. Anisotropic stress is an additional contribution to the energy-momentum of the form Σ_{ij}, which is only present when the consider matter can not be described as a perfect fluid. Anisotropic stress is only generated by neutrinos at stages that will not concern us in this work [14]. Hence, we usually set $\Sigma_{ij} \equiv 0$. In the absence of anisotropic stress, the gravitational potentials Φ and Ψ are equal to each other.

e.g. during the radiation and matter dominated eras. However, during the reheating stage, the equations governing the evolution of the gravitational potential are not well known. In order to relate the effect of Φ and Ψ to later times, we will embed them in the curvature perturbation $\mathcal{R}$ and study its effect on the energy density and matter distributions instead.

The coupling between the comoving curvature perturbation[7] $\mathcal{R}$ and the matter induces density perturbations $\delta \equiv \delta\rho/\bar{\rho}$ in the matter. Via this mechanism, quantum fluctuations in the inflation form the primordial seeds required for structure formation. Out of the primordial density fluctuations, eventually, all observed large-scale structure, such as galaxies and clusters of galaxies, are formed.

Furthermore, the density fluctuations are also responsible for the temperature anisotropies in the CMB. In particular, density perturbations of the photons in the primordial plasma are directly related to temperature fluctuations in the CMB via [60]:

$$\frac{\delta T}{\bar{T}} = 4\frac{\delta\rho_\gamma}{\bar{\rho}_\gamma} \equiv 4\delta_\gamma, \tag{3.1.13}$$

this equation is also commonly referred to as the brightness function. As the CMB anisotropies are generated by inflationary quantum fluctuations, measuring their properties (i.e. the power spectrum) opens an observational window to constrain inflationary models (see section 3.2). This concludes our brief discussion on the mechanism by which inflation can cause CMB anisotropies and the generation of primordial seeds. The concepts outlined in this section will be worked out in detail in the next chapters.

3.2 From Theory to Observations and Back

As described in the previous section, treating inflation quantum mechanically results in a profound prediction: the quantum fluctuations in the inflaton cause anisotropies in the CMB and provide the primordial seeds for all observed structure in the universe. In order to test this prediction, comparison with observations should be made. Theoretical predictions can be connected to observations via statistical methods. In the existing literature, both the original papers as well pedagogical reviews or lecture notes, the invoked statistical concepts (such as the power spectrum) are not discussed in detail.[8] Therefore, we will discuss the statistical methods that go into inflationary calculations here in detail.

3.2.1 Singlevariate Statistics

Here we will discuss relevant concepts from singlevariate statistics, mainly to set up the notation and to provide a basis for the discussion on multivariate statistics in the next subsection. Consider a single random variable X, drawn from some probability density function (PDF) $\rho(x)$. The PDF is properly normalized such that:

$$\int \rho(x)\, dx \equiv 1, \tag{3.2.1}$$

[7] During the post-inflationary evolution of the universe, it is convenient to regard $\mathcal{R}$ itself as the gravitational potential, although this is formally not true.

[8] Exceptions to this trend are Ref. [15] (in particular the appendix A.6.) and [84].

where the integration is over the domain of x (which could be infinite). We define the expectation value for X as:

$$\mathrm{E}[X] \equiv \int x\rho(x)\, dx, \tag{3.2.2}$$

the expectation value is also referred to as the mean or average and often denoted as $\mu = \mathrm{E}[X]$. The variance of X is given by:

$$\mathrm{Var}[X] \equiv \mathrm{E}[(X-\mu)^2] = \mathrm{E}[X^2] - \mathrm{E}[X]^2, \tag{3.2.3}$$

and is also denoted as $\sigma^2 = \mathrm{Var}[X]$, where σ is the standard deviation. Equivalently, the variance may be defined as the covariance of the random variable X with itself:

$$\mathrm{Var}[X] \equiv \mathrm{Cov}(X, X). \tag{3.2.4}$$

The mean and variance can be used to characterize the considered distribution. However, typically the mean and variance do not completely specify the distribution. (Only for the Gaussian or normal distribution they do).

Cumulants.—More systematically, the distribution function can be characterized in terms of the cumulants κ_n. The cumulants are defined using the cumulant generating function $K(t)$, which is related to the moment generating function $M(t)$:

$$M(t) \equiv \mathrm{E}(e^{tX}), \tag{3.2.5}$$

via the natural logarithm:

$$K(t) = \ln M(t). \tag{3.2.6}$$

Expressing the cumulant generating function as a Maclaurin series:

$$K(t) = \sum_{n=1}^{\infty} K^{(n)}(0)\frac{t^n}{n!} = \mu t + \frac{1}{2}\sigma^2 t^2 + \mathcal{O}(t^3), \tag{3.2.7}$$

the n-th cumulant κ_n can be obtained by differentiating the above expression n times and evaluating the result at $t = 0$, that is:

$$\kappa_n \equiv K^{(n)}(t)\Big|_{t=0}. \tag{3.2.8}$$

The first and second cumulants correspond to the mean and variance, respectively.[9]

[9] The second and third cumulants ($\kappa_{2,3}$) coincide with the second and third central moments ($\mu_{2,3}$), defined via:

$$\mu_n = \mathrm{E}[(X - \mathrm{E}[X])^n] = \int (x-\mu)^n \rho(x)\, dx. \tag{3.2.9}$$

However, the first cumulant is the expected value, whereas the first central moment vanishes. As mentioned, the second and third cumulants are respectively the second and third central moments (the second central moment is the variance). Higher cumulants cannot be connected straightforwardly to central moments, but are rather more complicated polynomial functions of the moments [1].

Gaussian Distribution.—As will be described extensively below, Gaussian statistics is of significant importance in (primordial) cosmology. The PDF for a singlevariate Gaussian distribution with mean μ and variance σ^2 is given by:

$$\rho(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]. \tag{3.2.10}$$

For a gaussian distribution, the moment generating function is given by:

$$M(t) = \exp\left[\mu t + \frac{1}{2}\sigma^2 t^2\right], \tag{3.2.11}$$

and hence the cumulants generating function is given by:

$$K(t) = \ln M(t) = \mu t + \frac{1}{2}\sigma^2 t^2. \tag{3.2.12}$$

Therefore, we conclude that all but the first two cumulants vanish ($\kappa_{n\geq 3} = 0$). The first two are simple functions of the mean and variance and read:

$$\kappa_1 = \mu, \qquad \kappa_2 = \sigma^2. \tag{3.2.13}$$

Hence, a Guassian distribution is completely specified by its first and second cumulant or, equivalently, its mean and variance.

3.2.2 Multivariate Statistics

The discussion on the statistics of a single random variable can now be extended to an arbitrary number of random variables. For definiteness, we consider k random variables which are collected in a so-called random vector $\boldsymbol{X}$ as:

$$\boldsymbol{X} \equiv (X_1, \ldots, X_k)^{\mathrm{T}}, \tag{3.2.14}$$

where T denotes the transpose. The vector $\boldsymbol{X}$ can formally be defined as a *random* vector if any linear combination of the k components obeys a univariate distribution. The expectation value or mean generalizes to a mean vector $\boldsymbol{\mu}$ containing the expectation values of all k random variables, yielding:

$$\boldsymbol{\mu} \equiv \mathrm{E}[\boldsymbol{X}] = (\mathrm{E}[X_1], \ldots, \mathrm{E}[X_k])^{\mathrm{T}} \tag{3.2.15}$$

Lastly, the variance or covariance of the single random variable with itself generalizes to the $k \times k$ covariance matrix $\boldsymbol{\xi}$, defined as:

$$\begin{aligned} \boldsymbol{\xi} &\equiv \mathrm{E}[(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^{\mathrm{T}}] \\ \xi_{ij} &= \mathrm{Cov}[X_i, X_j] \equiv \langle X_i X_j \rangle; \qquad (1 \leq i, j \leq k). \end{aligned} \tag{3.2.16}$$

The covariance matrix $\xi_{ij} = \langle X_i X_j \rangle$ is also called the two-point correlation function.

Multivariate Normal Distribution.—In case the k-dimensional random vector $\boldsymbol{X}$ obeys a multi-variate Gaussian distribution, the multivariate probability density function is given by:

$$\rho_{\boldsymbol{X}}(x_1, \ldots, x_k) = \frac{1}{\sqrt{(2\pi)^k |\boldsymbol{\xi}|}} \exp\left[-\frac{1}{2}\sum_{ij} x_i (\xi)^{-1}_{ij} x_j\right], \tag{3.2.17}$$

where $|\boldsymbol{\xi}| \equiv \det \boldsymbol{\xi}$ is the determinant of the covariance matrix. In analogy to singlevariate Gaussian distribution, which is completely characterized by its mean and variance, the multivariate Gaussian distribution is completely characterized by its one- and two-point correlation functions.

Isserlis' Theorem.—In case $\boldsymbol{X}$ is normally distributed, higher order correlation functions $\langle X_i \cdots X_j \rangle$ can be evaluated using Isserlis' theorem,[10] stating that higher order correlation functions can be expanded in terms of two point correlation functions. The theorem states that if $(X_1, \ldots X_{2n})$ is a zero-mean normal random vector, higher order correlation functions can be expanded as:

$$\langle X_1 \cdots X_{2n} \rangle = \sum \prod \langle X_i X_j \rangle = \sum \prod \xi_{ij}, \tag{3.2.18}$$

$$\langle X_1 \cdots X_{2n-1} \rangle = 0, \tag{3.2.19}$$

where the notation $\sum \prod$ implies summation over all distinct ways of partitioning $X_1 \cdots X_{2n}$ into pairs two point correlators $\langle X_i X_j \rangle$ and each summand is the product of n two-point correlators. For instance, for $n = 2$ (the 4-point correlator) the result of Isserlis' theorem is:

$$\langle X_1 X_2 X_3 X_4 \rangle = \xi_{12}\xi_{34} + \xi_{13}\xi_{24} + \xi_{14}\xi_{23}. \tag{3.2.20}$$

Notice that for purely Guassian quantities φ, the odd correlation functions ($\langle \varphi_1 \cdots \varphi_n \rangle$ with n odd) vanish on account of Isserlis' theorem.

3.2.3 Non-Gaussianity

As Gaussian statistics will satisfy Isserlis' theorem, deviations from pure Gaussianity will manifest as violations of Isserlis' theorem. For instance, for a variable φ that is not precisely Gaussian distributed, the three point correlation function $\langle \varphi\varphi\varphi \rangle$ will be small but non-vanishing.[11] This notion of non-Gaussianity (NG) provides a powerful probe for inflation from both the observational and theoretical side. In the case of slow roll single-field inflation, the three point correlation function for the comoving curvature perturbation, denoted as $\langle \mathcal{R}\mathcal{R}\mathcal{R} \rangle$, is suppressed by the slow roll parameters and hence the level of NG is small.

There is an intuitive reason why $\langle \mathcal{R}\mathcal{R}\mathcal{R} \rangle$ is supressed by single-field slow roll models generically predict a small level of NG. In order to achieve accelerated expansion, the inflaton potential must be very flat (as quantified by $\varepsilon_V, \eta_V \ll 1$). Flatness of the potential can only be achieved by suppressing inflaton (self-)interactions and other sources of non-linearity. In this way, possible sources of large NG are minimized and hence only the weak coupling to gravity results in NG. Therefore, in case the observational level of NG is large, the complete single field scenario can be ruled out, making NG the most stringent test for single-field inflation to date.

[10] This theorem is sometimes referred to as Wick's theorem, but we will reserve this name for the well-known theorem in QFT, which describes how to expand higher order expectation values $\langle \mathcal{O}_1 \cdots \mathcal{O}_n \rangle$ in terms of two-point correlators.

[11] Notice that the three point correlation function is of particular interest, as it is the lowest order statistic that can discriminate between pure Gaussianity and deviations.

3.2.4 Statistical Homogeneity and Isotropy

We will now apply the above concepts to a generic (random) scalar field $\mathcal{Q}(\boldsymbol{x})$, whose value at each point in real space is determined by Gaussian distribution, and examine the implications of imposing statistical homogeneity and isotropy on the field (in analogy with the Cosmological Principle). The field $\mathcal{Q}$ can be represented as a random vector in which each entry corresponds to the random value of the field at a particular location in space.[12]

By definition, the two point correlation function of the field is given by:[13]

$$\xi_{ij} = \langle \mathcal{Q}(\boldsymbol{x}_i)\mathcal{Q}(\boldsymbol{x}_j)\rangle. \tag{3.2.21}$$

Let $\boldsymbol{r}_{ij}$ be the vector connecting $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, i.e. $\boldsymbol{r}_{ij} \equiv \boldsymbol{x}_j - \boldsymbol{x}_i$, the two point correlator can then be written as:

$$\xi_{ij} = \langle \mathcal{Q}(\boldsymbol{x}_i)\mathcal{Q}(\boldsymbol{x}_i + \boldsymbol{r}_{ij})\rangle. \tag{3.2.22}$$

Now, on account of statistical homogeneity (translational invariance), it is required that ξ_{ij} does not depend on $\boldsymbol{x}_{i,j}$, but only on the vector connecting them ($\boldsymbol{r}_{ij}$), hence:

$$\xi_{ij} \neq \xi_{ij}(\boldsymbol{x}_i, \boldsymbol{x}_j). \qquad \text{(Homogeneity)} \tag{3.2.23}$$

On account of statistical isotropy (rotational invariance), the correlator should be independent of the direction of $\boldsymbol{r}_{ij}$ and can therefore only depend on its magnitude $|\boldsymbol{r}_{ij}|$, yielding:

$$\xi_{ij} = \xi_{ij}(\boldsymbol{x}_i, |\boldsymbol{r}_{ij}|). \qquad \text{(Isotropy)} \tag{3.2.24}$$

In the above expression, we still allowed for dependence on $\boldsymbol{x}_i$ since isotropy alone is doest forbid dependence on spatial coordinates. However, when combining homogeneity and isotropy as imposed by the Cosmological Principle, we find that the correlator can only depend on the magnitude of $\boldsymbol{r}_{ij}$, leaving:

$$\xi_{ij} = \xi_{ij}(|\boldsymbol{r}_{ij}|). \qquad \text{(Homogeneity and Isotropy)} \tag{3.2.25}$$

For notation convenience, we will drop the labels i, j and use primes to denote different spatial vectors from now on.

The Fourier transform of the field $\mathcal{Q}(\boldsymbol{x})$ at location $\boldsymbol{x}$ is given by:

$$\mathcal{Q}_{\boldsymbol{k}} = \int d^3\boldsymbol{x}\ \mathcal{Q}(\boldsymbol{x})\ e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}. \tag{3.2.26}$$

In momentum space, the two point correlator is denoted as $\langle \mathcal{Q}_{\boldsymbol{k}}\mathcal{Q}_{\boldsymbol{k}'}\rangle$ and can be evaluated to be:

$$\langle \mathcal{Q}_{\boldsymbol{k}}\mathcal{Q}_{\boldsymbol{k}'}\rangle = (2\pi)^3\delta^{(3)}(\boldsymbol{k} + \boldsymbol{k}')P(k), \tag{3.2.27}$$

based on statistical homogeneity and isotropy. The quantity $P(k)$ is defined as $P(k) \equiv |\mathcal{Q}_k|^2$ and only depends on the magnitude of the wavevector by rotational invariance. Explicitly, it is the Fourier transform of the two point correlator $P(k)$:

$$P(k) = \int d^3\boldsymbol{r}\ \xi(\boldsymbol{r})\ e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}}. \tag{3.2.28}$$

[12] Note that since the scalar field encompasses an infinite number of degrees of freedom, the dimension of the corresponding random vector will tend to infinity: $k \to \infty$.

[13] Note that in this context, the labels i, j on the spatial coordinates are *not* to be regarded as spatial indices running from 1 to 3.

Derivation: Fourier Space Correlator

Here, we will derive the above result explicitly. The momentum space correlator can be written as:

$$\begin{aligned}\langle \mathcal{Q}_{\boldsymbol{k}} \mathcal{Q}_{\boldsymbol{k}'} \rangle &= \int d^3\boldsymbol{x} \int d^3\boldsymbol{x}' \, \langle \mathcal{Q}(\boldsymbol{x}) \mathcal{Q}(\boldsymbol{x}') e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \, e^{-i\boldsymbol{k}'\cdot\boldsymbol{x}'} \\ &= \int d^3\boldsymbol{x} \int d^3\boldsymbol{r} \, \langle \mathcal{Q}(\boldsymbol{x}) \mathcal{Q}(\boldsymbol{x}') \rangle \, e^{-i(\boldsymbol{k}+\boldsymbol{k}')\cdot\boldsymbol{x}} \, e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}} \\ &= \int d^3\boldsymbol{x} \, P(\boldsymbol{k}') \, e^{-i(\boldsymbol{k}+\boldsymbol{k}')\cdot\boldsymbol{x}} = (2\pi)^3 \delta^3(\boldsymbol{k}+\boldsymbol{k}') P(\boldsymbol{k}'). \end{aligned} \tag{3.2.29}$$

To go from the first to the second line, we used the fact that the two coordinates are related as $\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{r}$. In the last line we used the definition of the momentum space Dirac delta function and we defined $P(\boldsymbol{k}')$ as:

$$P(\boldsymbol{k}') = \int d^3\boldsymbol{r} \, \xi(\boldsymbol{r}) \, e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}}. \tag{3.2.30}$$

Above, we concluded that the correlator only depends on the magnitude of the separation vector $r \equiv |\boldsymbol{r}|$. In momentum space, this is equivalent to assuming that P is only dependent on the magnitude of the 3-momentum vector. Lastly, we interchange the prime and unprimed variables and obtain the desired result:

$$\langle \mathcal{Q}_{\boldsymbol{k}} \mathcal{Q}_{\boldsymbol{k}'} \rangle = (2\pi)^3 \delta^{(3)}(\boldsymbol{k}+\boldsymbol{k}') P(k). \tag{3.2.31}$$

Variance and Qauntum Field Operators

The variance of $\mathcal{Q}$, denoted $\sigma_{\mathcal{Q}}^2 = \langle \mathcal{Q}^2(\boldsymbol{x}) \rangle$, follows from the two-point correlation function $\xi_{\mathcal{Q}}(\boldsymbol{r}) \equiv \langle \mathcal{Q}(\boldsymbol{x}) \mathcal{Q}(\boldsymbol{x}') \rangle$ by setting $\boldsymbol{r} = 0$. This yields:

$$\sigma_{\mathcal{Q}}^2 = \langle \mathcal{Q}^2(\boldsymbol{x}) \rangle = \xi_{\mathcal{Q}}(0) = \int_0^\infty \frac{dk}{k} \mathcal{P}_{\mathcal{Q}}. \tag{3.2.32}$$

In case the quantity of interest $\mathcal{Q}$ corresponds to a quantum operator or field, denoted using a hat $\hat{\mathcal{Q}}$, the variance can be identified with the expectation value of the square of the operator $|\hat{\mathcal{Q}}|^2 = \hat{\mathcal{Q}}^\dagger \hat{\mathcal{Q}}$ in the vacuum/ground state $|0\rangle$.[14] That is:

$$\boxed{\langle |\hat{\mathcal{Q}}|^2 \rangle \equiv \langle 0 | \hat{\mathcal{Q}}^\dagger \hat{\mathcal{Q}} | 0 \rangle = \int_0^\infty \frac{dk}{k} \mathcal{P}_{\hat{\mathcal{Q}}}.} \tag{3.2.33}$$

In particular, this result will be used to compute the power spectrum of the quantized inflaton fluctuation field $\hat{f}$ ($f \equiv a\delta\phi$).

3.2.5 The Power Spectrum: Connecting Theory with Observations

Using the forward Fourier transform, the power spectrum in Fourier space can be defined by exploiting the spherical geometry of the $d^3\boldsymbol{k}$ integral. The Fourier power spectrum for the

[14] The dagger † indicates the conjugate transpose of the operator.

variable $\mathcal{Q}$ is denoted by $\mathcal{P}_{\mathcal{Q}}$ and reads:

$$\boxed{\mathcal{P}_{\mathcal{Q}} = \frac{k^3}{2\pi^2}|\mathcal{Q}_{\boldsymbol{k}}|^2.} \tag{3.2.34}$$

The power spectrum, which is, in essence, the Fourier counterpart of the correlation function, will be of crucial importance in relating predictions of inflation to cosmological observations. For a Gaussian random field, all statistical information in contained in the power spectrum (or equivalently the two-point correlation function in real space). In other words, for a Gaussian random field, the power spectrum will not posses any scale dependence by means of k.

We will show that the power spectrum for $\mathcal{R}$ is indeed almost scale-invariant as the spectral index, parametrizing the (logarithmic) dependence of the power spectrum on k, is supressed by slow roll parameters. Hence, $\mathcal{R}$ is well described as a Gaussian random field, which is in line with the small level of NG generated by single field inflation (section 3.2.3).

Derivation: Power Spectrum

The differential volume element d^3k in Fourier space can be expressed in terms of the radial coordinate k and angular coordinates θ and ϕ as:

$$d^3k \equiv k^2 \sin\theta \, d\theta \, d\phi \, dk. \tag{3.2.35}$$

Furthermore, the dot product in the exponent $\boldsymbol{k} \cdot \boldsymbol{r}$ yields:

$$\boldsymbol{k} \cdot \boldsymbol{r} = kr\cos\theta, \tag{3.2.36}$$

where $k = |\boldsymbol{k}|$ and $r = |\boldsymbol{r}|$. Substituting these expressions in the forward Fourier transform of the correlation function and evaluating expanding the integral over d^3k into explicit integrals over k, θ and ϕ yields:

$$\xi_{\mathcal{Q}}(\boldsymbol{r}) = \frac{1}{(2\pi)^3}\int_0^\infty dk \int_0^\pi \sin\theta \, d\theta \int_0^{2\pi} d\phi \; k^2|\mathcal{Q}_{\boldsymbol{k}}|^2 \; e^{ikr\cos\theta}. \tag{3.2.37}$$

The integral over θ yields the first Bessel function of the first kind:

$$\int_0^\pi e^{ikr\cos\theta}\sin\theta \, d\theta = \int_{-1}^{+1} e^{ikr\cos\theta} \, d(\cos\theta) = 2\frac{\sin kr}{kr} = 2j_0(kr). \tag{3.2.38}$$

Integrating over ϕ as well yields the following result:

$$\xi_{\mathcal{Q}}(\boldsymbol{r}) = \int_0^\infty dk \, \frac{4\pi k^2}{(2\pi)^3} \, |\mathcal{Q}_{\boldsymbol{k}}|^2 \, \frac{\sin kr}{kr} = \int_0^\infty \frac{dk}{k}\left[\frac{k^3}{2\pi^2}|\mathcal{Q}_{\boldsymbol{k}}|^2\right]\frac{\sin kr}{kr} = \int_0^\infty \frac{dk}{k} \, \mathcal{P}_{\mathcal{Q}} \, j_0(kr). \tag{3.2.39}$$

In the last line, the (dimensionless) *power spectrum* $\mathcal{P}_{\mathcal{Q}}$ is defined as:

$$\mathcal{P}_{\mathcal{Q}} \equiv \frac{k^3}{2\pi^2}|\mathcal{Q}_{\boldsymbol{k}}|^2. \tag{3.2.40}$$

This gives the desired result.

3.2.6 The Central Limit Theorem

Even without resorting to theoretical motivations to argue that cosmological perturbations (such as $\mathcal{R}$) will be Gaussian to a very high degree, there exists a general theorem in statistics that constrains cosmological perturbations to be approximately Gaussian. This theorem is known as the central limit theorem (CLT) and states that Gaussianity will always result from a superposition of a large number of random processes [84]. The power of the CLT lies in the fact that the underlying random processes do not have to be drawn from a Gaussian distribution; the only requirement for CLT is that the underlying distribution has a finite variance. We will not go over the details here; the CLT is illustrated in Appendix D.1.

3.3 Gaussian Random Fields

On account of well-motivated theoretical considerations (see sections 3.2.3 and 3.2.5) or the central limit theorem, we expect cosmological perturbations to be nearly Gaussian. Therefore, we can model cosmological perturbations to a first approximation as Gaussian random fields, which we will introduce formally in this section (see also Refs. [67, 81, 84]). Gaussian random fields are fields that are completely specified by their one- and two-point correlation functions (i.e. by their mean, which is often set to zero by a field redefinition and the variance). Non-vanishing behavior of higher order correlation functions then encode deviations from pure Gaussianity, i.e. they describe non-Gaussian (NG) contributions to the perturbations.

Consider a generic position-dependent scalar field $\mathcal{Q}(\boldsymbol{x})$. In principle, the field may also posses temporal dependence, but it will not be included explicitly in the following discussion. The Fourier transform can be expressed as:

$$\mathcal{Q}_{\boldsymbol{k}} = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \mathcal{Q}_{\boldsymbol{k}}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}. \tag{3.3.1}$$

Without loss of generality, the Fourier mode of the field $\mathcal{Q}_{\boldsymbol{k}}$ can be written in the following way:

$$\mathcal{Q}_{\boldsymbol{k}} \equiv \alpha_{\boldsymbol{k}} + i\beta_{\boldsymbol{k}}, \tag{3.3.2}$$

where $\alpha_{\boldsymbol{k}}$ and $\beta_{\boldsymbol{k}}$ are real parameters. The reality of the field $\mathcal{Q}$ implies $\alpha_{\boldsymbol{k}} = \alpha_{-\boldsymbol{k}}$ and $\beta_{\boldsymbol{k}} = -\beta_{-\boldsymbol{k}}$.

A given field configuration is completely specified by means of the parameters $\alpha_{\boldsymbol{k}}$ and $\beta_{\boldsymbol{k}}$, which are themselves parametrized by the wavevector $\boldsymbol{k}$. In order to generate a random field configuration, one can regard $\alpha_{\boldsymbol{k}}$ and $\beta_{\boldsymbol{k}}$ as random variables and assign a PDF to them. In particular, for a Gaussian random field, $\alpha_{\boldsymbol{k}}$ and $\beta_{\boldsymbol{k}}$ are drawn from a Gaussian distribution:

$$\rho(\alpha_{\boldsymbol{k}}, \beta_{\boldsymbol{k}}) = \frac{1}{\pi\sigma_{\boldsymbol{k}}^2} \exp\left[-\frac{\alpha_{\boldsymbol{k}}^2 + \beta_{\boldsymbol{k}}^2}{\sigma_{\boldsymbol{k}}^2}\right], \tag{3.3.3}$$

where the normalization factor $(\pi\sigma_{\boldsymbol{k}}^2)^{-1}$ is chosen such that the probability density function integrate to unity when integrating over both $\alpha_{\boldsymbol{k}}$ and $\beta_{\boldsymbol{k}}$. Notice that the above result applies to solely one $\boldsymbol{k}$-mode. Generalizing to all possible modes $\boldsymbol{k}$, the PDF becomes a *functional* of all possible field configurations, as parametrized by different wavevectors $\boldsymbol{k}$. The PDF functional reads:

$$\rho[\mathcal{Q}_{\boldsymbol{k}}] = \frac{1}{\pi\sigma_{\boldsymbol{k}}^2} \exp\left[-\frac{|\mathcal{Q}_{\boldsymbol{k}}|^2}{\sigma_{\boldsymbol{k}}^2}\right], \tag{3.3.4}$$

where we define the amplitude of the field mode as $|\mathcal{Q}_{\boldsymbol{k}}| \equiv \sqrt{\alpha_{\boldsymbol{k}}^2 + \beta_{\boldsymbol{k}}^2}$. The PDF functional formalizes the statement that all field configurations are drawn from a Gaussian distribution. Observe that the functional above only states the amplitude of the field, denoted as $|\mathcal{Q}_{\boldsymbol{k}}|$, is drawn from a Gaussian distribution. The phase, however, is drawn from a flat distribution. Furthermore, notice that in general, the variance may depend on the wavevector. However, in assuming statistical isotropy, the variance $\sigma_{\boldsymbol{k}}^2/2$ depends only on the wavenumber ($\sigma_{\boldsymbol{k}}^2 \to \sigma_k^2$) and is the same for both $\alpha_{\boldsymbol{k}}$ and $\beta_{\boldsymbol{k}}$. For a Gaussian random field, the variance (i.e. two-point correlation function) contains all the statistical information and all higher order correlation functions will vanish.

The expectation value $\langle \mathcal{A} \rangle$ of some observable $\mathcal{A}$ are taken over the ensemble of all possible realizations of $\mathcal{A}$. In this case, observables themselves become functionals of the field configuration, i.e. $\mathcal{A} = \mathcal{A}[\mathcal{Q}_{\boldsymbol{k}}]$. The expectation value of the observable is given in terms of an integral over all possible field configurations:

$$\boxed{\langle \mathcal{A}[\mathcal{Q}_{\boldsymbol{k}}] \rangle = \int \mathcal{D}\mathcal{Q}_{\boldsymbol{k}}\ \mathcal{A}[\mathcal{Q}_{\boldsymbol{k}}]\ \rho[\mathcal{Q}_{\boldsymbol{k}}],} \tag{3.3.5}$$

where the integrand notation $\mathcal{D}\mathcal{Q}_{\boldsymbol{k}}$ implies the integration is over all possible field configurations.

For a Gaussian random field, the expectation value for a given observable $\mathcal{A}$ can be written explicitly by using the fact that the integration measure over all field configurations can be written as:

$$\mathcal{D}\mathcal{Q}_{\boldsymbol{k}} = \prod_{\boldsymbol{k}} d\alpha_{\boldsymbol{k}}\ d\beta_{\boldsymbol{k}}, \tag{3.3.6}$$

yielding the following expression for $\mathcal{A} = \mathcal{A}[\mathcal{Q}_{\boldsymbol{k}}]$:

$$\mathcal{A}[\mathcal{Q}_{\boldsymbol{k}}] = \int \prod_{\boldsymbol{k}} d\alpha_{\boldsymbol{k}}\ d\beta_{\boldsymbol{k}}\ \frac{1}{\pi\sigma_{\boldsymbol{k}}^2}\ \mathcal{A}[\mathcal{Q}_{\boldsymbol{k}}]\ \exp\left[-\frac{|\mathcal{Q}_{\boldsymbol{k}}|^2}{\sigma_{\boldsymbol{k}}^2}\right]. \tag{3.3.7}$$

Using the above expression, the correlation function of two field modes, characterized by momenta $\boldsymbol{k}$ and $\boldsymbol{k}'$, can be written as:

$$\langle \mathcal{Q}_{\boldsymbol{k}} \mathcal{Q}_{\boldsymbol{k}'} \rangle = \langle \alpha_{\boldsymbol{k}} \alpha_{\boldsymbol{k}'} \rangle - \langle \beta_{\boldsymbol{k}} \beta_{\boldsymbol{k}'} \rangle = \sigma_k^2\ \delta^{(3)}(\boldsymbol{k} + \boldsymbol{k}'). \tag{3.3.8}$$

To derive this result, we used the fact that cross terms in α and β vanish and that the expectation value of $\langle \alpha_{\boldsymbol{k}} \alpha_{\boldsymbol{k}'} \rangle$ gives:

$$\langle \alpha_{\boldsymbol{k}} \alpha_{\boldsymbol{k}'} \rangle = \int d^3\boldsymbol{k}\ \alpha_{\boldsymbol{k}} \alpha_{\boldsymbol{k}}\ \rho[\mathcal{Q}_{\boldsymbol{k}}]\ \delta^{(3)}(\boldsymbol{k} + \boldsymbol{k}') = \frac{\sigma_k^2}{2}\ \delta^{(3)}(\boldsymbol{k} + \boldsymbol{k}'). \tag{3.3.9}$$

For the term $\langle \beta_{\boldsymbol{k}} \beta_{\boldsymbol{k}'} \rangle$, we would obtain the same result but we can use the reality constraint $\beta_{\boldsymbol{k}} = -\beta_{-\boldsymbol{k}}$ to compensate the for the minus sign between the terms in Eq. 6.4.52.

With the above results at hand, we can now compute the real space two-point correlation function for the Guassian random field, which is given by:

$$\begin{aligned}
\langle \mathcal{Q}(\boldsymbol{x}) \mathcal{Q}(\boldsymbol{x}') &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \frac{d^3\boldsymbol{k}'}{(2\pi)^3}\ \langle \mathcal{Q}_{\boldsymbol{k}} \mathcal{Q}_{\boldsymbol{k}'} \rangle\ e^{i\boldsymbol{k}\cdot\boldsymbol{x}} e^{i\boldsymbol{k}'\cdot\boldsymbol{x}'} \\
&= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \frac{d^3\boldsymbol{k}'}{(2\pi)^3}\ \sigma_k^2\ \delta^{(3)}(\boldsymbol{k} + \boldsymbol{k}')\ e^{i\boldsymbol{k}\cdot\boldsymbol{x}} e^{i\boldsymbol{k}'\cdot\boldsymbol{x}'} \\
&= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \frac{\sigma_k^2}{(2\pi)^3}\ e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')}.
\end{aligned} \tag{3.3.10}$$

On account of Eq. 3.2.27, we find the $P(k)$ is equal to:

$$P(k) = \frac{\sigma_k^2}{(2\pi)^3}. \tag{3.3.11}$$

Note that for a scale invariant two-point correlation function in real space, i.e. for one that satisfies the following invariance equation under the rescaling of the coordinates $\boldsymbol{x} \to \lambda \boldsymbol{x}$ with the scaling parameter $\lambda > 0$:

$$\langle \mathcal{Q}(\boldsymbol{x})\mathcal{Q}(\boldsymbol{x'})\rangle \equiv \langle \mathcal{Q}(\lambda\boldsymbol{x})\mathcal{Q}(\lambda\boldsymbol{x'}), \tag{3.3.12}$$

it follows that $P(k)$ exhibits the following momentum dependence,

$$P(k) \propto k^{-3}. \tag{3.3.13}$$

on account of the Dirac-delta identity $\delta^{(3)}(\lambda(\boldsymbol{k}+\boldsymbol{k'})) = \lambda^{-3}\delta^{(3)}(\boldsymbol{k}+\boldsymbol{k'})$. For a Gaussian random field, the power spectrum $\mathcal{P}$ which is proportional to $k^3P(k)$ is independent of wavenumber k and hence scale invariant:

$$\mathcal{P}_{\mathcal{Q}} = \frac{k^3}{2\pi^2}P(k) \propto k^0. \tag{3.3.14}$$

3.4 Connection to Inflation

When applying the statistical methods outlined above to compute the power spectrum for inflaton fluctuations in the single-field scenario, we will find that the power spectrum $\mathcal{P}_{\mathcal{R}}$ is indeed highly Gaussian, i.e. the dependence on k will be very small (in good agreement with Eq. 3.3.14 and hence $\mathcal{R}$ is well described as a Gaussian random field). In particular, using the slow-roll approximation we will show that the scale-dependence of the power spectrum, as measured by the spectral index $n_{\rm s}$, is indeed suppressed by slow roll parameters:

$$n_{\rm s} - 1 \equiv \frac{d\ln\mathcal{P}_{\mathcal{R}}}{d\ln k} + \mathcal{O}(dn_{\rm s}/d\ln k) = -2\varepsilon - \eta = 2\eta_{\rm V} - 6\varepsilon_{\rm V}. \tag{3.4.1}$$

Here $\mathcal{O}(dn_{\rm s}/d\ln k)$ indicates that we neglect the (minor) dependence of the spectral index itself on k. Note that a perfectly scale-invariant power spectrum corresponds to the limit $n_{\rm s} - 1 \to 0$. Parametrizing the k-dependence of $n_{\rm s}$ would amount to including so-called spectral runnings (derivatives of $n_{\rm s}$ with respect to k) in the spectral index.

The almost scale invariant power spectrum as predicted by single-field inflation is indeed supported by observations on the CMB as made by Planck, from which the power spectrum for $\mathcal{R}$ could be inferred. Ignoring spectral runnings, the power spectrum can be parametrized as [3, 84]:

$$\boxed{\mathcal{P}_{\mathcal{R}}(k) = \frac{k^3}{2\pi^2}|\mathcal{R}_k|^2 = A_{\rm s}\left(\frac{k}{k_*}\right)^{n_{\rm s}-1}.} \tag{3.4.2}$$

Here, $k_* = 0.05\,\mathrm{Mpc}^{-1}$ is a pivot scale, $A_{\rm s}$ is the scalar amplitude and $n_{\rm s}$ is the spectral index, governing the scale dependence of the power spectrum. Planck constrained the spectral index to be [3]:

$$n_{\rm s} = 0.9645 \pm 0.0049. \tag{3.4.3}$$

Parameter	Parametrization	Bound/Value [3]	Dataset
Pivot Scale	$k_* = a_* H_*$	$0.05\,\mathrm{Mpc}^{-1}$	–
Scalar Amplitude	$\ln(10^{10} A_\mathrm{S})$	3.094 ± 0.034	TT, TE, EE+lowP
Scalar Spectral Index	$n_\mathrm{s} = 2\eta_\mathrm{V} - 6\epsilon_\mathrm{V} + 1$	0.9645 ± 0.0049	TT, TE, EE+lowP
Tensor Amplitude	A_T	–	–
Tensor Spectral Index	$n_\mathrm{t} = -2\epsilon_\mathrm{V}$	–	–
Tensor-to-Scalar Ratio	$r = 16\epsilon_\mathrm{V}$	< 0.103	TT+lowP

Table 3.1: Inflationary parameters and constraints on their values (if available) by Planck [3]. Notice that constraints on the tensor quantities A_t and n_t are not available as primordial gravitational waves have not been observed (yet).

Notice that the exponent $n_\mathrm{s}-1$ will be very close to zero, supporting nearly scale-invariance as predicted by inflation. In the full Planck parametrisation of the power spectrum, higher order derivatives of $dn_\mathrm{s}/d\ln k$ (i.e. spectral runnings), are included in the exponent of (k/k_*).[15]

For the primordial gravitational wave (tensor) power spectrum, the approach is similar to the spectrum of scalar fluctuations discussed above. In this case, however, the two different polarization directions of the waves, as indicated by $(+,\times)$, should be taken into account in the definition of the power spectrum:

$$\boxed{\mathcal{P}_E = \frac{k^3}{2\pi^2}\left(|E^+|^2 + |E^\times|^2\right) = A_\mathrm{T}\left(\frac{k}{k_*}\right)^{n_\mathrm{t}},} \tag{3.4.5}$$

where A_T and $n_\mathrm{t} = -2\epsilon_\mathrm{V}$ are the tensor amplitude and spectral index, respectively. As for the scalar power spectrum, the contribution of the spectral running is ignored in the exponent. Up till now, primordial gravitational waves have not been detected. However, they can be constrained using CMB observation by means of the so-called tensor-to-scalar ratio r, defined as:

$$r \equiv \frac{\mathcal{P}_E(k_*)}{\mathcal{P}_\mathcal{R}(k_*)} = \frac{A_\mathrm{T}^2}{A_\mathrm{S}^2} = 16\epsilon_\mathrm{V}, \tag{3.4.6}$$

this relation will be derived explicitly in section 5.6. From the data, Planck found an upper bound to this parameter which reads [3]:

$$r_{0.002} < 0.11, \tag{3.4.7}$$

where the subscript indicates that the tensor-to-scalar ratio is computed at the pivot scale $k_* = 0.002\,\mathrm{Mpc}^{-1}$, instead of the value $0.05\,\mathrm{Mpc}^{-1}$ used earlier.

Notice that the connection between theory and experiment by means of the spectral index and tensor-to-scalar ratio can be used to discriminate between different inflationary models, corresponding to different potential functions of the scalar field $V(\phi)$. To test whether an inflationary model is favored or rejected by the Planck data, one computes the slow-roll

[15] The exponent in the Planck parametrization is written as a series expansion of (higher orders of) the spectral running:

$$n_\mathrm{s} - 1 + \frac{1}{2}\frac{dn_\mathrm{s}}{d\ln k}\ln\frac{k}{k_*} + \frac{1}{6}\frac{d^2 n_\mathrm{s}}{d\ln k^2}\left(\ln\frac{k}{k_*}\right)^2 + \cdots. \tag{3.4.4}$$

parameters ($\epsilon_{\rm V}$, $\eta_{\rm V}$), which are directly related to the shape of the potential. Subsequently, the spectral index and tensor-to-scalar ratio can be predicted using Eq. 3.4.1 and Eq. 3.4.6. In case the predicted value lies within the experimental bounds of $n_{\rm s}$ and $n_{\rm t}$, the considered model is favored by the Planck data.

All the relevant inflationary parameters introduced in this subsection are tabulated in Tab. 3.1, accopanied by constraints on their values as determined by Planck [3]. Notice that there are constraints on the tensor power spectrum, as primordial gravitational waves have not been detected yet.

3.5 Intuition from a Toy Model

Here, we will discuss a toy model that captures most of the essential features about the physics of scalar field fluctuations in an inflationary background (i.e. a quasi De-Sitter background, see Appendix C.2). The aim of this toy model is to illustrate the concepts described in this chapter in a more quantitative way, without the need to perform a full-detail computationally demanding analysis – this will be done in the upcoming chapters.

We consider a scalar field χ, which is *not* the inflaton, and assume no direct coupling between χ and ϕ. That is, χ is taken to be a *free* field and its potential $V(\chi)$ is only sourced by self-interactions. Such a field is called a *spectator* or test field during inflation and were studied in e.g. [73, 74, 86, 88]. In addition, the coupling to gravity is assumed to minimal, meaning the coupling constant ξ between the Ricci scalar R and χ vanishes in the total Lagrangian $\mathcal{L}_\chi$. The total action $S_{\rm TM}$ for the toy model (TM) can therefore be written as:

$$S_{\rm TM} = \int d^4x\sqrt{-g}\left[\frac{1}{2}M_{\rm pl}R^2 - \frac{1}{2}g^{\mu\nu}\partial_\mu\chi\partial_\nu\chi - V(\chi)\right]. \tag{3.5.1}$$

Since by definition the inflaton field dominates the energy density in the universe during inflation, the energy density of χ is much smaller than the inflaton energy density:

$$\rho_\chi \ll \rho_\phi. \tag{3.5.2}$$

For that reason, fluctuations in χ do not affect the dynamics of the background considerably: the coupling between the metric fluctuations $\delta g_{\mu\nu}(\boldsymbol{x},t)$ and $\delta\chi(\boldsymbol{x},t)$ is small. Therefore, the backreaction (see section 3.1.2) of the perturbed metric on the field χ can be neglected completely. This simplifies the analysis significantly since in the absence of a backreaction the evolution equation for the fluctuation $\delta\chi$ can be obtained trivially by perturbing the equation of motion for χ at linear order.

The evolution of the field χ is governed by the Klein-Gordon equation, written conveniently in terms of the d'Alembertian operator $\Box\chi = V_\chi$. Evaluating d'Alembertian for the FRW background and keeping the spatial gradient term yields:

$$\ddot{\chi} + 3H\dot{\chi} - \frac{1}{a^2}\partial^2\chi + V_\chi = 0, \tag{3.5.3}$$

which is identical to Eq. 2.5.7 except for the gradient term (which is omitted in Eq. 2.5.7). For a homogeneous field, the spatial gradient $\partial^2\chi$ vanishes (we defined $\partial^2 \equiv \partial_i\partial^i$). However, in order to obtain the evolution equation for fluctuations $\delta\chi(\boldsymbol{x},t)$, we will perturb the above equation to linear order and acting with ∂^2 on $\delta\chi(\boldsymbol{x},t)$ will not yield zero. For that reason, we keep the gradient term in the above equation.

To first order, the total field $\chi(\boldsymbol{x}, t)$ can be written as the sum of an homogeneous background $\overline{\chi}(t)$ and perturbation $\delta\chi(\boldsymbol{x}, t)$ in the following way:

$$\chi(\boldsymbol{x}, t) = \overline{\chi}(t) + \delta\chi(\boldsymbol{x}, t). \tag{3.5.4}$$

Inserting this form for $\chi(\boldsymbol{x}, t)$ into Eq. 3.5.3 and extracting all terms linear in $\delta\chi(\boldsymbol{x}, t)$ gives the following evolution equation:

$$\delta\ddot{\chi}(\boldsymbol{x}, t) + 3H\delta\dot{\chi}(\boldsymbol{x}, t) - \frac{1}{a^2}\partial^2\delta\chi(\boldsymbol{x}, t) + V_{\chi\chi}\delta\chi(\boldsymbol{x}, t) = 0. \tag{3.5.5}$$

Moving to Fourier space amounts to replacing ∂^2 by $-k^2$ and labelling the Fourier modes $\delta\chi_{\boldsymbol{k}}(t)$, yielding:

$$\delta\ddot{\chi}_{\boldsymbol{k}}(t) + 3H\delta\dot{\chi}_{\boldsymbol{k}}(t) + \frac{k^2}{a^2}\delta\chi_{\boldsymbol{k}}(t) + V_{\chi\chi}\delta\chi_{\boldsymbol{k}}(t) = 0. \tag{3.5.6}$$

3.5.1 Sub- and Super-Horizon Behaviour of Fluctuations

First, we assume the scalar field χ to be massless and set the potential to zero. We study the qualitative behavior of the fluctuation modes $\delta\chi_{\boldsymbol{k}}$ in the sub- and superhorizon limit.

- ▷ *Sub-Horizon Limit.*—For modes deep inside the horizon, k satisfies the condition $k/aH \gg 1$, and hence the damping term $3H\delta\dot{\chi}_{\boldsymbol{k}}$ can be neglected in the evolution equation, leaving:

$$\delta\ddot{\chi}_{\boldsymbol{k}} + \frac{k^2}{a^2}\delta\chi_{\boldsymbol{k}} = 0. \tag{3.5.7}$$

  Except for the time-dependent prefactor of $\delta\chi_{\boldsymbol{k}}$, this evolution equation resembles that of an harmonic oscillator. Therefore, the solution for the fluctuation will have an oscillatory character inside the horizon.

- ▷ *Super-Horizon Limit.*—For modes far outside the horizon, $k/aH \ll 1$ and the term proportional to k^2 can be neglected, yielding:

$$\delta\ddot{\chi}_{\boldsymbol{k}} + 3H\delta\dot{\chi}_{\boldsymbol{k}} = 0. \tag{3.5.8}$$

  The solution to this equation gives:

$$\delta\chi_{\boldsymbol{k}}(t) = \frac{\mathcal{C}_1}{3H}e^{-3Ht} + \mathcal{C}_2, \tag{3.5.9}$$

  where $\mathcal{C}_1$ and $\mathcal{C}_2$ are time-independent constants. Hence, on super-horizon scales the solution consists of a exponentially decay mode and a constant mode. We conclude that the fluctuation freezes in outside the horizon. The constancy of the fluctuation outside the horizon is very similar to the advocated constancy of the comoving curvature perturbation $\mathcal{R}$ on these scales (this feature of $\mathcal{R}$ will be proven in section 6.3).

3.5.2 Fluctuations from a Massive Scalar Field

Now, we turn on the potential again and consider a potential containing solely a quadratic mass term m_χ. We assume that the mass is small relative to the Hubble scale, i.e. $m_\chi^2/H^2 \ll 1$. The potential function reads:

$$V(\chi) = \frac{1}{2}m_\chi^2\chi^2, \tag{3.5.10}$$

so that the second derivative is given by $V_{\chi\chi} = m_\chi^2$. By making the field redefinition $\sigma(\boldsymbol{x}, t) \equiv a\delta\chi(\boldsymbol{x}, t)$ and moving to conformal time $d\tau \equiv dt/a$ one obtains the following equation of motion for the redefined field:

$$\sigma_{\boldsymbol{k}}'' + \left[k^2 - \frac{a''}{a} + a^2 m_\chi^2\right]\sigma_{\boldsymbol{k}} = 0. \tag{3.5.11}$$

Notice that the above evolution equation has the form of that for an harmonic oscillator with time-dependent frequency:

$$\omega_k^2(\tau) \equiv k^2 + M^2(\tau), \tag{3.5.12}$$

where the mass term is $M^2(\tau)$ is defined as $M^2(\tau) \equiv a^2 m_\chi^2 - a''/a$. The derived equation of motion can be obtained via the variational principle from the following action:

$$S[\sigma_{\boldsymbol{k}}, \sigma_{\boldsymbol{k}}'] = \frac{1}{2}\int d\tau d^3\boldsymbol{x}\Big[\sigma_{\boldsymbol{k}}' - (k^2 + M^2(\tau))\sigma_{\boldsymbol{k}}\Big]. \tag{3.5.13}$$

This action and the corresponding Lagrangian $\mathcal{L}_\sigma$ will be used later on to canonically quantize the fluctuations in the field.

Inflation is well-described by a quasi De-Sitter background, for which the time-variation of the Hubble parameter is small, as quantified with $\varepsilon \equiv -\dot{H}/H^2 \ll 1$. Using conformal time, the definition for ε can be rewritten in the following differential equation for the Hubble radius:

$$\frac{d}{d\tau}\left(\frac{1}{aH}\right) \equiv \varepsilon - 1. \tag{3.5.14}$$

To first order in the parameter ε, the scale factor can be obtained from the above equation as:

$$a(\tau) = -\frac{1}{H\tau}(1+\varepsilon). \tag{3.5.15}$$

Using this form for the scale factor, the time-dependent mass parameter $M^2(\tau)$ can be written to first order as:

$$M^2(\tau) = \frac{1}{\tau^2}\left(3\eta_\chi - 2 - 3\varepsilon\right). \tag{3.5.16}$$

Here, we defined the parameter $\eta_\chi \equiv m_\chi^2/3H^2 \ll 1$ in analogy with the Hubble flow parameter $\eta \equiv \dot{\varepsilon}/H\varepsilon$. In the above expansion for $M^2(\tau)$, only terms linear in η_χ and ε are considered. In defining the parameter $\nu_\chi \equiv 3/2 + \varepsilon - \eta_\chi$, the mass term can be written to first order in ε and η_χ as:

$$M^2(\tau) = -\frac{\nu_\chi^2 - 1/4}{\tau^2}. \tag{3.5.17}$$

Finally, the evolution equation for the field $\sigma_{\boldsymbol{k}}$ can be written as:

$$\sigma_{\boldsymbol{k}}'' + \left[k^2 - \frac{\nu_\chi^2 - 1/4}{\tau^2}\right]\sigma_{\boldsymbol{k}} = 0. \tag{3.5.18}$$

The solution to this equation for ν_χ real is given by the linear combination of Hankel functions of the first and second kind $H_{\nu_\chi}^{(1,2)}(x)$, which are given by:

$$H_\nu^{(1)}(x) \equiv J_\nu(x) + iY_\nu(x), \qquad H_\nu^{(2)}(x) \equiv J_\nu(x) - iY_\nu(x), \tag{3.5.19}$$

where $J_\nu(x)$ and $Y_\nu(x)$ are the Bessel functions of the first and second kind, respectively. For constants $c_{1,2}$ to be determined by the initial conditions, the generic solution for $\sigma_{\boldsymbol{k}}$ is given by:

$$\sigma_{\boldsymbol{k}}(\tau) = \sqrt{-\tau}\left[c_1 H^{(1)}_{\nu_\chi}(-k\tau) + c_2 H^{(2)}_{\nu_\chi}(-k\tau)\right]. \tag{3.5.20}$$

Notice that the solution only depends on the magnitude of the momentum mode $k = |\boldsymbol{k}|$ and hence from now on we omit the explicit vector notation and write σ_k instead.

3.5.3 Initial Condition and Bunch-Davies Vacuum

In order to completely fix the dynamics of the field modes σ_k in terms of explicit expressions for the constants $c_{1,2}$, we need to determine the initial condition for the mode function. At early times, as described by the limit $|k\tau| \gg 1$ the time-dependent mode frequency $\omega_k(\tau)$ becomes time-*in*dependent:

$$\omega_k^2(k,\tau) = k^2\left[1 - \frac{\nu_\chi^2 - 1/4}{(k\tau)^2}\right] \xrightarrow{|k\tau| \gg 1} k^2. \tag{3.5.21}$$

In this regime, the equation of motion can be written simply as:

$$\sigma_k'' + k^2\sigma_k = 0, \tag{3.5.22}$$

for which the generic solution can be written as:

$$\sigma_k(\tau) = \mathcal{C}_1 e^{-ik\tau} + \mathcal{C}_2 e^{+ik\tau}. \tag{3.5.23}$$

Together with the appropriate expressions for the constants $\mathcal{C}_{1,2}$, this solution for the field mode σ_k in the early time limit forms the initial condition for the generic solution (Eq. 3.5.20). We determine the constants $\mathcal{C}_{1,2}$ by requiring the field σ_k to be in the lowest energy state or vacuum state. This procedure can only be performed by quantizing the field $\sigma(\boldsymbol{x}, t)$, i.e. within the context of quantum field theory (QFT), and will be shown in detail for the inflaton field in section 5.4. Constructing the (quantum operator for the) Hamiltonian $\hat{H}$ for the system and minimizing its expectation value $E_0 \equiv \langle 0|\hat{H}|0\rangle$ in the vacuum state of the field, denoted as $|0\rangle$, yields the following conditions on $\mathcal{C}_{1,2}$:

$$\mathcal{C}_1 = \frac{1}{\sqrt{2k}}, \qquad \mathcal{C}_2 = 0. \tag{3.5.24}$$

Hence, the initial condition for the generic solution of $\sigma_k(\tau)$ (Eq. 3.5.20) is given by:

$$\boxed{\lim_{|k\tau|\gg 1} \sigma_k(\tau) = \frac{1}{\sqrt{2k}} e^{-ik\tau}.} \tag{3.5.25}$$

The constructed vacuum from which this initial condition for the mode function is derived is called the *Bunch-Davies* vacuum.

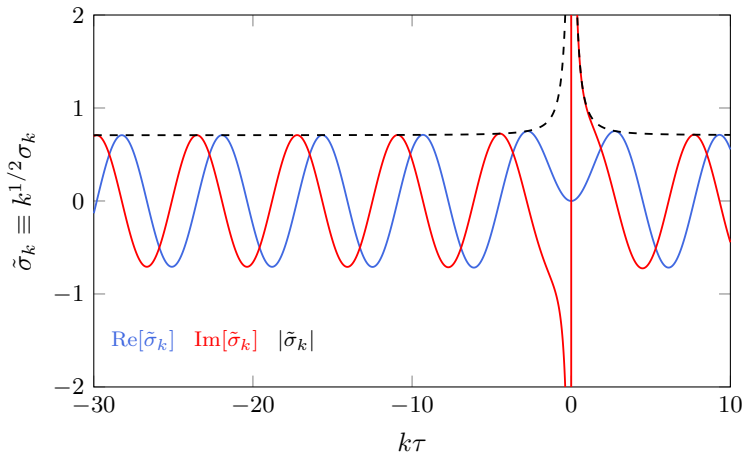


Figure 3.4: Plot of the final solution for the mode function as given by Eq. 3.5.29 in the rescaled form $\tilde{\sigma}_k \equiv k^{1/2}\sigma_k$. The real, imaginary and absolute parts are shows separately in different colors. The unimportant phase factor $\Delta^{(+)}$ is dropped for all plotted functions.

3.5.4 Solution to the Mode Function

Now the above initial condition can be used to determine the constants $c_{1,2}$ in Eq. 3.5.20. Consider the $|k\tau| \gg 1$ limit of the Hankel functions $H_\nu^{(1,2)}(-k\tau)$:

$$\lim_{|k\tau|\gg 1} H_\nu^{(1)} = \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{-k\tau}}e^{-ik\tau} \times \Delta^{(-)}, \qquad \lim_{|k\tau|\gg 1} H_\nu^{(2)} = \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{-k\tau}}e^{+ik\tau} \times \Delta^{(+)}. \tag{3.5.26}$$

Here, we defined $\Delta^{(\pm)} \equiv \exp(\pm i\pi(\nu + 1/2)/2$ and $\Delta^{(+)}\Delta^{(-)} = 1$. In this limit, the mode function can then be written as:

$$\lim_{|k\tau|\gg 1} \sigma_k(\tau) = \sqrt{-\tau}\left[c_1\sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{-k\tau}}e^{-ik\tau} \times \Delta^{(-)} + c_2\sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{-k\tau}}e^{+ik\tau} \times \Delta^{(+)}\right]. \tag{3.5.27}$$

In order to bring the above equation in agreement with the initial condition as given by Eq. 3.5.25, the constants are set to:

$$c_1 \equiv \frac{\sqrt{\pi}}{2} \times \Delta^{(+)}, \qquad c_2 \equiv 0. \tag{3.5.28}$$

Therefore, the appropriate mode function σ_k in line with the initial condition reads:

$$\boxed{\sigma_k(\tau) = \sqrt{-\tau}\left[\frac{\sqrt{\pi}}{2}\Delta^{(+)}\right] H_{\nu_\chi}^{(1)}(-k\tau).} \tag{3.5.29}$$

The final form for the mode function as given by the above equation is plotted in Fig. 3.4 (in the slightly rescaled form $\tilde{\sigma}_k = k^{1/2}\sigma_k$). Notice that $|\tilde{\sigma}_k|$ is constant or growing for all negative values of $k\tau$. This implies that, irrespective of the considered k-mode, the quantum fluctuation in the field $\delta\chi_k \equiv a\sigma_k$ grows as time evolves until $k\tau = 0$. During inflation, $a \simeq e^{\mathcal{H}\tau}$ and hence the quantum fluctuation grows exponentially fast. Therefore, quantum fluctuations are indeed stretched to cosmological scales very rapidly.

3.5.5 Canonical Quantization of the Fluctuation Field

Up till now, the (rescaled) field perturbation $\sigma(\boldsymbol{x}, t) \equiv a\delta\chi(\boldsymbol{x}, t)$ is regarded as a *classical* quantity, whereas it arises from purely quantum mechanical effects. In order to examine the consequences of σ being caused by quantum effects, rather than classical effects, we will quantize the field according to the method of canonical quantization.

To quantize the field conform the method of canonical quantization, we first define the conjugate momentum π to the field σ as:

$$\pi \equiv \frac{\partial \mathcal{L}_\sigma(\sigma, \sigma')}{\partial \sigma'} = \sigma', \tag{3.5.30}$$

where $\mathcal{L}_\sigma$ is easily inferred from the action given in Eq. 3.5.13. We promote the classical fields σ and π to quantum operator fields, denoted with hats:

$$\sigma \longrightarrow \hat{\sigma}, \quad \pi \longrightarrow \hat{\pi}. \tag{3.5.31}$$

We impose equal time canonical commutation relations (CCR) on the fields by means of the commutator:

$$[\hat{\sigma}(\tau, \boldsymbol{x}), \hat{\pi}(\tau, \boldsymbol{y})] \equiv i\delta^{(3)}(\boldsymbol{x} - \boldsymbol{y}). \tag{3.5.32}$$

The above commutator reflects on *locality*: modes at different spatial locations ($\boldsymbol{x}$ and $\boldsymbol{y}$) evolve independently and hence the corresponding operators commute. In Fourier space, the CCR condition becomes:

$$[\hat{\sigma}_{\boldsymbol{k}}(\tau), \hat{\pi}_{\boldsymbol{k}'}(\tau)] = i\delta^{(3)}(\boldsymbol{k} + \boldsymbol{k}'). \tag{3.5.33}$$

The operator $\hat{\sigma}_{\boldsymbol{k}}$ in Fourier space can be expanded in terms of a single time-independent operator $\hat{a}_{\boldsymbol{k}}$, its Hermitian conjugate and the complex mode function $\sigma_k(\tau)$ as follows [14]:

$$\hat{\sigma}_{\boldsymbol{k}} = \sigma_k(\tau)\hat{a}_{\boldsymbol{k}} + \sigma_k^*(\tau)\hat{a}_{\boldsymbol{k}}^\dagger. \tag{3.5.34}$$

Here, the operators $\hat{a}_{\boldsymbol{k}}$ and $\hat{a}_{\boldsymbol{k}}^\dagger$ may be regarded as annihilation and creation operators. They are defined to act on the vacuum state of the field, denoted by $|0\rangle$, in the following way [14]:

$$\hat{a}_{\boldsymbol{k}}|0\rangle = \langle 0|\hat{a}_{\boldsymbol{k}}^\dagger = 0, \tag{3.5.35}$$

and excited states are generated by multiple applications of the raising operator [14]:

$$|m_{\boldsymbol{k}_1}, n_{\boldsymbol{k}_2, \cdots}\rangle = \frac{1}{\sqrt{m!n!\cdots}} \left[(\hat{a}_{\boldsymbol{k}_1}^\dagger)^m (\hat{a}_{\boldsymbol{k}_2}^\dagger)^n \cdots \right] |0\rangle, \tag{3.5.36}$$

where $m_{\boldsymbol{k}_1}$ denotes the number of particles with momentum $\boldsymbol{k}_1$. The square root prefactor accounts for proper normalization.

The creation and annihilation operators satisfy the following commutation relations:

$$[\hat{a}_{\boldsymbol{k}}, \hat{a}_{\boldsymbol{q}}] = [\hat{a}_{\boldsymbol{k}}^\dagger, \hat{a}_{\boldsymbol{q}}^\dagger] = 0, \qquad [\hat{a}_{\boldsymbol{k}}, \hat{a}_{\boldsymbol{q}}^\dagger] = [\hat{a}_{\boldsymbol{q}}, \hat{a}_{\boldsymbol{k}}^\dagger] = (2\pi)^3\delta^{(3)}(\boldsymbol{k} + \boldsymbol{q}), \tag{3.5.37}$$

where the factor of $(2\pi)^3$ is included due to the obeyed Fourier convention. In terms of the creation and annihilation operators ($\hat{a}_{\boldsymbol{k}}$, $\hat{a}_{\boldsymbol{k}}^\dagger$), the field $\hat{f}$ and its conjugate momentum $\hat{\pi}$ read:

$$\hat{\sigma}(\tau, \boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \left[\sigma_k(\tau)\hat{a}_{\boldsymbol{k}} + \sigma_k^*(\tau)\hat{a}_{\boldsymbol{k}}^\dagger \right] e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{3.5.38}$$

$$\hat{\pi}(\tau, \boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \left[\sigma_k'(\tau)\hat{a}_{\boldsymbol{k}} + (\sigma_k^*(\tau))'\hat{a}_{\boldsymbol{k}}^\dagger \right] e^{i\boldsymbol{k}\cdot\boldsymbol{x}}. \tag{3.5.39}$$

For the above definitions of the operator fields $\hat{\sigma}(\tau, \boldsymbol{x})$ and $\hat{\pi}(\tau, \boldsymbol{x})$ to satisfy the canonical commutation relations (CCR), the following constraint on the field mode σ_k, its temporal derivative σ_k' and the complex conjugates of those should be enforced:

$$W[\sigma_k, \sigma_k'] = (-i)\left[\sigma_k(\sigma_{-k}^*)' - \sigma_k^* \sigma_{-k}'\right] \equiv 1. \tag{3.5.40}$$

The function $W[\sigma_k, \sigma_k^*]$ is called the Wroskian and the above condition is referred to as the Wroskian normalization condition.

3.5.6 Correlation Function

Now that we constructed the explicit solution to the mode function $\sigma_k(\tau)$ and performed canonical quantization, we can compute the two-point correlation function and its Fourier-equivalent, the power spectrum.

In section 3.1, we stated that the mean, i.e. the one-point correlation function $\langle 0|\hat{f}|0\rangle$, of the quantum fluctuations in the inflaton field vanishes due to their random nature. In contrast to the mean, the variance or two-point correlation function $\langle 0||\hat{f}|^2|0\rangle$ is non-zero, which will be shown explicitly in section 5.3. Here, we will prove that these statements also hold for the field fluctuations in the toy model. In addition, it will be shown that the power spectrum (Eq. 3.2.34) arises naturally from the computation of the two-point correlation function.

First, we consider the average $\langle\hat{\sigma}\rangle$ or one-point correlation function in the vacuum state $|0\rangle$. Using the fact that $\hat{a}_{\boldsymbol{k}}$ annihilates the vacuum, it is straightforward to show that the average vanishes:

$$\boxed{\langle\hat{\sigma}\rangle \equiv \langle 0|\hat{\sigma}|0\rangle = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \langle 0|\left[\sigma_k(\tau)\hat{a}_{\boldsymbol{k}} + \sigma_k^*(\tau)\hat{a}_{\boldsymbol{k}}^\dagger\right]|0\rangle\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = 0,} \tag{3.5.41}$$

where the contractions show explicitly how the terms vanish: $\hat{a}_{\boldsymbol{k}}^\dagger$ annihilates $\langle 0|$ and $\hat{a}_{\boldsymbol{k}}$ has the same effect but on $|0\rangle$. Hence, we have shown that the 1-point correlation function vanishes.

Now we consider the 2-point correlation function:

$$\langle|\hat{\sigma}|^2\rangle \equiv \langle 0|\hat{\sigma}^\dagger(\tau, \boldsymbol{x})\hat{\sigma}(\tau, \boldsymbol{x})|0\rangle. \tag{3.5.42}$$

In terms of creation and annihilation operators, the 2-point correlation function can be written as:

$$\begin{aligned}
\langle|\hat{\sigma}|^2\rangle &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \int \frac{d^3\boldsymbol{q}}{(2\pi)^3} \langle 0|(\sigma_k^*\hat{a}_{\boldsymbol{k}}^\dagger + \sigma_k\hat{a}_{\boldsymbol{k}})(\sigma_q\hat{a}_{\boldsymbol{q}} + \sigma_q^*\hat{a}_{\boldsymbol{q}}^\dagger)|0\rangle \\
&= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \int \frac{d^3\boldsymbol{q}}{(2\pi)^3} \langle 0|\sigma_k\sigma_q^*\hat{a}_{\boldsymbol{k}}\hat{a}_{\boldsymbol{q}}^\dagger|0\rangle \\
&= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \int \frac{d^3\boldsymbol{q}}{(2\pi)^3} \sigma_k\sigma_q^* \langle 0|\hat{a}_{\boldsymbol{k}}\hat{a}_{\boldsymbol{q}}^\dagger|0\rangle,
\end{aligned} \tag{3.5.43}$$

where the contractions show which terms vanish. The commutation relation for the creation and annihilation operators in Fourier space will be used to rewrite the expectation value $\langle 0|\hat{a}_{\boldsymbol{k}}\hat{a}_{\boldsymbol{q}}^\dagger|0\rangle$:

$$\langle 0|[\hat{a}_{\boldsymbol{k}}, \hat{a}_{\boldsymbol{q}}^\dagger]|0\rangle = \langle 0|\hat{a}_{\boldsymbol{k}}\hat{a}_{\boldsymbol{q}}^\dagger|0\rangle - \langle 0|\hat{a}_{\boldsymbol{q}}^\dagger\hat{a}_{\boldsymbol{k}}|0\rangle = \langle 0|\hat{a}_{\boldsymbol{k}}\hat{a}_{\boldsymbol{q}}^\dagger|0\rangle. \tag{3.5.44}$$

Notice that the second term after the first equality sign vanishes due to the contractions indicated. On account of this result, the 2-point correlation function can be written as follows:

$$\begin{aligned}\langle|\hat{\sigma}|^2\rangle &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\int \frac{d^3\boldsymbol{q}}{(2\pi)^3}\sigma_k\sigma_q^*\langle 0|[\hat{a}_{\boldsymbol{k}},\hat{a}^\dagger_{\boldsymbol{q}}]|0\rangle \\ &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\int \frac{d^3\boldsymbol{q}}{(2\pi)^3}\sigma_k\sigma_q^* \times (2\pi)^3\delta^{(3)}(\boldsymbol{k}+\boldsymbol{q}) \\ &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}|\sigma_k|^2 = \int d\ln k\, \mathcal{P}_\sigma. \end{aligned} \tag{3.5.45}$$

In going from the first to the second line, we used the expression for the commutator (Eq. 3.5.37) in terms of the 3-momentum delta function. In going to the third, the delta function $\delta^{(3)}(\boldsymbol{k}+\boldsymbol{q})$ is used to evaluate the $\boldsymbol{q}$-integral. From the last line it follows that the 2-point correlation function of the fluctuation field $\hat{\sigma}$ is proportional to the square of the Fourier mode function. In the last equality, the power spectrum for σ is defined as:

$$\boxed{\mathcal{P}_\sigma \equiv \frac{k^3}{2\pi^2}|\sigma_k|^2,} \tag{3.5.46}$$

in accordance with result presented in Eq. 3.2.34.

3.5.7 Power Spectrum

Now the power spectrum of fluctuations in the field $\chi(\boldsymbol{x},t)$ can be computed. Since we know that $\delta\chi$ freezes on super-horizon scales, a convenient moment to compute their power spectrum is after horizon exit, corresponding to $|k\tau|\to 0$, since after horizon exit the fluctuation modes do not evolve appreciably anymore.

In the limit $|k\tau|\to 0$, the Hankel function of the first kind can be written as:

$$\lim_{|k\tau|\to 0} H_\nu^{(1)}(-k\tau) = \sqrt{\frac{2}{\pi}}e^{-i\pi/2}\, 2^{\nu-3/2}\frac{\Gamma(\nu)}{\Gamma(3/2)}(-k\tau)^{-\nu}. \tag{3.5.47}$$

So that the mode function can be written as:

$$\lim_{|k\tau|\to 0} \sigma_k(\tau) = 2^{\nu-3/2}\Delta^{(+)}e^{-i\pi/2}\frac{\Gamma(\nu)}{\Gamma(3/2)}\frac{1}{\sqrt{2k}}(-k\tau)^{1/2-\nu_\chi}. \tag{3.5.48}$$

Using the fact that during inflation $\tau = -1/aH$, the absolute value of the mode function reads:

$$\lim_{|k\tau|\to 0} |\sigma_k| = \frac{C(\nu)}{\sqrt{2k}}\left(\frac{k}{aH}\right)^{1/2-\nu_\chi}, \qquad C(\nu) \equiv 2^{\nu_\chi-3/2}\frac{\Gamma(\nu)}{\Gamma(3/2)} \to 1. \tag{3.5.49}$$

In the limit $\nu_\chi \to 3/2$ (i.e. for ε, $\eta_\chi \ll 1$), the constant $C(\nu)$ approaches unity. Going back to the unscaled fluctuation $\delta\chi(\boldsymbol{x},t) = \sigma(\boldsymbol{x},t)/a(t)$ and setting the constant $C(\nu)$ to unity, the power spectrum $\mathcal{P}_{\delta\chi}$ can be written as:

$$\boxed{\mathcal{P}_{\delta\chi} = \frac{k^3}{2\pi^2}|\delta\chi_k|^2 = \left(\frac{H}{2\pi}\right)^2\left(\frac{k}{aH}\right)^{3-2\nu_\chi}.} \tag{3.5.50}$$

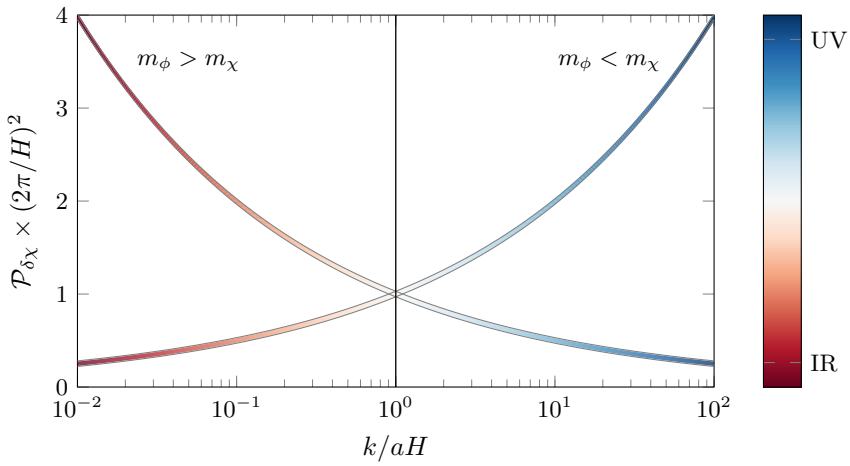


Figure 3.5: Illustration of the power spectrum for $V(\chi)$ and $V(\phi)$ both quadratic. For $m_\phi > m_\chi$ ($m_\phi < m_\chi$) the spectrum is red (blue). The spectral index is set to $n_\chi = \pm 0.3$ and the Hubble parameter is set to unity: $H \equiv 1$.

The amplitude $A_{\delta\chi}$ and spectral index n_χ for the spectrum, in analogy with the parametrization of Eq. 3.4.2, are given by:

$$A_{\delta\chi} \equiv \left(\frac{H}{2\pi}\right)^2, \qquad n_\chi - 1 \equiv \frac{d\ln \mathcal{P}_{\delta\chi}}{d\ln k} = 2\eta_\chi - 2\varepsilon. \tag{3.5.51}$$

Because of the smallness of the parameters ε (by definition of a quasi De-Sitter stage) and η_χ (by assumption of $m_\chi^2 \ll H^2$), the spectrum $\mathcal{P}_{\delta\chi}$ is nearly scale-invariant, i.e. independent of the wavenumber k. The small tilt in the power spectrum arises because (a) the field χ is assumed to be massive and (b) the Hubble parameter is not exactly constant over time during inflation (as quantified by the small but not-zero value for ε). For a massless scalar field in a pure De-Sitter background, the spectrum has amplitude $(H/2\pi)^2$ and is completely scale independent.

Conform traditional terminology, a spectrum with index $n_\chi > 1$ is said to be *blue*, since the power increases towards the ultraviolet regime ($k/aH \gg 1$). Conversely, for $n_\chi < 1$ the spectrum is said to be *red* as the power is higher in the infrared ($k/aH \ll 1$). This is easily illustrated by choosing a simple quadratic inflaton potential:

$$V(\phi) = \frac{1}{2} m_\phi^2 \phi^2. \tag{3.5.52}$$

In the slow-roll regime, the parameter ε can be replaced by the slow roll parameter $\epsilon_{\rm v}$ and the potential is related to H via $V(\phi) = 3M_{\rm pl}^2 H^2$. Using this relation, the slow roll parameter $\epsilon_{\rm v}$ can be computed to be $\epsilon_{\rm v} = m_\phi^2/3H^2$ (conform Eq. 2.7.5). The spectral index n_χ is then given by:

$$n_\chi - 1 \equiv 2\eta_\chi - 2\varepsilon = \frac{2}{3H^2}(m_\chi^2 - m_\phi^2). \tag{3.5.53}$$

Hence, for $m_\phi > m_\chi$ ($m_\phi < m_\chi$) we find that $n_\chi < 1$ ($n_\chi > 1$) and the spectrum is red (blue). In Fig. 3.5, the power spectrum is plotted in case the choice $n_\chi = \pm 0.3$ for illustrational purposes. The color scheme indicates the UV and IR regimes.

3.5.8 Quantum to Classical Transition

We conclude this section by discussing the quantum-to-classical transition of fluctuations in the toy model. In section 3.1, it was argued that fluctuations in the inflaton lose their quantum nature on super-horizon scales and hence $\delta\phi$ can be regarded as a classical stochastic field after horizon exit. Here, we will formalize this statement for fluctuations in the toy model field χ.

In the super-horizon limit, $\sigma_k(\tau)$ and $\sigma_k'(\tau)$ can be written as:

$$\lim_{|k\tau|\to 0} \sigma_k(\tau) = -\frac{1}{\sqrt{2k^3}}\frac{i}{\tau}, \qquad \lim_{|k\tau|\to 0} \sigma_k'(\tau) = \frac{1}{\sqrt{2k^3}}\frac{i}{\tau^2}. \tag{3.5.54}$$

In this limit, the field operator $\hat{\sigma}$ and its momentum conjugate $\hat{\pi}$ become:

$$\hat{\sigma}(\tau, \boldsymbol{x}) = -\frac{i}{\tau}\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\frac{1}{\sqrt{2k^3}}\left[\hat{a}_{\boldsymbol{k}} - \hat{a}^\dagger_{\boldsymbol{k}}\right] e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{3.5.55}$$

$$\hat{\pi}(\tau, \boldsymbol{x}) = \frac{i}{\tau^2}\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\frac{1}{\sqrt{2k^3}}\left[\hat{a}_{\boldsymbol{k}} - \hat{a}^\dagger_{\boldsymbol{k}}\right] e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = -\frac{1}{\tau}\hat{f}(\tau, \boldsymbol{x}). \tag{3.5.56}$$

Notice that the two operators become *proportional* to each other on super-horizon scales. Hence, the canonical commutation relation (CCR) Eq. 3.5.32 vanishes:

$$\boxed{\left[\hat{\sigma}(\tau, \boldsymbol{x}), \hat{\pi}(\tau, \boldsymbol{x})\right] \xrightarrow{|k\tau|\to 0} 0.} \tag{3.5.57}$$

The vanishing commutator implies that the field loses its quantum nature on super-horizon scales and hence it can be identified with a classical field [14, 15].

Chapter 4

Cosmological Perturbation Theory

"The role of the infinitely small in nature is infinitely great."

— Louis Pasteur

In order to study perturbations around a perfectly homogeneous and isotropic background universe, the evolution equations governing their dynamics are required. As long as the perturbations remain small relative to the corresponding background quantities, they can be studied within the context of (first order) cosmological perturbation theory (CPT). In this chapter, the formalism used to derive the evolution equations for perturbations around the FRW background universe will be introduced in detail. For extensive literature on CPT, we refer the reader to Refs. [65, 69], for more pedagogical introductions to the subject, see [14, 15, 25, 74].

The approach in this chapter will be two-fold. First of all, we will derive the evolution equations for perturbations in a single fluid by perturbing the background Einstien field equations (EFE's). The cases of a perfect fluid specified at background level by energy density ρ and pressure P, a scalar field and non-perfect fluid which generates anisotropic stress at first order in perturbations will be considered.[1] Secondly, we will extend the analysis to multiple (perfect) fluids and in particular multiple scalar fields, as they will be useful in the context the multi-field inflationary scenario.

In addition to the perturbed EFE's, any complementary evolution equations that describe the dynamics of the fluid(s) should be perturbed to first order as well. In particular, this applies to the Klein-Gordon equation (Eq. 2.5.2), describing the evolution of inflaton field (in the single field scenario). Hence, we will also perturb this evolution equation to first order. The resulting set of equations can then be used in the next chapter to obtain a single equation of motion for the inflaton fluctuation, which is known as the Mukhanov-Sasaki equation (Eq. 5.2.14). In the analysis, we will find that from the set of perturbed equation, it follows that perturbations are adiabatic for a single scalar field.

As this chapter lies at the core of the upcoming chapters and is therefore extensive and detailed, we will first provide an outline of the content to be treated in this chapter in the next section.

[1] As we will discuss in detail in section 4.8.1, the Cosmological Principle constrains any background fluid to behave as a perfect fluid, so that deviations from a perfect fluid start at first order in perturbation.

4.1 Outline and Preliminaries

To start, we will first outline the specific aim and strategy employed in this chapter. Subsequently, we will introduce a number of technical subtleties, to be discussed in more detail in the upcoming sections. Finally, we review the main results on the background evolution of the FRW universe. This is done (a) to collect all results of the preceding chapters relevant for this chapter and (b) to get familiar with the usage of conformal time as the evolution variable.

4.1.1 Outline

Before going into the detailed mathematical derivations, it proves useful to describe the primary aim and corresponding approach of this chapter. The goal of this chapter will be to derive the equations governing the evolution of perturbations around a homogeneous FRW background for a generic perfect fluid and the specific case of the inflaton scalar field. To obtain those evolution equations, we perturb the EFE's to first order as follows:

$$\overline{G}_{\mu\nu} + \delta G_{\mu\nu} = 8\pi G \left[\overline{T}_{\mu\nu} + \delta T_{\mu\nu}\right], \tag{4.1.1}$$

where the overlined quantities correspond to the zeroth-order or background variables and the first order perturbations are denoted using the δ-notation. As the perturbed EFE's should hold at all orders seperately we can extract the first order equation as:

$$\delta G_{\mu\nu} = 8\pi G\, \delta T_{\mu\nu}. \tag{4.1.2}$$

The left-hand-side of the perturbed EFE's (i.e. the perturbed Einstein tensor $\delta G_{\mu\nu}$) is independent of the considered energy and matter source and can therefore be computed around an FRW background, regardless of the considered constituents filling the FRW universe. In the first sections of this chapter, the perturbed Einstein tensor will be constructed. This is done in a number of steps. First of all, the metric tensor will be perturbed to first order:

$$g_{\mu\nu}(t, \boldsymbol{x}) = \bar{g}_{\mu\nu}(t) + \delta g_{\mu\nu}(t, \boldsymbol{x}), \tag{4.1.3}$$

and the chosen gauge will be introduced. Then, the expressions for the perturbed connections or Christoffel symbols will be constructed:

$$\Gamma^{\sigma}_{\mu\nu}(t, \boldsymbol{x}) = \bar{\Gamma}^{\sigma}_{\mu\nu}(t) + \delta\Gamma^{\sigma}_{\mu\nu}(t, \boldsymbol{x}). \tag{4.1.4}$$

Subsequently, the perturbed Ricci tensor $\delta R_{\mu\nu}$ and scalar δR will be derived. Together with the first order correction to the metric $\delta g_{\mu\nu}$, they can be used to construct $\delta G_{\mu\nu}$ as follows:

$$\delta G_{\mu\nu} = \delta R_{\mu\nu} + \frac{1}{2}\delta g_{\mu\nu} R + \frac{1}{2} g_{\mu\nu}\delta R. \tag{4.1.5}$$

To find the complete first order perturbed EFE's the right-hand-side, containing the perturbed energy momentum tensor $\delta T_{\mu\nu}$, should be computed as well. As mentioned before, this will be done explicitly for a generic perfect fluid (characterized at background level by an energy density $\bar{\rho}$ and pressure $\bar{P}$) and the inflaton scalar field.

4.1.2 SVT Decomposition and the Gauge Problem

The outline described above comes with a number of technical subtleties, which we will introduce here and treat in more detail in the upcoming sections. First of all, as described in section 3.1, quantum fluctuations in the inflaton causes perturbations in the metric tensor and energy-momentum tensor. That is, a scalar perturbation ($\delta\phi$) induces tensor perturbations. As a result, the perturbed EFE's will constitute a set of coupled differential equations, coupling the evolution of scalar quantities to vector and tensor quantities, and are therefore difficult to solve.[2] To make the perturbed EFE's manageable a procedure called scalar-vector-tensor (SVT) decomposition will be performed. By means of the SVT decomposition a generic perturbation can be decomposed into scalar, vector and tensor components.[3] This decomposition is powerful since at linear order it can be proven that the scalar, vector and tensor components evolve independently. In other words, the SVT decomposition reduces the coupled EFE's to uncoupled differential equations for the scalar, vector and tensors components. Hence, one can study scalar perturbations and ignore the (possible) existence of vector and tensor perturbations (and vice versa). This simplifies the mathematical analysis in cosmological perturbation theory considerably.

The second technical subtlety to be addressed is the *gauge* problem. In defining the perturbations around the FRW metric, one implicitly chooses a specific coordinate system, also known as the *gauge*. The perturbations are therefore not uniquely defined. By switching to a different gauge, the values of the perturbations may change or fictitious perturbations may even be generated that are solely caused by the gauge transformation. This issue is known as the gauge problem and can be solved by working consistently in a completely specified gauge or by constructing gauge-invariant perturbations.

4.1.3 FRW Background

Here, we will review the relevant background or zeroth order results for the FRW universe. The unperturbed differential line element $d\bar{s}^2$ reads:

$$d\bar{s}^2 = a^2(\tau)\Big[-d\tau^2 + \delta_{ij}dx^i dx^j\Big], \tag{4.1.6}$$

using conformal time $d\tau = dt/a$ as the evolution variable. The corresponding metric is diagonal and the non-vanishing components are given by:

$$\bar{g}_{00} = -a^2(\tau), \qquad \bar{g}_{ij} = a^2(\tau)\delta_{ij}. \tag{4.1.7}$$

The Christoffel symbols for the FRW background are given by:

$$\bar{\Gamma}^0_{00} = \mathcal{H}, \qquad \bar{\Gamma}^i_{0j} = \mathcal{H}\delta^i_j, \qquad \bar{\Gamma}^0_{ij} = \mathcal{H}\delta_{ij}, \tag{4.1.8}$$

and the other Christoffel symbols are zero or related to the above by symmetry in the lower pair of indices. The diagonal Ricci tensor and scalar are given by:

$$\bar{R}_{00} = -3\frac{a''}{a} + 3\mathcal{H}^2, \qquad \bar{R}_{ij} = \Big(\frac{a''}{a} + \mathcal{H}^2\Big)\delta_{ij}, \qquad \bar{R} = \frac{6}{a^2}\frac{a''}{a}. \tag{4.1.9}$$

[2] Although we already advocated that vector perturbations are not generated in the single field inflationary scenario, we will consider them here for the sake of completeness.

[3] For vectors, the SVT decomposition is known as *Helmholtz's Theorem*, which states that a generic vector field A^i can always be decomposed into the divergence of a scalar A and divergenceless vector $\hat{A}^i$ as follows: $A^i = \partial^i A + \hat{A}^i$, where $\hat{A}^i$ is subject to the constraint $\partial_i \hat{A}^i = 0$.

The Einstein tensor is diagonal as well, since it is composed of the diagonal Ricci and metric tensors, and reads:

$$\bar{G}_{00} = 3\mathcal{H}^2, \qquad \bar{G}_{ij} = \Big(\mathcal{H}^2 - 2\frac{a''}{a}\Big)\delta_{ij}. \tag{4.1.10}$$

The two Friedmann and continuity equations, governing the evolution of the background Hubble parameter $\mathcal{H}$ and the energy density $\bar{\rho}$ and $\bar{P}$ are given by:

$$\mathcal{H}^2 = \frac{\bar{\rho}a^2}{3M_{\rm pl}^2}, \tag{4.1.11}$$

$$\mathcal{H}' = -\frac{a^2}{6M_{\rm pl}^2}(\bar{\rho} + 3\bar{P}), \tag{4.1.12}$$

$$\bar{\rho}' = -3\mathcal{H}(\bar{\rho} + \bar{P}). \tag{4.1.13}$$

In addition, we define the equation of state and sound speed as $w \equiv \bar{P}/\bar{\rho}$ and $c_s^2 \equiv \bar{P}'/\bar{\rho}'$, respectively. Using the above equations, one can find the following constraint equations for w and c_s^2:

$$\frac{\mathcal{H}}{\mathcal{H}^2} = -\frac{1}{2}(1+3w), \qquad \frac{w'}{w+1} = 3\mathcal{H}(w - c_s^2). \tag{4.1.14}$$

Note on Notation

Up till now, we have consistently denoted zeroth order (background) and first-order contributions with the bar- and δ-notation, respectively. However, in the remaining part of this chapter, the explicit distinction between the zeroth order and first order contributions will not be made always and it should be evident from the context what the order of the quantity is. Typically, perturbations will still be denoted with the δ-notation, whereas the bar over the background variables will be often be omitted. To illustrate this ignorance, in case the variables ρ and $\delta\rho$ appear simultaneously in the same expression, ρ will always refer to the background energy density and $\delta\rho$ denotes the first order perturbation.

4.2 Scalar-Vector-Tensor Decomposition

In section 3.1, we discussed globally how quantum fluctuations in the inflaton field causes perturbations in, for instance, the metric $\delta g_{\mu\nu}$ and the energy-momentum tensor $\delta T_{\mu\nu}$ via the equations of motion (EFE's). That is, we found that a scalar perturbation ($\delta\phi$) induces perturbations in tensor quantities (e.g. $\delta g_{\mu\nu}$ and $\delta T_{\mu\nu}$).[4]

In this section, it will be shown how generic perturbations can be decomposed into scalar, vector and tensor (SVT) components: this is called the SVT decomposition. The reason for SVT decomposition is that the scalar, vector and tensor components evolve independently. More formally, the SVT approach reduces the (linearly perturbed) Einstein equations, which are coupled differential equations, to a set of uncoupled differential equations. That is, one can study scalar perturbations and ignore the existence of tensor perturbations and vice versa. This simplifies mathematical computations in cosmological perturbation theory considerably. In Appendix D.3, we prove that SVT components indeed evolve independently at first order.

[4]Furthermore, it was argued that inflation does not produce vector perturbations and even if they were produced, they would rapidly decay away with the expansion. See Appendix D.2

The SVT decomposition is based on the so-called Helmholtz decomposition theorem for 3-vectors. Hence, this theorem will be discussed first. Then, the SVT decomposition will be demonstrated for a generic scalar α, vector β_i and traceless symmetric tensor γ_{ij} ($\gamma_{ii} = 0$, $\gamma_{ij} = \gamma_{ji}$). First, the decomposition is performed in real space. Subsequently, the decomposition in Fourier space is also given. The review of SVT decomposition given here is partially based on Refs. [15, 21].

4.2.1 Helmholtz Decomposition Theorem

As mentioned above, the SVT decomposition is based on the Helmholtz decomposition theorem for 3-vectors. This theorem states that any 3-vector field $\boldsymbol{\omega}(\boldsymbol{x})$ can be decomposed into longitudinal (parallel) and transverse (perpendicular) components. That is, $\boldsymbol{\omega}$ can be written as:

$$\boldsymbol{\omega} = \boldsymbol{\omega}^{\parallel} + \boldsymbol{\omega}^{\perp}, \tag{4.2.1}$$

where $\boldsymbol{\omega}^{\parallel}$ and $\boldsymbol{\omega}^{\perp}$ are the longitudinal and transverse components, respectively. The terminology *longitudinal* and *transverse* originates from the fact that in Fourier space $\boldsymbol{\omega}^{\parallel}$ is parallel to the wave-vector $\boldsymbol{k}$ and $\boldsymbol{\omega}^{\perp}$ is perpendicular to $\boldsymbol{k}$. By construction, the longitudinal and transverse parts are curl- and divergence-free:

$$\nabla \times \boldsymbol{\omega}^{\parallel} = \nabla \cdot \boldsymbol{\omega}^{\perp} = 0. \tag{4.2.2}$$

The decomposition into longitudinal and parallel components always exist, but is not unique. That is, one can always add a constant to $\omega_i^{\parallel}$ or $\omega_i^{\perp}$.

In index notation, the above constraints on $\boldsymbol{\omega}^{\parallel}$ and $\boldsymbol{\omega}^{\perp}$ can be written as:

$$\nabla \cdot \boldsymbol{\omega}^{\perp} = \delta^{ij}\partial_i\omega_j^{\perp} = \partial^j\omega_j^{\perp} = 0 \qquad \text{and} \qquad (\nabla \times \boldsymbol{\omega}^{\parallel})_i = \varepsilon_{ijk}\partial_j\omega_k^{\parallel} = 0. \tag{4.2.3}$$

where ε_{ijk} is the Levi-Cevita symbol.[5] On account of the divergence constraint $\nabla \cdot \boldsymbol{\omega}^{\perp} = 0$, the parallel component can be written as the gradient of a scalar:

$$\omega_i^{\parallel} = \partial_i\omega, \tag{4.2.4}$$

for some scalar ω. Hence, the generic 3-vector $\boldsymbol{\omega}$ can be written in terms of the gradient of a scalar and a divergence-free vector:

$$\omega_i = \partial_i\omega + \hat{\omega}_i^{\perp}, \tag{4.2.5}$$

where the vector is divergence-free, i.e. $\partial^i\hat{\omega}_i^{\perp} = 0$. We will adopt the convention that hatted quantities are divergence-free.

4.2.2 Real Space Decomposition

Now that the general principle behind SVT decomposition is introduced, the decomposition can be performed for the generic scalar α, vector β_i and traceless symmetric tensor γ_{ij}

[5] The Levi-Cevita symbol ε_{ijk} is defined to be 0 for $i = j = k$, $+1$ for an even permutation of (i, j, k) and -1 for an odd permutation of (i, j, k).

($\gamma_{ii} = 0$). Obviously, the scalar α cannot be decomposed into different components. Hence, the decomposition of α reads:

$$\alpha = \alpha^{\mathrm{S}}, \tag{4.2.6}$$

where the superscript S indicates that α is 'decomposed' into only a scalar component. On account of Helmholtz theorem, the vector β_i can be decomposed into the gradient of a scalar and a divergenceless vector:

$$\beta_i = \beta_i^{\mathrm{S}} + \hat{\beta}_i^{\mathrm{V}}. \tag{4.2.7}$$

Here, $\beta_i^{\mathrm{S}} = \partial_i \beta$ is the gradient of the scalar β and $\hat{\beta}_i^{\mathrm{V}}$ is the divergenceless vector, i.e. $\partial^i \hat{\beta}_i^{\mathrm{V}} = 0$.

Similarly, the traceless tensor γ_{ij} can be decomposed into scalar, vector and tensor components as:

$$\gamma_{ij} = \gamma_{ij}^{\mathrm{S}} + \gamma_{ij}^{\mathrm{V}} + \gamma_{ij}^{\mathrm{T}}, \tag{4.2.8}$$

where the tensor component is divergenceless and satisfies: $\partial^i \hat{\gamma}_{ij}^{\mathrm{T}} = 0$. In this case, both indices i and j are separately decomposed into longitudinal and transverse components. The scalar and vector components can therefore be written as:

$$\gamma_{ij}^{\mathrm{S}} = \left(\partial_i \partial_j - \frac{1}{3}\delta_{ij}\partial^2\right)\hat{\gamma} \equiv \partial_{\langle i}\partial_{j\rangle}\gamma, \tag{4.2.9}$$

$$\gamma_{ij}^{\mathrm{V}} = \frac{1}{2}(\partial_i \hat{\gamma}_j + \partial_j \hat{\gamma}_i) \equiv \partial_{(i}\hat{\gamma}_{j)}, \tag{4.2.10}$$

where the Laplacian ∂^2 equals $\partial_i \partial^i$ and the vector $\hat{\gamma}_j$ satisfies the constraint $\partial^i \hat{\gamma}_i = 0$. The purely tensorial part obeys the constraint:

$$\partial_i \hat{\gamma}_{ij}^{\mathrm{T}} = 0. \tag{4.2.11}$$

4.2.3 Fourier Space Decomposition

To go from real space to Fourier space, the partial derivative ∂_j is replaced by ik_j where i is the imaginary unit, so that the gradient transforms as $\partial^2 = \partial_i \partial^i \to -k^2$. In Fourier space the vector β_i is decomposed as:

$$\beta_i = \beta_i^{\mathrm{S}} + \beta_i^{\mathrm{V}} = ik_i\beta + \hat{\beta}_i^{\mathrm{V}} \equiv \frac{ik_i}{k}\tilde{\beta} + \hat{\beta}_i^{\mathrm{V}}, \tag{4.2.12}$$

where $\tilde{\beta} \equiv k\beta$ and $k^i \hat{\beta}_i^{\mathrm{V}} = 0$ as the vector contribution is divergence-free. In Fourier space, the traceless tensor SVT-components become:

$$\begin{aligned} \gamma_{ij}^{\mathrm{S}} &= \left(-k_i k_j + \frac{1}{3}\delta_{ij}k^2\right)\gamma = \left(-\frac{k_i k_j}{k^2} + \frac{1}{3}\delta_{ij}\right)\tilde{\gamma}, \\ \gamma_{ij}^{\mathrm{V}} &= \frac{i}{2}\left(k_i \hat{\gamma}_j + k_j \hat{\gamma}_i\right) = \frac{i}{2k}\left(k_i \tilde{\gamma}_j + k_j \tilde{\gamma}_i\right), \end{aligned} \tag{4.2.13}$$

where $\tilde{\gamma} \equiv k^2\gamma$ and $\tilde{\gamma}_i \equiv k\hat{\gamma}_i$, obeying the contraint $k^i \gamma_i = 0$. The Fourier-space constraint on the tensor component reads:

$$k^i \gamma_{ij}^{\mathrm{T}} = 0. \tag{4.2.14}$$

On account of rotational invariance of the background, we can choose $\boldsymbol{k}$ to be along the 3-axis without loss of generality so that $\boldsymbol{k} = (0, 0, k)$. Then $\gamma^{\mathrm{S}}_{ij}, \gamma^{\mathrm{V}}_{ij}$ and γ^{T}_{ij} can be written in matrix representation as:

$$\gamma^{\mathrm{S}}_{ij} = \frac{1}{3}\begin{bmatrix} \gamma & 0 & 0 \\ 0 & \gamma & 0 \\ 0 & 0 & -2\gamma \end{bmatrix}, \quad \gamma^{\mathrm{V}}_{ij} = \frac{i}{2}\begin{bmatrix} 0 & 0 & \gamma_1 \\ 0 & 0 & \gamma_2 \\ \gamma_1 & \gamma_2 & 0 \end{bmatrix}, \quad \gamma^{\mathrm{T}}_{ij} = \begin{bmatrix} \gamma^{\times} & \gamma^{+} & 0 \\ \gamma^{+} & -\gamma^{\times} & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{4.2.15}$$

4.3 The Perturbed Metric and Gauge Problem

In this section, we will write down the most general first perturbed metric and differential line element around a flat background FRW spacetime. Accordingly, the SVT decomposition of the introduced perturbations will be performed and the gauge problem will be discussed in more detail. Furthermore, our notation for the perturbed metric will be related to other conventions used in the literature.

The first order perturbed components of the metric tensor can be written as:

$$\delta g_{00} = -2a^2\Phi, \tag{4.3.1}$$

$$\delta g_{0i} = a^2 B_i, \tag{4.3.2}$$

$$\delta g_{ij} = -2a^2(\Psi\delta_{ij} - E_{ij}). \tag{4.3.3}$$

The scalar perturbations Φ and Ψ are formally called the lapse and spatial curvature perturbation. However, as mentioned earlier, they are also often referred to as gravitational potentials. This nomenclature originates from the fact that in the Newtonian gauge (to be introduced later in this section), Ψ and Φ act as gravitational potentials. The vector perturbation B_i is called the shift and finally E_{ij} is called the shear tensor, which is symmetric ($E_{ij} = E_{ji}$) and traceless ($\delta^{ij}E_{ij} = E^i_i = 0$). The perturbed line element is given by:

$$ds^2 = a^2(\tau)\Big[-(1+2\Phi)d\tau^2 + 2B_i dx^i d\tau + \big((1-2\Psi)\delta_{ij} + 2E_{ij}\big)dx^i dx^j\Big]. \tag{4.3.4}$$

On account of SVT decomposition, we can decompose the lapse vector B_i and shear tensor E_{ij}. The shift vector is decomposed as:

$$B_i = \partial_i B + \hat{B}_i, \tag{4.3.5}$$

where hatted quantities denote divergenceless quantities: $\partial^i \hat{B}_i = 0$. Since the shear tensor E_{ij} is traceless and symmetric, it can be decomposed in the same way as the tensor γ_{ij} considered in the previous section, that is:

$$E_{ij} = E^{\mathrm{S}}_{ij} + E^{\mathrm{V}}_{ij} + E^{\mathrm{T}}_{ij}. \tag{4.3.6}$$

The scalar, vector and tensor components can be written:

$$E^{\mathrm{S}}_{ij} = \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\partial^2\right)E \equiv \partial_{\langle i}\partial_{j\rangle}E, \tag{4.3.7}$$

$$E^{\mathrm{V}}_{ij} = \frac{1}{2}(\partial_i\hat{E}_j + \partial_j\hat{E}_i), \tag{4.3.8}$$

$$E^{\mathrm{T}}_{ij} = \hat{E}_{ij}, \tag{4.3.9}$$

where hatted quantities are again divergenceless, i.e. $\partial^i \hat{E}_i = \partial^i \hat{E}_{ij} = 0$. For convenience, the term in E^{S}_{ij} proportional to $\partial^2 E$ can be absorbed in the spatial curvature perturbation Ψ as follows:

$$\Psi + \frac{1}{3}\partial^2 E \to \Psi. \tag{4.3.10}$$

Then, the scalar contribution to the spatial perturbed metric, denoted $\delta g^{\mathrm{S}}_{ij}$, can be written as:

$$\delta g^{\mathrm{S}}_{ij} = -2a^2(\Psi\delta_{ij} - \partial_i\partial_j E). \tag{4.3.11}$$

4.3.1 The Gauge Problem

As we will show now, the perturbations in the metric come with an important technical subtlety. The perturbations in $\delta g_{\mu\nu}(\boldsymbol{x}, \tau)$ are *not* uniquely defined. That is to say, they are dependent on the choice of coordinates, also known as the *gauge* choice [14, 21, 60, 74]. Stated otherwise, in writing down the perturbed line element (or metric $\delta g_{\mu\nu}$), we have chosen a specific time slicing and defined specific spatial coordinates on these time slices (the threading). Switching to a different coordinate system:

$$x^\mu \to \tilde{x}^\mu = x^\mu + \xi^\mu(\tau, \boldsymbol{x}) \tag{4.3.12}$$

may change the values of the perturbed quantities or may even generate fictitious perturbations that are solely caused by the coordinate transformation and therefore have *no* physical meaning. This issue is known as the *gauge* problem.

To illustrate the possible generation of fictitious perturbations due to transformations between different gauges, we will give two examples.

▷ *FRW Space-Time and Spatial Translations.*—Consider the flat background FRW space-time, described by the line element in Eq. 4.1.6. We will make a change of spatial coordinates $\boldsymbol{x} \to \tilde{\boldsymbol{x}}$ quantified by the spatial shift $\boldsymbol{\xi}(\boldsymbol{x}, \tau)$ as follows:

$$\tilde{\boldsymbol{x}} = \boldsymbol{x} + \boldsymbol{\xi}(\boldsymbol{x}, \tau), \tag{4.3.13}$$

where the shift is taken to be small so that it can be treated as a first order perturbation. Taking the differential of the above transformation equation yields:

$$dx^i = d\tilde{x}^i - \partial_\tau \xi^i d\tau - \partial_k \xi^i d\tilde{x}^k. \tag{4.3.14}$$

Using the above relation between dx^i and $d\tilde{x}^i$, we can expand $\delta_{ij}dx^i dx^j$ to first order in the shift perturbation as:

$$\delta_{ij}dx^i dx^j = (\delta_{ij} - 2\partial_{(i}\xi_{j)})d\tilde{x}^i d\tilde{x}^j - 2\xi'_j d\tau dx^j \tag{4.3.15}$$

where we defined $\xi'_i \equiv \partial_\tau \xi_i$ as usual. Finally, in terms of the coordinates $(\tilde{\boldsymbol{x}}, \tau)$ the flat FRW space-time is described by the differential line element:

$$ds^2 = a^2(\tau)\Big[-d\tau^2 - 2\xi'_j d\tau dx^j + (\delta_{ij} - 2\partial_{(i}\xi_{j)})d\tilde{x}^i d\tilde{x}^j\Big]. \tag{4.3.16}$$

Comparing to the most generically perturbed line element (Eq. 4.3.4), we conclude that we have generated the metric perturbations $B_j \equiv \xi'_j$ and $\hat{E}_j \equiv \xi_j$. These perturbations are fictitious as they are generated by the spatial transformation and are also referred to as gauge modes.

▷ *Lapse and Temporal Shifts.*—As a second example, let us consider the lapse perturbation $\Phi(\boldsymbol{x}, \tau)$ subject to a change in the temporal coordinate. The shift is defined by:

$$\tilde{\tau} = \tau + \xi^0(\boldsymbol{x}, \tau), \tag{4.3.17}$$

where $\xi^0(\boldsymbol{x}, \tau)$ quantifies the shift and is assumed to be a first order perturbation as before. Under such a transformation $\tau \to \tilde{\tau}$, the lapse changes as follows:

$$\Phi(\boldsymbol{x}, \tau) = \Phi(\boldsymbol{x}, \tilde{\tau} - \xi^0) = \Phi(\boldsymbol{x}, \tilde{\tau}) - \Phi'(\boldsymbol{x}, \tilde{\tau})\, \xi^0(\boldsymbol{x}, \tau). \tag{4.3.18}$$

Notice that in making the temporal transformation $\tau \to \tilde{\tau}$, a second order perturbation $\delta\Phi(\boldsymbol{x}, \tau) \equiv \Phi'(\boldsymbol{x}, \tilde{\tau})\, \xi^0(\boldsymbol{x}, \tau)$ in the lapse was generated.

▷ *Energy Density and Temporal Shifts.*—Finally, we consider the effect of a change in the temporal coordinate:

$$\tilde{\tau} = \tau + \xi^0(\boldsymbol{x}, \tau), \tag{4.3.19}$$

on the background energy density $\bar{\rho}(\tau)$. The energy density then changes as:

$$\rho(\tau) \to \rho(\tilde{\tau}) = \bar{\rho}(\tau) + \bar{\rho}'\xi^0. \tag{4.3.20}$$

Observe that in shifting the temporal coordinate, the energy density receives a fictitious perturbation, given by:

$$\delta\rho \equiv \bar{\rho}'\xi^0. \tag{4.3.21}$$

In a similar manner, a real density perturbation could be removed by choosing a different time slicing of the spatial hypersurfaces.

The above examples show that fictitious perturbations are already generated when performing simple space-time transformations. Therefore, a concise approach is required to discriminate between physical and fictitious gauge transformations. There are two approaches to overcome the gauge problem and both will be discussed briefly below. The first is to construct gauge-invariant perturbations, i.e. perturbations that do not change under a gauge transformation. For scalar perturbations, those gauge-invariant combinations were derived by Bardeen [12] and are called the Bardeen potentials. The second approach is to work consistently in a specific gauge and choose the gauge transformation:

$$x^\mu \to \tilde{x}^\mu = x^\mu + \xi^\mu(\tau, \boldsymbol{x}), \tag{4.3.22}$$

such that one moves from the generic frame (as specified by the Eq. 4.3.4) to the chosen gauge.

4.3.2 Gauge Transformations of Metric Perturbations

To be able to use the second approach discussed above, we have to derive the explicit gauge transformations of perturbations in the metric. We will derive the transformation equations by exploiting the invariance of the differential line element ds^2 under gauge transformations:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = \tilde{g}_{\mu\nu}d\tilde{x}^\mu d\tilde{x}^\nu. \tag{4.3.23}$$

This approach is the same as the one taken in Ref. [66] to derive gauge-transformations of the metric perturbations. As before, the gauge transformation is defined in terms of the 4-shift ξ^μ, in terms of explicit temporal and spatial transformations we can write:

$$\tilde{\tau} = \tau + \xi^0, \qquad \tilde{x}^i = x^i + \partial^i \xi + \hat{\xi}^i, \tag{4.3.24}$$

where the spatial shift is decomposed into the gradient of a scalar ξ and a divergenceless vector $\hat{\xi}^i$. At first order, the 4-shift vector is the same in both frames, $\xi^\mu(\boldsymbol{x}, \tau) = \xi^\mu(\tilde{\boldsymbol{x}}, \tilde{\tau})$, so that the total differentials of ξ^0, $\partial^i \xi$ and $\hat{\xi}^i$ can be taken with respect to the variables in the tilde frame, yielding:[6]

$$d\xi^0 = \xi^{0\prime} d\tilde{\tau} + \partial_i \xi^0 d\tilde{x}^i, \tag{4.3.25}$$

$$d\xi = \xi' d\tilde{\tau} + \partial_j \xi d\tilde{x}^j, \tag{4.3.26}$$

$$d\hat{\xi}^i = \hat{\xi}^{i\prime} d\tilde{\tau} + \partial_j \hat{\xi}^i d\tilde{x}^j. \tag{4.3.27}$$

The temporal and spatial differentials in both frames are therefore related to each other at first order via:

$$d\tau = d\tilde{\tau} - \xi^{0\prime} d\tilde{\tau} + \partial_i \xi^0 d\tilde{x}^i, \tag{4.3.28}$$

$$dx^i = d\tilde{x}^i - (\partial^i \xi' + \hat{\xi}^{i\prime}) d\tilde{\tau} - \partial_j (\partial^i \xi + \hat{\xi}^i) d\tilde{x}^j. \tag{4.3.29}$$

The perturbed line element:

$$ds^2 = a^2(\tau)\Big[-(1+2\Phi)d\tau^2 + 2B_i dx^i d\tau + \Big((1-2\Psi)\delta_{ij} + 2\partial_i\partial_j E + 2\partial_{(i}\hat{E}_{j)} + 2\hat{E}_{ij}\Big) dx^i dx^j \Big], \tag{4.3.30}$$

can now be written in the tilde frame by using the first order relation between the differentials in both frames. To transform the overall scale factor, we use the fact that a is has only temporal dependence, so that under a shift $\tau \to \tilde{\tau} = \tau + \xi^0$, the scale factor transforms as:

$$a(\tau) = a(\tilde{\tau}) - \xi^0 a'(\tilde{\tau}) = a(\tilde{\tau})\Big[1 - \mathcal{H}\xi^0\Big]. \tag{4.3.31}$$

Substituting the obtained first order expressions relating $d\tau$, dx^i and $a(\tau)$ to the tilde frame into the perturbed line element ds^2 and keeping only terms at linear order in both perturbations and the shift variable yields [66]:

$$\begin{aligned} ds^2 = a^2(\tilde{\tau})\Big[&-\big(1 + 2(\Phi - \mathcal{H}\xi^0 - \xi^{0\prime})\big) d\tilde{\tau}^2 \\ &+ 2\partial_i (B + \xi^0 - \xi') d\tilde{\tau} d\tilde{x}^i + 2(\hat{B}_i - \hat{\xi}_i') d\tilde{\tau} d\tilde{x}^i \\ &+ \Big[\big(1 - 2(\Psi + \mathcal{H}\xi^0)\big)\delta_{ij} + 2\partial_i\partial_j (E - \xi) + 2(\partial_{(i}\hat{E}_{j)} - \partial_{(i}\hat{\xi}_{j)}) + 2\hat{E}_{ij}\Big] d\tilde{x}^i d\tilde{x}^j \Big]. \end{aligned} \tag{4.3.32}$$

Now we can extract the gauge transformations for the perturbations by comparing the above expression to the line element in the tilde frame, which is equivalent to Eq. 4.3.30 with

[6] The choice to differentiate with respect to the tilde frame coordinates is an arbitrary choice; one could equally well perform the differentiations with respect to the untilded frame.

tildes on the perturbations and coordinates. Notice that the tensor perturbation does not transform: $\tilde{\hat{E}}_{ij} = \hat{E}_{ij}$. For the scalar perturbations, we find:

$$\tilde{\Phi} = \Phi - \mathcal{H}\xi^0 - \xi^{0\prime}, \tag{4.3.33}$$

$$\tilde{B} = B + \xi^0 - \xi', \tag{4.3.34}$$

$$\tilde{\Psi} = \Psi + \mathcal{H}\xi^0, \tag{4.3.35}$$

$$\tilde{E} = E - \xi. \tag{4.3.36}$$

For the divergence-free vectors, we get:

$$\tilde{\hat{B}}_i = \hat{B}_i - \hat{\xi}_i, \qquad \partial_{(i}\tilde{\hat{E}}_{j)} = \partial_{(i}\hat{E}_{j)} - \partial_{(i}\hat{\xi}_{j)}. \tag{4.3.37}$$

4.3.3 Gauge Transformations of Matter Scalars and Vectors

Having discussed the gauge transformations of metric perturbations, we will now discuss the gauge transformations of the physical scalars and vectors, following the discussion in [66]. Any scalar σ which is homogeneous at zeroth-order (i.e. in the FRW background the scalar possesses no spatial dependence) can be written as:

$$\sigma(\tau, \boldsymbol{x}) = \bar{\sigma}(\tau) + \delta\sigma(\tau, \boldsymbol{x}). \tag{4.3.38}$$

Examples of physical scalars are the energy density ρ, spatial curvature, scalar shear. Physical scalars depend on the chosen time slicing of the hypersurfaces of constant time and hence a gauge transformation of a physical scalar depends on the temporal shift ξ^0. However, physical scalars are independent of the choice of spatial coordinates on the hypersurfaces of constant time and hence gauge transformation do not posses dependence on ξ^i. Explicitly, a scalar perturbation $\delta\sigma$ transforms as follows:

$$\delta\tilde{\sigma} = \delta\sigma - \xi^0\bar{\sigma}'. \tag{4.3.39}$$

Similarly, physical 3-vectors such as matter velocity $\boldsymbol{v}$, only depend on the spatial threading of the hypersurface and are independent of the chosen time slicing. Hence, a gauge transformation for $\boldsymbol{v}$ only depends on the spatial shift ξ^i and not on the temporal shift ξ^0, and can be written as:

$$\tilde{v}^i = v^i + \xi^{i\prime}. \tag{4.3.40}$$

In case the vector can be written as the gradient of a scalar, i.e. $v^i = \partial^i v$, the transformation equation becomes:

$$\tilde{v} = v + \xi'. \tag{4.3.41}$$

Finally, the transformation equation for a divergence-free vector $\boldsymbol{v}$ is:

$$\tilde{v}^i = v^i + \hat{\xi}^i. \tag{4.3.42}$$

4.3.4 Bardeen Potentials

Because of the gauge-dependence of the metric perturbations, Bardeen [12] proposed to consider solely perturbations that are manifestly gauge-invariant. For scalar perturbations, the scalar gauge transformations can be used to eliminate two scalar metric perturbations. Hence,

according to Bardeen, the remaining scalar perturbations should be used to construct gauge-invariant combinations. Using the gauge transformations for Φ, B, Ψ and E, Bardeen constructed the following two gauge-invariant potentials:

$$\Phi_{\mathrm{B}} \equiv \Phi + \mathcal{H}(B - E') + (B - E')', \tag{4.3.43}$$

$$\Psi_{\mathrm{B}} \equiv \Psi - \mathcal{H}(B - E'). \tag{4.3.44}$$

One could easily verify, by using the transformation equations for the scalar perturbations, that Φ_{B} and Ψ_{B} are indeed gauge-invariant. As we will show in the next subsection, the Bardeen potentials coincide with the scalar metric perturbations in the Newtonian gauge. Hence, one may conclude that the Newtonian gauge is preferred over other gauge choices. However, it turns out that gauge-invariant perturbations can always be constructed when an unambiguous time-slicing is chosen [66].

4.4 The Newtonian Gauge

As mentioned in the previous section, the second solution to overcome the gauge problem is to work consistently in an unambiguously defined gauge. Throughout this work, we will work mostly in the Newtonian gauge, which is constructed by choosing the temporal and spatial shifts (ξ_0 and ξ) such that the scalar lapse B and scalar shear E vanish. Mathematically, this gauge choice[7] is simplest, as the perturbed metric remains diagonal since $g_{0i} = 0$. Therefore, we will compute most of the calculations in this work in the Newtonian gauge. To move from an arbitrary gauge to the Newtonian gauge, we choose ξ^0 and ξ such that $\tilde{E}$ and $\tilde{B}$ both vanish. On account of Eq. 4.3.36, we can uniquely chose:

$$\xi = E, \tag{4.4.1}$$

so that $\tilde{E}$ becomes zero. Furthermore, using Eq. 4.3.34, we can set:

$$\xi^0 = \xi' - B = E' - B, \tag{4.4.2}$$

so that $\tilde{B}$ vanishes. The remaining scalar perturbations are then Φ and Ψ and they coincide with the Bardeen potentials since $B = E = 0$ in the Newtonian gauge. The perturbed line element in this gauge becomes:

$$ds^2 = a^2(\tau)\Big[- (1 + 2\Phi)\, d\tau^2 + (1 - 2\Psi)\delta_{ij} dx^i dx^j \Big]. \tag{4.4.3}$$

Notice that this line element is very similar to the line element obtained in the weak field limit of general relativity (see e.g. [31] for a discussion on the weak field limit of GR). Hence, the perturbations Φ and Ψ may be regarded as gravitational potentials. The non-zero metric components corresponding to the above line element are given by:

$$g_{00} = -a^2(1 + 2\Phi), \qquad g_{ij} = a^2(1 - 2\Psi)\delta_{ij}, \tag{4.4.4}$$

and the metric perturbations are therefore:

$$\delta g_{00} = -2a^2\Phi, \qquad \delta g_{ij} = -2a^2\Psi\delta_{ij}. \tag{4.4.5}$$

[7] Other popular gauges are discussed in Appendix D.4.

The inverse metric components g^{00} and g^{ij} then become:[8]

$$g^{00} \equiv (g_{00})^{-1} = -(1-2\Phi)/a^2, \qquad g^{ij} \equiv (g_{ij})^{-1} = (1+2\Psi)\delta^{ij}/a^2. \tag{4.4.6}$$

The perturbations are now given explicitly by:

$$\delta g^{00} = 2\Phi/a^2, \qquad \delta g^{ij} = 2\Psi\delta^{ij}/a^2. \tag{4.4.7}$$

In the Newtonian gauge, the physics appears rather intuitively, as the constant time hypersurfaces (the slicing) are orthogonal to the worldlines of observers at rest on the coordinate system comoving with the expansion (the threading). Furthermore, the great virtue of the Newtonian gauge is that it allows to obtain analytic results for the evolution of inflationary perturbations. In the absence of anisotropic stress, i.e. when the energy-momentum tensor $T^{\mu\nu}$ contains no stress contribution ($\Sigma^{ij} = 0$), the two remaining scalar perturbations are equal: $\Psi = \Phi$.

4.5 The Comoving Curvature Perturbation

The comoving curvature perturbation, which is defined as follows in the comoving orthogonal gauge (recall Eq. D.4.7) for scalar perturbations:

$$\mathcal{R} = \Psi - \mathcal{H}(v+B), \tag{4.5.1}$$

is of significant importance in relating inflationary predictions to late-time observables, such as the CMB, since it is constant on super-horizon scales. In this section, we will provide a geometrical definition of the (comoving) curvature perturbation.[9]

The perturbed spatial 3-metric γ_{ij} (for a spatially flat universe) is given by:

$$\gamma_{ij} = (1-2\Psi)\delta_{ij} + 2E_{ij}, \tag{4.5.2}$$

where the tensor perturbation E_{ij} is symmetric and traceless ($E^i_i = 0$). The intrinsic curvature on spatial hypersurfaces is given by ${}^{(3)}R$ and can be written as:

$${}^{(3)}R = g^{ij}R_{ij} = \frac{\gamma^{ij}}{a^2}R_{ij}. \tag{4.5.3}$$

We will derive that, when considering scalar perturbations only, ${}^{(3)}R$ is related to the curvature perturbation Ψ in the following way:

$${}^{(3)}R = \frac{4}{a^2}\partial^2\Psi. \tag{4.5.4}$$

It should be noted that in the above definition of Ψ, the term proportional to $\partial^2 E$ resulting from $E^{\mathrm{S}}_{ij} = \partial_{\langle i}\partial_{j\rangle}E$ is already absorbed (conform Eq. 4.3.10).

The above result may be considered as the definition of the curvature perturbation Ψ. However, notice that the curvature perturbation is gauge-dependent (the transformation

[8] To obtain the inverse perturbed metric in the Newtonian gauge, we use the fact that the metric perturbations Φ and Ψ are assumed to be small relative to unity, and hence the first order approximation $(1+\alpha)^{-1} = 1-\alpha$ for $\alpha \ll 1$ can be used.

[9] Notice the difference between the curvature perturbation, which is denoted as Ψ and defined in an arbitrary gauge, and the comoving curvature perturbation $\mathcal{R}$, which is defined solely in the comoving gauge.

equation is given Eq. 4.3.35). In the comoving gauge, the curvature perturbation Ψ is equivalent to the comoving curvature perturbation $\mathcal{R}$. To obtain the gauge-invariant expression for $\mathcal{R}$ from Ψ as defined in an arbitrary gauge, we use the fact $v = B = 0$ in the comoving gauge and hence linear combinations of those two can always be added to Ψ to obtain a gauge-invariant combination. From the transformation equations (Eqs. 4.3.33–4.3.36), we find that the manifestly gauge-invariant combination is given by:

$$\mathcal{R} \equiv \Psi - \mathcal{H}(v + B), \tag{4.5.5}$$

which is in correspondence with the definition of $\mathcal{R}$ given earlier. Note that in the comoving gauge, the intrinsic curvature of spatial hypersurfaces is related directly to $\mathcal{R}$:

$$^{(3)}R = \frac{4}{a^2}\partial^2\mathcal{R}. \qquad \text{(Comoving Gauge)} \tag{4.5.6}$$

The expression for the intrinsic curvature $^{(3)}R$ (Eq. 4.5.4) will be derived explicitly below.

Derivation: Comoving Curvature Perturbation

Here, we will derive Eq. 4.5.4 for the intrinsic curvature of the spatial hypersurfaces from first principles. The perturbed spatial hypersurfaces are described by the perturbed spatial metric:

$$\gamma_{ij} = \Big[(1 - 2\Psi)\delta_{ij} + 2E_{ij}\Big], \tag{4.5.7}$$

where γ_{ij} is related to the spatial part of the perturbed FRW metric via $g_{ij} = a^2\gamma_{ij}$ (Eq. 1.2.4) and the tensor perturbation satisfies $E^i_i = 0$. It will turn out to be convenient to have the zeroth order expression for γ^{ij} at hand as well:

$$\gamma^{ij} = \delta^{ij} + \mathcal{O}(\Psi, E^{ij}), \tag{4.5.8}$$

where $\mathcal{O}(\Psi, E^{ij})$ indicates the above expression only holds to zeroth order in perturbations.

By definition, the intrinsic curvature of the spatial hypersurfaces is given by:

$$^{(3)}R \equiv g^{ij}R_{ij} = \frac{\gamma^{ij}}{a^2}R_{ij}. \tag{4.5.9}$$

In terms of the Christoffel symbol $^{(3)}\Gamma^i_{jk}$, which reads:

$$^{(3)}\Gamma^i_{jk} = \frac{1}{2}\gamma^{il}(\partial_j\gamma_{kl} + \partial_k\gamma_{jl} - \partial_l\gamma_{jk}), \tag{4.5.10}$$

the intrinsic curvature becomes:

$$a^2\ ^{(3)}R = \gamma^{ik}\partial^{(3)}_l\Gamma^l_{ik} - \gamma^{ik}\partial^{(3)}_k\Gamma^l_{il} + \mathcal{O}(\Gamma^2), \tag{4.5.11}$$

where $\mathcal{O}(\Gamma^2)$ indicates terms such as $\gamma^{ik}{}^{(3)}\Gamma^l_{ij}{}^{(3)}\Gamma^m_{lm}$ which should in principle be included. However, they are second order in perturbations and hence they can be neglected in first order analysis. In the above expression, the perturbed expression for $^{(3)}\Gamma^l_{ik}$ to first order is given by:

$$^{(3)}\Gamma^l_{ik} = -\partial_i\Psi\delta^l_k + \partial_iE^l_k - \partial_k\Psi\partial^l_i + \partial_kE^l_i + \partial^l\Psi\delta_{ik} - \partial^lE_{ik}, \tag{4.5.12}$$

as this expression is already first order we can use $\gamma^{ik} = \delta^{ik}$. Note that furthermore that the background term in the expansion of γ_{ik} is not present in ${}^{(3)}\Gamma^l_{ik}$ as the spatial derivatives acting on the background term δ_{ik} yield zero.

Substituting the above result in the first order expression for ${}^{(3)}R$ and using $E^i_i = 0$ gives, after some manipulations:

$$a^2\ {}^{(3)}R = 4\partial_l\partial^l\Psi + 2\partial_l\partial^k E^l_k = 4\partial^2\Psi + 2\partial_i\partial_j E^{ij}. \tag{4.5.13}$$

Now using that for scalar perturbation we can write the tensorial perturbation E^{ij} can be written as (Eq. 4.3.7):

$$E^{\mathrm{S}}_{ij} = \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\partial^2\right)E \equiv \partial_{\langle i}\partial_{j\rangle}E. \tag{4.5.14}$$

Substituting this result in to the expression for the intrinsic curvature on spatial hypersurfaces yields:

$${}^{(3)}R = \frac{4}{a^2}\partial^2\left(\Psi + \frac{1}{3}\partial^2 E\right). \tag{4.5.15}$$

Absorbing the second term between parentheses into the first as before, $\Psi + \partial^2 E/3 \to \Psi$, we obtain the desired result.

4.6 Adiabatic and Isocurvature Perturbations

So far, we have considered generic perturbations and did not classify perturbations into different classes. However, it turns out that in a thermodynamic context, perturbations can be divided into two distinct types: they can either be of the adiabatic or isocurvature type.

4.6.1 Adiabatic Perturbations

By definition, a generic adiabatic or curvature perturbation δX can be expressed as a time shift $\delta t(t, \boldsymbol{x})$ of the background field $\bar{X}$ [14, 60, 74]:

$$\delta X \equiv \dot{\bar{X}}\ \delta t(t, \boldsymbol{x}). \qquad \text{(Adiabicity)} \tag{4.6.1}$$

That is, in phase space, the trajectory of an adiabatic perturbation δX coincides with that of the background $\bar{X}$. Stated differently, the characteristic property of adiabatic perturbations is that the local state of the fluid in the perturbed universe, for instance described by the energy density $\rho(t, \boldsymbol{x})$ and pressure $P(t, \boldsymbol{x})$, is the same as in the unperturbed universe at a slightly different moment in time $t + \delta t(t, \boldsymbol{x})$.

The above defining condition of adiabicity can be derived from the left diagram in Fig. 4.1, by invoking the fact that the phase space trajectory of the perturbation coincides with the background trajectory. Consider the (adiabatic) perturbation denoted in field phase space by the vector:

$$\boldsymbol{P}_{\delta X} = (\delta X, \delta\dot{X}), \tag{4.6.2}$$

originating from the point $P_{\delta X}$, which lies on the background trajectory $\dot{\bar{X}}(\bar{X})$ of the field X in phase space. Notice that the perturbation vector has the same slope as the background

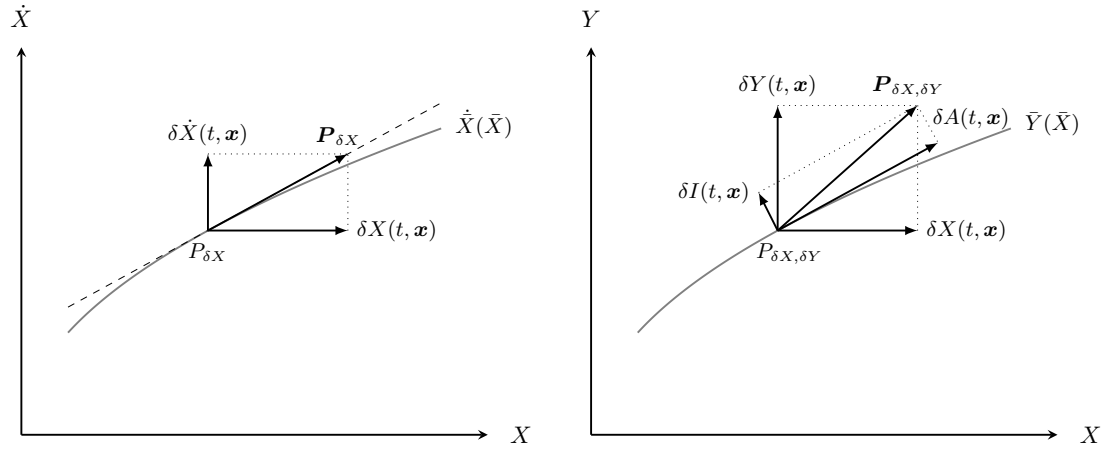


Figure 4.1: Left: Decomposition of an adiabatic perturbation $\boldsymbol{P}_{\delta X}$ in phase space $(X, \dot{X})$. Right: Decomposition of a generic perturbation $\boldsymbol{P}_{\delta X,\delta Y}$ in field space (X, Y) into adiabatic δA and isocurvature δI contributions.

trajectory at the point $P_{\delta X}$. Hence, we can write:

$$\frac{\delta\dot{X}(t,\boldsymbol{x})}{\delta X(t,\boldsymbol{x})} = \frac{d\dot{\bar{X}}(t)}{d\bar{X}(t)}, \tag{4.6.3}$$

from which we can deduce that:

$$\delta X(t,\boldsymbol{x}) = d\bar{X}(t) = \dot{\bar{X}}(t)\ \delta t(t,\boldsymbol{x}), \tag{4.6.4}$$

which is indeed the defining property of adiabatic perturbations.

For a universe filled with multiple fluids (labeled by f), adiabatic perturbations correspond to perturbations induced by a common local shift in time. If the fluids f are described by density ρ and pressure P, we find that adiabatic perturbations are defined as:

$$\delta\rho_f(t,\boldsymbol{x}) \equiv \bar{\rho}_f(t+\delta t(t,\boldsymbol{x})) - \bar{\rho}_f(t) = \dot{\bar{\rho}}_f\ \delta t(t,\boldsymbol{x}), \qquad \delta P_f(t,\boldsymbol{x}) = \dot{\bar{P}}_f\ \delta t(t,\boldsymbol{x}), \tag{4.6.5}$$

where the shift $\delta t(t,\boldsymbol{x})$ is the same for all fluids f. This time shift can thus be written as:

$$\delta t(t,\boldsymbol{x}) = \frac{\delta\rho_f(t,\boldsymbol{x})}{\dot{\bar{\rho}}_f(t)} = \frac{\delta P_f(t,\boldsymbol{x})}{\dot{\bar{P}}_f(t)}. \tag{4.6.6}$$

Assuming furthermore that there is no energy transfer between the different fluids, they are seperately conserved and we can use the continuity equation $\dot{\bar{\rho}}_f = -3H(1+w_f)\bar{\rho}_f$ to find:

$$\frac{\delta_f}{1+w_f} = \frac{\delta_{f'}}{1+w_{f'}}, \tag{4.6.7}$$

for two fluids f and f' and we defined $\delta_f \equiv \delta\rho_f/\bar{\rho}_f$. The above equation is often taken as the definiton of adiabicity [14, 90, 92]. Notice that single field inflation produced adiabatic perturbations, as they are all generated by the time shift $\delta t(t,\boldsymbol{x})$ induced in the inflaton field via quantum fluctuations (see introduction to chapter 3):

$$\delta t(t,\boldsymbol{x}) = \frac{\delta\phi(t,\boldsymbol{x})}{\dot{\bar{\phi}}(t)}. \tag{4.6.8}$$

Notice that in the multi-field scenario the perturbations will not necessarily be adiabatic. Consider for instance two dynamically important fields, ϕ and χ, during inflation. The time shifts in those two fields, as induced by quantum fluctuations in the fields, do not have to be the same and hence:

$$\frac{\delta\phi}{\dot{\bar{\phi}}} \neq \frac{\delta\chi}{\dot{\bar{\chi}}}, \tag{4.6.9}$$

which violates the adiabatic condition. This naturally leads to the complement of adiabatic perturbations: isocurvature perturbations.

4.6.2 Isocurvature Perturbations

Isocurvature perturbations cause the perturbed solution to deviate from the background solution, that is, for X and Y we have:

$$\frac{\delta X}{\dot{\bar{X}}} \neq \frac{\delta Y}{\dot{\bar{Y}}}. \tag{4.6.10}$$

Taking Eq. 4.6.6 as the defining condition for adiabatic density perturbations, we can define isocurvature fluctuations in the energy density as:

$$\mathcal{S} \equiv H\left(\frac{\delta P}{\dot{P}} - \frac{\delta\rho}{\dot{\rho}}\right). \tag{4.6.11}$$

Using this expression for $\mathcal{S}$ and the sound speed c_s^2, defined as:[10]

$$c_s^2 = \frac{\bar{P}'}{\bar{\rho}'}, \tag{4.6.12}$$

we find that we can relate δP to $\delta\rho$ as follows:

$$\delta P = c_s^2\left(\delta\rho - 3(\bar{\rho} + \bar{P})\mathcal{S}\right). \tag{4.6.13}$$

The above relation is valid in any gauge and for adiabatic perturbations, we have $\mathcal{S} = 0$. Therefore, we find that perturbations in energy density and pressure are simply related via the sound speed. The above definition can be generalized to the isocurvature perturbation between any two matter quantities x and y in the following way:

$$\mathcal{S}_{xy} = H\left(\frac{\delta x}{\dot{\bar{x}}} - \frac{\delta y}{\dot{\bar{y}}}\right). \tag{4.6.14}$$

As single-field inflation predicts solely adiabatic perturbations, we have $\mathcal{S} = \mathcal{S}_{ff'} = 0$ for any variables f and f'.

In the right diagram of Fig. 4.1, we show how a generic perturbation $\boldsymbol{P}_{\delta X,\delta Y}$ in the field space of X and Y can be decomposed into an adiabatic part δA, parallel to the background trajectory $\bar{Y}(\bar{X})$, and an isocurvature contribution δI, perpendicular to the background trajectory. The usual decomposition along the X and Y axes is also shown.

[10] It should be noted that both the equation of state $w \equiv \bar{P}/\bar{\rho}$ and c_s^2 are always defined at zeroth order in perturbations.

4.7 Einstein Tensor

In this section, the Einstein tensor will be perturbed to first order around a flat FRW background. As described in the previous section, the strategy will be to first perturb the Christoffel symbols, then the Ricci tensor and finally the Einstein tensor.

4.7.1 Perturbed Christoffel Symbols

The Christoffel symbol (or affine connection) is defined in term of the metric and its derivatives as (Eq. A.2.1):

$$\Gamma^{\mu}_{\rho\sigma} = \frac{1}{2} g^{\mu\lambda} \left(\partial_\rho g_{\sigma\lambda} + \partial_\sigma g_{\rho\lambda} - \partial_\lambda g_{\rho\sigma} \right). \tag{4.7.1}$$

Perturbing the Christoffel symbol to first order gives:

$$\delta\Gamma^{\mu}_{\rho\sigma} = \frac{1}{2} \delta g^{\mu\lambda} \left(\partial_\rho g_{\sigma\lambda} + \partial_\sigma g_{\rho\lambda} - \partial_\lambda g_{\rho\sigma} \right) + \frac{1}{2} g^{\mu\lambda} \left(\partial_\rho \delta g_{\sigma\lambda} + \partial_\sigma \delta g_{\rho\lambda} - \partial_\lambda \delta g_{\rho\sigma} \right). \tag{4.7.2}$$

Here, it is understood that the unbarred metric corresponds to the background metric and the first order perturbations are denoted using the δ-notation.

The perturbed Christoffel symbols can be computed directly using the metric perturbations and the non-vanishing components are:

$$\begin{aligned}
\delta\Gamma^{0}_{00} &= \Phi', \\
\delta\Gamma^{0}_{0i} &= \partial_i \Phi, \\
\delta\Gamma^{i}_{00} &= \partial^i \Phi, \\
\delta\Gamma^{i}_{0j} &= -\Psi' \delta^i_j, \\
\delta\Gamma^{0}_{ij} &= -2\mathcal{H}\Phi\delta_{ij} - 2\mathcal{H}\Psi\delta_{ij} - \Psi'\delta_{ij}, \\
\delta\Gamma^{i}_{jk} &= -\partial_j \Psi \delta^i_k - \partial_k \Psi \delta^i_j + \partial^i \Psi \delta_{jk}.
\end{aligned} \tag{4.7.3}$$

Note that we work in conformal time and hence $\mathcal{H} \equiv a'/a$ and the prime denotes a conformal time derivative: $\mathcal{Q}' \equiv \partial_0 \mathcal{Q} = \partial_\tau \mathcal{Q}$. Below, we will show the explicit derivations of the purely temporal and purely spatial perturbed Christoffel symbols.

Derivation: Perturbed Christoffel Symbols

Purely Temporal Component.—To derive $\delta\Gamma^0_{00}$ we set all indices equal to zero in $\delta\Gamma^{\mu}_{\rho\sigma}$ to obtain:

$$\delta\Gamma^0_{00} = \frac{1}{2}\delta g^{0\rho}(\partial_0 g_{\rho 0} + \partial_0 g_{0\rho} - \partial_\rho g_{00}) + \frac{1}{2} g^{0\rho}(\partial_0 \delta g_{\rho 0} + \partial_0 \delta g_{0\rho} - \partial_\rho \delta g_{00}). \tag{4.7.4}$$

Now we compute the first $(*)$ and second $(**)$ term in the above expression. Note that the first term vanishes for $\rho \neq 0$ since in the Newtonian gauge $\delta g^{0i} = 0$. Hence, we set ρ to zero and use $g_{00} = -a^2$ and $\delta g^{00} = 2\Phi/a^2$ to evaluate the first term as follows:

$$(*) = \frac{1}{2}\delta g^{00}\partial_0 g_{00} = -2\mathcal{H}\Phi. \tag{4.7.5}$$

Similar reasoning applies to the second term, which can be computed by using $g^{00} = -1/a^2$ and $\delta g_{00} = -2a^2\Phi$ as follows:

$$(**) = \frac{1}{2}g^{00}\partial_0\delta g_{00} = \frac{1}{a^2}\partial_0(a^2\Phi) = \Phi' + 2\mathcal{H}\Phi. \tag{4.7.6}$$

Combining the results for $(*)$ and $(**)$ gives the desired expression:

$$\delta\Gamma^0_{00} = \Phi'. \tag{4.7.7}$$

Purely Spatial Component.—To derive the purely spatial perturbed Christoffel symbol we consider $\delta\Gamma^i_{jk}$ and evaluate:

$$\delta\Gamma^i_{jk} = \frac{1}{2}\delta g^{i\rho}(\partial_j g_{\rho k} + \partial_k g_{j\rho} - \partial_\rho g_{jk}) + \frac{1}{2}g^{i\rho}(\partial_j\delta g_{\rho k} + \partial_k\delta g_{j\rho} - \partial_\rho\delta g_{jk}). \tag{4.7.8}$$

The first term vanishes completely in the Newtonian gauge. In case ρ is spatial the metric derivatives between parentheses vanish since the unperturbed metric is no function of spatial coordinates: $\partial_m g_{jk} = 0$. For $\rho = 0$, the expression vanishes as well since $\delta g^{i0} = 0$. The second term can be computed straightforwardly using the fact it vanishes unless $\rho = m$ (i.e. it is spatial). In that case, using $\delta g_{ij} = -2a^2\Psi$, the expression becomes:

$$(**) = -\partial_j\Psi\delta^i_k - \partial_k\Psi\delta^i_j + \partial^i\Psi\delta_{jk} = \delta\Gamma^i_{jk}. \tag{4.7.9}$$

4.7.2 Perturbed Ricci Tensor

Now that the perturbed Christoffel symbols are derived, we can use them to obtain the first-order perturbation of the Ricci tensor $\delta R_{\mu\nu}$. In general, the first order perturbed Ricci tensor can be expressed as:

$$\delta R_{\mu\nu} = \partial_\alpha\delta\Gamma^\alpha_{\mu\nu} - \partial_\mu\delta\Gamma^\alpha_{\nu\alpha} + \delta\Gamma^\alpha_{\sigma\alpha}\Gamma^\sigma_{\mu\nu} + \Gamma^\alpha_{\sigma\alpha}\delta\Gamma^\sigma_{\mu\nu} - \delta\Gamma^\alpha_{\sigma\nu}\Gamma^\sigma_{\mu\alpha} - \Gamma^\alpha_{\sigma\nu}\delta\Gamma^\sigma_{\mu\alpha}. \tag{4.7.10}$$

Defining the differential operator $\partial^2 \equiv \partial_i\partial^i$, the explicit components are given by:

$$\delta R_{00} = \partial^2\Phi + 3\Psi'' + 3\mathcal{H}\Phi' + 3\mathcal{H}\Psi', \tag{4.7.11}$$

$$\delta R_{0i} = 2\partial_i\Psi' + 2\mathcal{H}\partial_i\Phi, \tag{4.7.12}$$

$$\begin{aligned}\delta R_{ij} =& \Big[-2(a''/a)(\Phi+\Psi) - \mathcal{H}\Phi' - 5\mathcal{H}\Psi' + \partial^2\Psi - \Psi'' - 2\mathcal{H}^2(\Phi+\Psi)\Big]\delta_{ij} \\ &+ \partial_i\partial_j\Psi - \partial_i\partial_j\Phi.\end{aligned} \tag{4.7.13}$$

In the derivation of δR_{ij} we used the fact that the time derivative of the Hubble constant can be written as $\mathcal{H}' = a''/a - \mathcal{H}^2$. Below, the purely temporal component will be derived explicitly. Deriving the other two components is slightly more computationally demanding, but the strategy is the same as the one used below to derive δR_{00}.

Derivation: Temporal Component Perturbed Ricci Tensor

To obtain the purely temporal perturbed Ricci tensor component δR_{00}, we set μ and ν equal to zero in the expression for $\delta R_{\mu\nu}$ to get:

$$\delta R_{00} = \partial_\alpha \delta\Gamma^\alpha_{00} - \partial_0 \delta\Gamma^\alpha_{0\alpha} + \delta\Gamma^\alpha_{\sigma\alpha}\Gamma^\sigma_{00} + \Gamma^\alpha_{\sigma\alpha}\delta\Gamma^\sigma_{00} - \delta\Gamma^\alpha_{\sigma 0}\Gamma^\sigma_{0\alpha} - \Gamma^\alpha_{\sigma 0}\delta\Gamma^\sigma_{0\alpha}. \tag{4.7.14}$$

We first evaluate the frist two terms in the expression for δR_{00}, that is:

$$(1) \equiv \partial_\alpha \delta\Gamma^\alpha_{00} - \partial_0 \delta\Gamma^\alpha_{0\alpha}. \tag{4.7.15}$$

The first term $(*)$ in (1) gives:

$$(*) = \partial_0 \delta\Gamma^0_{00} + \partial_i \delta\Gamma^i_{00} = \Phi'' + \partial_i \partial^i \Phi = \Phi'' + \partial^2 \Phi, \tag{4.7.16}$$

and the second term $(**)$ yields:

$$(**) = -\partial_0 \delta\Gamma^\alpha_{0\alpha} = -\partial_0 \delta\Gamma^0_{00} - \partial_0 \delta\Gamma^i_{0i} = -\Phi'' + 3\Psi''. \tag{4.7.17}$$

Now, we consider the terms:

$$(2) \equiv \delta\Gamma^\alpha_{\sigma\alpha}\Gamma^\sigma_{00} + \Gamma^\alpha_{\sigma\alpha}\delta\Gamma^\sigma_{00}. \tag{4.7.18}$$

The first term in (2) yields:

$$(*) = \delta\Gamma^\alpha_{\sigma\alpha}\Gamma^\sigma_{00} = \delta\Gamma^0_{00}\Gamma^0_{00} + \delta\Gamma^0_{i0} + \delta\Gamma^i_{0i}\Gamma^0_{00} + \delta\Gamma^i_{ji}\Gamma^j_{00} = \mathcal{H}\Phi' - 3\mathcal{H}\Psi', \tag{4.7.19}$$

where we used the fact that zeroth order Christoffel symbols with two temporal indices vanish $\Gamma^0_{0i} = 0$ and the fact that $\Gamma^k_{ij} = 0$ since $\partial_k g_{ij} = 0$. The second term gives:

$$(**) = \Gamma^\alpha_{\sigma\alpha}\delta\Gamma^\sigma_{00} = 4\mathcal{H}\Phi'. \tag{4.7.20}$$

In a similar way, the last two terms in the expression for δR_{00}, labelled as (3), can be computed to give:

$$(3) \equiv -\delta\Gamma^\alpha_{\sigma 0}\Gamma^\sigma_{0\alpha} - \Gamma^\alpha_{\sigma 0}\delta\Gamma^\sigma_{0\alpha} = -2\mathcal{H}\Phi' + 6\mathcal{H}\Psi'. \tag{4.7.21}$$

Adding the obtained expressions for (1), (2) and (3) finally yields the desired expression for δR_{00} as:

$$\delta R_{00} = (1) + (2) + (3) = \partial^2\Phi + 3\Psi'' + 3\mathcal{H}\Phi' + 3\mathcal{H}\Psi'. \tag{4.7.22}$$

4.7.3 Perturbed Ricci Scalar

The Ricci scalar is defined as $R \equiv g^{\mu\nu} R_{\mu\nu}$ and its background value is:

$$R = \frac{6}{a^2}\frac{a''}{a}. \tag{4.7.23}$$

Perturbing the Ricci tensor to first order gives:

$$\delta R = \delta g^{\mu\alpha} R_{\alpha\mu} + g^{\mu\alpha}\delta R_{\alpha\mu}. \tag{4.7.24}$$

Using the background and perturbed metric components and the first order Ricci tensor, the expression for δR can be derived to be:

$$\delta R = \frac{1}{a^2}\bigg[-12\frac{a''}{a}\Phi + 4\partial^2\Psi - 2\partial^2\Phi - 6\Psi'' - 6\mathcal{H}\Phi' - 18\mathcal{H}\Psi'\bigg]. \tag{4.7.25}$$

In the derivation box below, the stated expression for δR will be derived.

Derivation: Variation of the Ricci Scalar

We label the two terms in the expression of δR in terms of the perturbed metric and Ricci tensor as:

$$\delta R = \delta g^{\mu\nu} R_{\mu\nu} + g^{\mu\nu}\delta R_{\mu\nu} \equiv (1) + (2). \tag{4.7.26}$$

Using the background expressions for $R_{\mu\nu}$, the first term gives:

$$(1) = \delta g^{00} R_{00} + \delta g^{ij} R_{ij} = \frac{6}{a^2}\bigg[(\mathcal{H}^2 - a''/a)\Phi + (\mathcal{H}^2 + a''/a)\Psi\bigg]. \tag{4.7.27}$$

where the off-diagonal summation terms vanish since $R_{0i} = 0$ and we used the implicit summation for the Kronecker delta: $\delta^i_i = 3$. The second term can be expanded into non-zero terms as follows:

$$(2) = g^{00}\delta R_{00} + g^{ij}\delta R_{ij} = -\frac{1}{a^2}\delta R_{00} + \frac{1}{a^2}\delta^{ij}\delta R_{ij}, \tag{4.7.28}$$

again diagonal summation terms vanish since $g^{0i} = 0$. Substituting the expressions for δR_{00} and δR_{ij}, we finally get:

$$(2) = \frac{1}{a^2}\bigg[-6\frac{a''}{a}(\Phi+\Psi) - 2\partial^2\Phi + 4\partial^2\Psi - 6\mathcal{H}\Phi' - 18\mathcal{H}\Psi' - 6\Psi'' - 6\mathcal{H}^2(\Phi+\Psi)\bigg]. \tag{4.7.29}$$

Adding terms (1) and (2) gives the desired expression for the variation of the Ricci scalar.

4.7.4 Perturbed Einstein Tensor

Now that we have explicit expressions for the variations of the metric, Ricci tensor and scalar, we can construct the perturbed Einstein tensor. The definition of the Einstein tensor reads:

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R. \tag{4.7.30}$$

The general expression for the first order perturbation $\delta G_{\mu\nu}$ of the Einstein tensor is:

$$\delta G_{\mu\nu} = \delta R_{\mu\nu} - \frac{1}{2}\delta g_{\mu\nu}R - \frac{1}{2}g_{\mu\nu}\delta R. \tag{4.7.31}$$

The explicit components are given by:

$$\delta G_{00} = -6\mathcal{H}\Psi' + 2\partial^2\Psi, \tag{4.7.32}$$

$$\delta G_{0i} = 2\partial_i\Psi' + 2\mathcal{H}\partial_i\Phi, \tag{4.7.33}$$

$$\begin{aligned}\delta G_{ij} = &\bigg[4\frac{a''}{a}(\Phi+\Psi) + 2\mathcal{H}\Phi' + 4\mathcal{H}\Psi' - \partial^2\Psi + \partial^2\Phi + 2\Psi'' - 2\mathcal{H}^2(\Phi+\Psi)\bigg]\delta_{ij} \\ &+ \partial_i\partial_j\Psi - \partial_i\partial_j\Phi.\end{aligned} \tag{4.7.34}$$

Below, the purely temporal component of the perturbed Einstein tensor will be derived explicitly.

Derivation: Temporal Component Perturbed Einstein Tensor

To derive the purely temporal component of the perturbed Einstein tensor, we simply set the indices μ and ν to zero and substitute the expressions for $R, \delta R_{00}, g_{00}$ and δg_{00} to obtain:

$$\delta G_{00} = \delta R_{00} - \frac{1}{2}\delta g_{00} R - \frac{1}{2} g_{00} \delta R = -6\mathcal{H}\Psi' + 2\partial^2 \Psi. \tag{4.7.35}$$

4.8 Energy-Momentum Tensor

Here, the energy-momentum tensor will be perturbed to first order. This will be done first for the generic energy-momentum tensor for a (perfect) fluid. Before we do this, the definition of a perfect fluid and the associated local rest frame will be reviewed. Subsequently, the energy-momentum tensor for a single scalar field (Eq. 2.6.2) will be perturbed.

4.8.1 Perfect Fluid and the Local Rest Frame

A relativistic fluid is described by its energy-momentum tensor $T^{\mu\nu}$, and the dynamics (the *fluid flow*) is described by the EFE's in the considered background $g_{\mu\nu}$. From the EFE's, it follows that the energy-momentum tensor is symmetric $T^{\mu\nu} = T^{\nu\mu}$. By definition, a fluid is a matter component that is described by an energy-momentum tensor that smoothly varies as function of position.

In a *local rest frame*, the momentum density T^{0i} is defined to vanish and the purely temporal component T^{00} is defined as the energy density:

$$T^{0i} \equiv 0, \qquad \rho \equiv T^{00}. \quad \text{(Local Rest Frame)} \tag{4.8.1}$$

The fluid 4-velocity in the local rest frame is given by $U^\mu = (1, 0, 0, 0)$, so that the spatial components vanish $U^i = v^i = 0$. Notice that the local rest frame applies to the FRW universe, where the spatial 3-velocity vanishes by isotropy.

Local Rest Frame and Four-Velocity

Consider a worldline $x^\mu = x^\mu(\eta)$, parametrized by proper time η. For a generic frame, the four 4-velocity is defined as:

$$U^\mu \equiv \frac{dx^\mu}{d\eta} = (U^0, U^i), \tag{4.8.2}$$

where U^i is related to the 3-velocity v^i via $U^i = U^0 v^i$. The 4-velocity satisfies the constraint equation:

$$g_{\mu\nu} U^\mu U^\nu = -1. \tag{4.8.3}$$

Expanding the constraint equation and using $U^i = U^0 v^i$, we can obtain an explicit expression for the temporal component of the 4-velocity:

$$\begin{aligned} g_{00}U^0U^0 + g_{ij}U^iU^j &= -U^0U^0 + g_{ij}(U^0v^i)(U^0v^j) \\ &= U^0U^0(v^2 - 1) \equiv -1, \end{aligned} \tag{4.8.4}$$

where we defined the magnitude of the 3-velocity as: $v^2 \equiv v_i v^i$. The last equality can be solved for U^0:

$$U^0 = \frac{1}{\sqrt{1 - v^2}}. \tag{4.8.5}$$

Note that $U^0 = dt/d\eta$ and hence we derived the usual definition of the relativistic gamma factor $U^0 = \gamma$. In a generic frame, then, the 4-velocity can be written as:

$$U^\mu = (\gamma, \gamma \boldsymbol{v}), \tag{4.8.6}$$

where $\boldsymbol{v}$ denotes the 3-velocity. In terms of 4-velocity, a *local rest frame* can be defined as a frame where $\boldsymbol{v} \equiv \mathbf{0}$, so that:

$$U^\mu = (1, \mathbf{0}) = \delta^\mu_0, \tag{4.8.7}$$

and $U_\mu = -\delta^0_\mu$. In conformal time (not to be confused with proper time), the 4-velocity in the local rest frame read $U^\mu = \delta^\mu_0/a$ and $U_\mu = -a\delta^0_\mu$.

In terms of 4-velocity U^μ, the energy-momentum tensor for a fluid in a generic frame can be written as:

$$T^{\mu\nu} = (\rho + P)U^\mu U^\nu + P g^{\mu\nu} + \Sigma^{\mu\nu}. \tag{4.8.8}$$

Here, ρ is the energy density, P is the isotropic pressure and $\Sigma^{\mu\nu}$ the anisotropic stress contribution. Now, we move to the local rest frame ($\rho \equiv T^{00}$, $T^{0i} \equiv 0$) and examine the constraints of this frame on the energy-momentum tensor. Considering the completely temporal compenent:

$$T^{00} = (\rho + P)U^0U^0 - P + \Sigma^{00} \equiv \rho. \tag{4.8.9}$$

Hence, in the local rest frame (for which $U^0 = 1$), we find the constraint $\Sigma^{00} = 0$. Regarding the component T^{0i}, the constraint equation is:

$$T^{0i} = (\rho + P)U^i + \Sigma^{0i} \equiv 0, \tag{4.8.10}$$

yielding $\Sigma^{0i} = 0$ since $U^i = 0$. Therefore, in the local rest frame only the spatial part of the anisotropic stress contribution is non-zero:

$$\Sigma^{\mu\nu} = \begin{bmatrix} 0 & 0 \\ 0 & \Sigma^{ij} \end{bmatrix}. \tag{4.8.11}$$

Per definition, a *perfect fluid* is a fluid that is isotropic in the local rest frame, i.e. $\Sigma^{ij} = 0$. In other words, in a local rest frame, the energy-momentum tensor is diagonal for a perfect fluid. Note that the inflaton is a perfect fluid, since its background energy momentum tensor satisfies (Eq. 2.6.2):

$$\bar{T}^{ij} \propto \bar{g}^{ij} \propto \delta^{ij} \tag{4.8.12}$$

and therefore the spatial part is diagonal, constraining the anisotropic stress contribution to be zero ($\Sigma^{ij} = 0$).

More generally, we have shown in section 1.3.1 that the Cosmological Principle constrains the (background) energy-momentum tensor $\bar{T}_{\mu\nu}$ for any matter constituent to take the form of that for a perfect fluid. Hence, in the cosmological context, anisotropic stress contributions only enter at linear order in perturbations, as the background universe matter content can be described by perfect fluids. At first order in perturbations, the anisotropic stress contribution is written using the δ-notation, i.e. $\delta\Sigma_{ij}$. Even if anisotropic stress is not present at zeroth order in perturbations, it is not *a priori* obvious that anisotropic stress will be absent at linear (or higher) order in the perturbations. In section 4.8.3 will we show that a scalar field, which adheres the Cosmological Principle at zeroth order, indeed also does not develop anisotropic stress at linear order.

4.8.2 Perturbed Energy-Momentum Tensor for a Perfect Fluid

We will perturb the energy-momentum tensor in the local rest frame (i.e. we consider a perfect fluid), and use conformal time as the time variable. To first order, the perturbed tensor for a non-perfect fluid can be written as:

$$\delta T^{\mu}_{\nu} = (\delta\rho + \delta P)U^{\mu}U_{\nu} + (\rho + P)(\delta U^{\mu}U_{\nu} + U^{\mu}\delta U_{\nu}) + \delta P\delta^{\mu}_{\nu} + \delta\Sigma^{\mu}_{\nu}, \tag{4.8.13}$$

and we go to the local rest frame by setting $\delta\Sigma^{\mu}_{\nu} \equiv 0$. Since the spatial components of the background 4-velocity are zero in the local rest frame ($U^i = U_i = 0$), the first order perturbation is linear in velocity v_i and reads: [11]

$$\delta U^i = v^i/a, \qquad \delta U_i = -av^i. \tag{4.8.14}$$

The temporal perturbation to the 4-velocity δU^0 depends on the metric potential Φ and reads:

$$\delta U^0 = -\Phi/a, \qquad \delta U_0 = -a\Phi. \tag{4.8.15}$$

So that the total first order perturbation to the 4-velocity δU^{μ} and δU_{μ} are given by:

$$\delta U^{\mu} = \frac{1}{a}(-\Phi, v^i), \qquad \delta U_{\mu} = a(-\Phi, v^i). \tag{4.8.16}$$

Derivation: Perturbation to Temporal Four-Velocity Component

In order to derive δU^0 and δU_0, we will perturb the constraint equation to first order:

$$\delta g_{\mu\nu}U^{\mu}U^{\nu} + 2U_{\mu}\delta U^{\mu} = 0 \quad \longrightarrow \quad \delta g_{00}U^0U^0 + 2U_0\delta U^0 = 0. \tag{4.8.17}$$

Inserting the expressions $U^0 = 1/a$, $U_0 = -a$ and $\delta g_{00} = -2a^2\Phi$, we obtain:

$$\delta U^0 = -\Phi/a, \tag{4.8.18}$$

[11] Since at zeroth order there is no velocity component, we assume that the velocity term introduced at first order is much smaller than unity (recall $c \equiv 1$), so that second order terms can be neglected and the analysis remains at first order in perturbations.

as stated above. To obtain δU_0, we perturb the relation $U_\mu = g_{\mu\nu}U^\nu$ to first order:

$$\delta U_\mu = \delta g_{\mu\nu}U^\nu + g_{\mu\nu}\delta U^\nu. \tag{4.8.19}$$

Expanding the summation over ν and setting μ to zero, we obtain:

$$\delta U_0 = \delta g_{00}U^0 + g_{00}\delta U^0 = -a\Phi. \tag{4.8.20}$$

We are now in the position to compute the first order components δT^0_0, δT^i_0 and δT^i_j. They can be computed to give:

$$\delta T^0_0 = (\delta\rho + \delta P)U^0U_0 + (\rho + P)(\delta U^0U_0 + U^0\delta U_0) + \delta P\delta^0_0 = -\delta\rho, \tag{4.8.21}$$

$$\delta T^i_j = (\delta\rho + \delta P)U^iU_j + (\rho + P)(\delta U^iU_j + U^i\delta U_j) + \delta P\delta^i_j = \delta P\delta^i_j, \tag{4.8.22}$$

$$\begin{aligned}\delta T^i_0 &= (\delta\rho + \delta P)U^iU_0 + (\rho + P)(\delta U^iU_0 + U^i\delta U_0) + \delta P\delta^i_0 \\ &= (\rho + P)\delta U^iU_0 = -(\rho + P)v^i.\end{aligned} \tag{4.8.23}$$

In the first line, we used $U^0U_0 = -1$ and the fact that the terms δU^0U_0 and $U^0\delta U_0$ cancel. In the remaining lines, the expressions for U^μ, δU^μ and the lower index variants of these are simply inserted. In conclusion, the first order perturbations of the energy-momentum tensor δT^μ_ν for a perfect fluid in the local rest frame read:

$$\delta T^0_0 = -\delta\rho, \qquad \delta T^i_0 = -(\rho + P)v^i, \qquad \delta T^0_i = (\rho + P)v_i, \qquad \delta T^i_j = \delta P\delta^i_j. \tag{4.8.24}$$

Notice that in case we include anisotropic stress (i.e. the fluid is non-perfect), only the purely spatial components changes:

$$\delta T^i_j = \delta P\delta^i_j + \delta\Sigma^i_j, \tag{4.8.25}$$

as the anisotropic stress tensor only affects the purely spatial components in a local rest frame.[12]

Extension to Multiple Perfect Fluids

The extension to the perturbed energy-momentum tensor for multiple (non-interacting) perfect fluids can be made straightforwardly by replacing the perturbations by the sum of perturbations from the individual fluids. For perfect fluids labeled by α the total pressure and energy density of the system can be written in the following way:

$$\rho = \sum_{(\alpha)} \rho_{(\alpha)}, \qquad P = \sum_{(\alpha)} P_{(\alpha)}. \tag{4.8.26}$$

Here we use the notation (α) as the label for the considered fluids to emphasize that label does not correspond to a four-vector index. The perturbation to the energy density and the pressure can be written similarly as:

$$\delta\rho = \sum_{(\alpha)} \delta\rho_{(\alpha)}, \qquad \delta P = \sum_{(\alpha)} \delta P_{(\alpha)}. \tag{4.8.27}$$

[12] Recall that the FRW background indeed satisfies the definition of a local rest frame.

The velocity potential v^i, in combination with the energy density and pressure, can be written for multiple fluids as follows:

$$(\rho + P)v^i = \sum_{(\alpha)} (\rho_{(\alpha)} + P_{(\alpha)}) v^i_{(\alpha)}. \tag{4.8.28}$$

4.8.3 Perturbed Energy-Momentum Tensor for a Scalar Field

Here, the perturbed energy-momentum tensor for a scalar field will be obtained. Recall that the energy-momentum tensor for a scalar field (Eq. 2.4.16) is given by (Eq. 2.6.2):

$$T^{(\phi)}_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - g_{\mu\nu} \left[\frac{1}{2} g^{\alpha\beta} \partial_\alpha \phi \partial_\beta \phi + V(\phi) \right]. \tag{4.8.29}$$

Notice that the inflaton is a perfect fluid, since from the above energy-momentum tensor it follows that $T^{(\phi)}_{ij} \propto g_{ij}$ and hence the spatial part is therefore diagonal, enforcing the anisotropic stress contribution to be zero: $\Sigma_{ij} = 0$. Perturbing the above equation to first order in the scalar field and metric yields the following expression for $\delta T_{\mu\nu}$:

$$\begin{aligned} \delta T_{\mu\nu} = {} & 2\partial_\mu \delta\phi \partial_\nu \phi - \delta g_{\mu\nu} \left[\frac{1}{2} g^{\alpha\beta} \partial_\alpha \phi \partial_\beta \phi + V(\phi) \right] \\ & - g_{\mu\nu} \left[\frac{1}{2} \delta g^{\alpha\beta} \partial_\alpha \phi \partial_\beta \phi + g^{\alpha\beta} \partial_\alpha \phi \partial_\beta \delta\phi + V_\phi(\phi) \delta\phi \right]. \end{aligned} \tag{4.8.30}$$

Evaluating the above perturbed equation for the purely temporal, mixed and spatial components gives:

$$\begin{aligned} \delta T_{00} &= \phi' \delta\phi' + 2a^2 V \Phi + a^2 V_\phi \delta\phi, \\ \delta T_{0i} &= (\partial_i \delta\phi) \cdot \phi', \\ \delta T_{ij} &= \left[\phi' \delta\phi' - (\Psi + \Phi)\phi'^2 + 2a^2 V \Psi - a^2 V_\phi \delta\phi \right] \delta_{ij}. \end{aligned} \tag{4.8.31}$$

In the derivation box below, the purely temporal perturbation δT_{00} will be derived explicitly.

Derivation: Temporal Perturbation Scalar Field Energy Momentum Tensor

To compute the temporal perturbation δT_{00}, we set $\mu = \nu = 0$ in Eq. 4.8.30. The first term yields:

$$2\partial_0 \delta\phi \partial_0 \phi = 2\phi' \delta\phi'. \tag{4.8.32}$$

The second term is computed to be:

$$-\delta g_{00} \left[\frac{1}{2} g^{00} \partial_0 \phi \partial_0 \phi + V(\phi) \right] = -\Phi \phi'^2 + 2a^2 \Phi V, \tag{4.8.33}$$

where we used $g^{00} = -1/a^2$ and $\delta g_{00} = -2a^2 \Phi$. The last term gives:

$$-g_{00} \left[\frac{1}{2} \delta g_{00} \phi'^2 + g^{00} \partial_0 \partial_0 \delta\phi + V_\phi \delta\phi \right] = \Phi \phi'^2 - \phi' \delta\phi' + a^2 V_\phi \delta\phi. \tag{4.8.34}$$

Combining the above results yields the stated expression for the purely temporal perturbation:

$$\delta T_{00} = \phi' \delta\phi' + 2a^2 V \Phi + a^2 V_\phi \delta\phi. \tag{4.8.35}$$

For later convenience, we also provide the perturbed energy-momentum tensor in mixed form δT^{μ}_{ν} below:

$$\delta T^0_0 = \frac{1}{a^2}\Big[\Phi\phi'^2 - \phi'\delta\phi' - a^2 V_\phi \delta\phi\Big], \tag{4.8.36}$$

$$\delta T^0_i = -\frac{1}{a^2}(\partial_i \delta\phi)\cdot\phi', \tag{4.8.37}$$

$$\delta T^i_j = \frac{1}{a^2}\Big[\phi'\delta\phi' - \Phi\phi'^2 - a^2 V_\phi \delta\phi\Big]\delta^i_j. \tag{4.8.38}$$

Notice that the purely spatial perturbation to the energy-momentum tensor is diagonal, since $\delta T^i_j \propto \delta^i_j$. This result implies that even when perturbing the inflaton energy-momentum tensor, no anisotropic stress contribution is generated. To see this explicitly, consider Eq. 4.8.25, which yields:

$$\delta T^i_j = \delta P \delta^i_j + \delta\Sigma^i_j. \tag{4.8.39}$$

As we found $\delta T^i_j \propto \delta^i_j$ for the perturbed scalar field energy-momentum tensor, the anisotropic stress perturbation must be zero $\delta\Sigma^i_j = 0$. Later on, this result will be used to equate the two gravitational potentials Φ and Ψ in the Newtonian gauge perturbed metric.

Extension to Multiple Scalar Fields

In analogy with the extension to multiple perfect fluids, the scalar field can be extended to N minimally coupled non-interacting scalar fields. The single-field Lagrangian generalizes to:

$$\mathcal{L} = -\frac{1}{2}\sum_{(\alpha)} g^{\mu\nu}\partial_\mu\phi_{(\alpha)}\partial_\nu\phi_{(\alpha)} - V(\phi_1, \cdots \phi_N), \tag{4.8.40}$$

where the label (α) takes on the values $1, ..., N$ and V represents the potential of the multiple scalar fields. Now, the perturbed energy-momentum becomes simply:

$$\delta T^0_0 = \frac{1}{a^2}\sum_{(\alpha)}\Big[\Phi\phi'^2_{(\alpha)} - \phi'_{(\alpha)}\delta\phi'_{(\alpha)} - a^2 V_{(\alpha)}\delta\phi_{(\alpha)}\Big], \tag{4.8.41}$$

$$\delta T^0_i = -\frac{1}{a^2}\sum_{(\alpha)}(\partial_i\delta\phi_{(\alpha)})\cdot\phi'_{(\alpha)}, \tag{4.8.42}$$

$$\delta T^i_j = \frac{1}{a^2}\sum_{(\alpha)}\Big[\phi'_{(\alpha)}\delta\phi'_{(\alpha)} - \Phi{\phi'_{(\alpha)}}^2 - a^2 V_{(\alpha)}\delta\phi_{(\alpha)}\Big]\delta^i_j, \tag{4.8.43}$$

in which $V_{(\alpha)} \equiv \partial V/\partial\phi_{(\alpha)}$.

4.9 Einstein Field Equations

Now that we have the expressions for the perturbed Einstein tensor and the energy-momentum for perfect fluids and multiple scalar fields, we can construct the perturbed field equations to first order. Firstly, we will construct them for a single scalar field and then for a single perfect fluid. The extension to multiple fluids or scalar fields can be made straightforwardly by plugging in the corresponding energy-momentum tensors discussed above; the perturbed

Einstein tensor does not change in the case of multiple matter components. For convenience, the perturbed EFE's will constructed in mixed form, i.e. we compute:

$$\delta G^{\mu}_{\nu} = 8\pi G\ \delta T^{\mu}_{\nu}, \tag{4.9.1}$$

and will encompass the dimensional constants of nature from now on in terms of the Planck mass $M^2_{\rm pl} = 1/8\pi G$.

To start, we will express the perturbed Einstein tensor in manifestly mixed tensor form δG^{μ}_{ν}. This form can be obtained from $\delta G_{\mu\nu}$ as follows:

$$\delta G^{\mu}_{\nu} = \delta g^{\mu\alpha} G_{\alpha\nu} + g^{\mu\alpha}\delta G_{\alpha\nu}. \tag{4.9.2}$$

Using the perturbed metric tensor, the explicit components of δG^{μ}_{ν} can be derived to be:

$$\delta G^0_0 = \frac{1}{a^2}\Big[6\mathcal{H}^2\Phi + 6\mathcal{H}\Psi' - 2\partial_i\partial^i\Psi\Big], \tag{4.9.3}$$

$$\delta G^0_i = \frac{1}{a^2}\Big[-2\partial_i\Psi' - 2\mathcal{H}\partial_i\Phi\Big], \tag{4.9.4}$$

$$\begin{aligned}\delta G^i_j = {} & \frac{1}{a^2}\Big[4\frac{a''}{a}\Phi + 2\mathcal{H}\Phi' + 4\mathcal{H}\Psi' - \partial_k\partial^k\Psi + 2\Psi'' + \partial_k\partial^k\Phi - 2\mathcal{H}^2\Phi\Big]\delta^i_j \\ & + \frac{1}{a^2}\partial_j\partial^i(\Psi - \Phi).\end{aligned} \tag{4.9.5}$$

Notice that the off-diagonal components of the spatial perturbation δG^i_j $(i \neq j)$ are simply given by $\partial_j\partial^i(\Psi - \Phi)/a^2$.

4.9.1 Single Scalar Field

Here we will give the perturbed Einstein Field equations in mixed form for a flat universe sourced by a single scalar field (for the perturbed energy-momentum tensor see Eqs. 4.8.41–4.8.43). The different equations of motion will be discussed separately.

▷ *Purely Temporal Equation.*—First we consider the perturbed field equation for $(\mu,\nu) = (0,0)$. Equation the expressions for the perturbed energy-momentum tensor and the Einstein tensor gives:

$$\boxed{3\mathcal{H}\left(\Psi' + \mathcal{H}\Phi\right) - \partial_i\partial^i\Psi = -\frac{1}{2M^2_{\rm pl}}\left(\phi'\delta\phi' + a^2V_\phi\delta\phi - \Phi\phi'\right).} \tag{4.9.6}$$

▷ *Mixed Spatial-Temporal Equation.*—Now we consider the case with $(\mu,\nu) = (0,i)$. The arising equation is now:

$$\boxed{\Psi' + \mathcal{H}\Phi = \frac{1}{2M^2_{\rm pl}}\delta\phi\ \phi'.} \tag{4.9.7}$$

▷ *Purely Spatial Equation.*—Lastly, we consider the purely spatial equation, that is we equate δG^i_j with $\delta T^i_j/M^2_{\rm pl}$. The resulting equation reads:

$$\begin{aligned}\Big[4\frac{a''}{a}\Phi + 2\mathcal{H}\Phi' + 4\mathcal{H}\Psi' - \partial_k\partial^k\Psi + 2\Psi'' + \partial_k\partial^k\Phi - 2\mathcal{H}^2\Phi\Big]\delta^i_j + \frac{1}{a^2}\partial_j\partial^i(\Psi - \Phi) \\ = \frac{1}{M^2_{\rm pl}}\Big[\phi'\delta\phi' - \Phi\phi'^2 - a^2V_\phi\delta\phi\Big]\delta^i_j.\end{aligned} \tag{4.9.8}$$

This last equation may appear as rather complex, but a significant simplification can be obtained by taking the traceless ($i \neq j$) part of this equation, which yields:

$$\partial_j \partial^i (\Psi - \Phi) = 0. \tag{4.9.9}$$

Hence, as advocated, we can set the two gravitational potentials equal to each other $\Psi \equiv \Phi$ and choose to work with only one of them, which we will choose to be Φ. Notice that the fact that we can set the two gravitational potentials equal to each other comes from the fact that the perturbed energy-momentum tensor for the scalar field does not contain an anisotropic stress term $\delta\Sigma^i_j$ and hence δT^i_j is diagonal, i.e. proportional to the Kronecker delta δ^i_j (see Eq. 4.8.38). By setting $\Psi = \Phi$ and using $a''/a = \mathcal{H}' + \mathcal{H}^2$, we obtain:

$$\boxed{\Phi'' + 3\mathcal{H}\Phi' + (2\mathcal{H}' + \mathcal{H}^2)\Phi = \frac{1}{2M_{\rm pl}^2}\Big[\phi'\delta\phi' - \Phi\phi'^2 - a^2 V_\phi \delta\phi\Big].} \tag{4.9.10}$$

4.9.2 Single Perfect Fluid

Similar to the scalar field, we will now give the perturbed field equations for a generic perfect fluid (perturbed energy-momentum tensor is given by Eq. 4.8.24). Apart from the different energy-momentum tensor, the procedure is completely analogous. The resulting field equations are (see also [56]):

$$\partial^2\Phi - 3\mathcal{H}(\Phi' + \mathcal{H}\Phi) = \Delta(a)\delta\rho, \tag{4.9.11}$$

$$-\partial_i(\Phi' + \mathcal{H}\Phi) = \Delta(a)(\rho + P)v_i, \tag{4.9.12}$$

$$\Phi'' + 3\mathcal{H}\Phi' + (2\mathcal{H}' + \mathcal{H}^2)\Phi = \Delta(a)\delta P, \tag{4.9.13}$$

here we defined for purely notational convenience $\Delta(a) \equiv a^2/2M_{\rm pl}^2 = 4\pi G a^2$.

As the perturbed energy-momentum tensor for a perfect fluid remains diagonal, no anisotropic contribution is present and hence both gravitational potentials are set equal to each other as before. Furthermore, the second field equation (Eq. 4.9.12) can be simplified further by considering scalar perturbations only, the velocity perturbation v^i can then be expressed as the gradient of the velocity potential, i.e. $v_i = \partial_i v$, with v the velocity potential. After this substitution, Eq. 4.9.12 can be written as:

$$-(\Phi' + \mathcal{H}\Phi) = \Delta(a)(\rho + P)v. \tag{4.9.14}$$

4.9.3 Non-Perfect Fluid

Now we will consider the perturbed field equations for perturbations in a generic non-perfect fluid, with the non-anisotropic stress contribution parametrized by $\delta\Sigma^i_j$. The perturbed energy-momentum tensor δT^μ_ν is the same as for a perfect fluid (Eq. 4.8.24), except for the purely spatial component, which is given by Eq. 4.8.25. Notice that in this case, the two gravitational potentials Φ and Ψ are typically not equal:

$$\Phi \neq \Psi. \tag{4.9.15}$$

The perturbed field equations can be written as:

$$\boxed{\begin{aligned} \partial^2\Psi - 3\mathcal{H}(\Psi' + \mathcal{H}\Phi) &= \Delta(a)\delta\rho, && (4.9.16)\\ -\partial_i(\Psi' + \mathcal{H}\Phi) &= \Delta(a)(\rho + P)v_i, && (4.9.17)\\ \Psi'' + 2\mathcal{H}\Psi' + \mathcal{H}\Phi' + 2(a''/a)\Phi + \partial^2(\Phi - \Psi)/2 - \mathcal{H}^2\Phi &= \Delta(a)\delta P, && (4.9.18)\\ \partial_j\partial^i(\Psi - \Phi) &= 2\Delta(a)\delta\Sigma^i_j. && (4.9.19) \end{aligned}}$$

The last two equations above express the trace and trace-free parts of the (i, j) field equation, respectively.

4.10 Klein-Gordon Equation

Up till now, we have solely perturbed the Einstein field equations to first order. In addition, any other dynamical equation complementing the field equations should be perturbed as well. In particular, for a single scalar field in a generic space-time background as described by $g_{\mu\nu}$, the dynamics is governed by the Klein-Gordon equation (Eq. 2.5.2):

$$\Box\phi \equiv \frac{1}{\sqrt{-g}}\partial_\nu\left(\sqrt{-g}g^{\mu\nu}\partial_\mu\phi\right) = V_\phi. \tag{4.10.1}$$

Perturbing this equation gives:

$$\delta(\Box\phi) = V_{\phi\phi}\delta\phi. \tag{4.10.2}$$

Now, our task is to evaluate the perturbed d'Alemberatian operator $\delta(\Box\phi)$ using the first order perturbed FRW metric in the Newtonian Gauge. For generic $g_{\mu\nu}$, the perturbed d'Alemberatian is given by:

$$\begin{aligned} \delta(\Box\phi) = {} & \delta\Big(\frac{1}{\sqrt{-g}}\Big)\partial_\nu\left(\sqrt{-g}g^{\mu\nu}\partial_\mu\phi\right)\\ & + \frac{1}{\sqrt{-g}}\partial_\nu\left(\delta(\sqrt{-g})g^{\mu\nu}\partial_\mu\phi\right)\\ & + \frac{1}{\sqrt{-g}}\partial_\nu\left(\sqrt{-g}\,\delta(g^{\mu\nu})\partial_\mu\phi\right)\\ & + \frac{1}{\sqrt{-g}}\partial_\nu\left(\sqrt{-g}g^{\mu\nu}\partial_\mu(\delta\phi)\right). \end{aligned} \tag{4.10.3}$$

For perturbations around an FRW universe in Newtonian gauge, evaluating the above expression and substitution into the perturbed Klein-Gordon equation yields:

$$2\Phi\phi'' + (3\Psi' + \Phi' + 4\mathcal{H}\Phi)\phi' - \delta\phi'' - 2\mathcal{H}\delta\phi' + \partial^2(\delta\phi) = a^2\, V_{\phi\phi}\delta\phi. \tag{4.10.4}$$

To simplify this result, we will invoke the background Klein-Gordon equation in conformal time:

$$\phi'' + 2\mathcal{H}\phi' = -V_\phi a^2, \tag{4.10.5}$$

which can be derived analogous to Eq. 2.5.7 by using the conformal time metric. Now, we can rewrite the perturbed Klein-Gordon equation as:

$$(3\Psi' + \Phi')\phi' - \delta\phi'' - 2\mathcal{H}\delta\phi' + \partial^2(\delta\phi) = a^2\left(V_{\phi\phi}\delta\phi + 2\Phi V_\phi\right). \tag{4.10.6}$$

In the derivation box below, we will show how to evaluate the first term in the variation of d'Alembertian (Eq. 4.10.3).

Partial Derivation: Perturbed Klein-Gordon Equation

We will evaluate the first term in Eq. 4.10.3 by using the conformal time metric, for which $g \equiv \det g_{\mu\nu} = -a^8$ so that $\sqrt{-g} = a^4$. Specifically, we want to compute the term:

$$\delta(\Box\phi) \supset \delta\Big(\frac{1}{\sqrt{-g}}\Big)\partial_\nu\left(\sqrt{-g}g^{\mu\nu}\partial_\mu\phi\right). \tag{4.10.7}$$

The variation of $1/\sqrt{-g}$ can be computed using:

$$\delta\Big(\frac{1}{\sqrt{-g}}\Big) = \frac{\delta\sqrt{-g}}{g} = \frac{\delta g}{2(\sqrt{-g})^3}, \qquad \delta g = g\, g^{\alpha\beta}\delta g_{\alpha\beta}. \tag{4.10.8}$$

Evaluating the above expression for the conformal time FRW metric yields:

$$\delta\Big(\frac{1}{\sqrt{-g}}\Big) = -\frac{1}{2a^4}g^{\alpha\beta}\delta g_{\alpha\beta} = \frac{1}{a^4}(3\Psi - \Phi), \tag{4.10.9}$$

since $g^{\alpha\beta}\delta g_{\alpha\beta} = 2\Phi - 6\Psi$. Hence, we obtain:

$$\begin{aligned}\delta(\Box\phi) &\supset \delta\Big(\frac{1}{\sqrt{-g}}\Big)\partial_\nu\left(\sqrt{-g}g^{\mu\nu}\partial_\mu\phi\right)\\ &= \frac{1}{a^4}(3\Psi - \Phi)\partial_\nu\left(\sqrt{-g}\,\bar{g}^{\mu\nu}\partial_\mu\bar{\phi}\right),\end{aligned} \tag{4.10.10}$$

where we used the fact that the second line is already first order in perturbations, so the remaining variables can evaluated at zeroth order. At zeroth order, we find:

$$\sqrt{-g}\,\bar{g}^{\mu\nu}\partial_\mu\bar{\phi} = -a^2\phi'' - 2aa'\phi'. \tag{4.10.11}$$

Finally, we thus find that the first term in the variation of $\delta(\Box\phi)$ is given by:

$$\delta(\Box\phi) \supset \frac{1}{a^4}(\Phi - 3\Psi)(a^2\phi'' + 2aa'\phi'). \tag{4.10.12}$$

Similarly, the remaining three terms can be evaluated explicitly and adding them yields the l.h.s. of Eq. 4.10.4.

Part III

Quantum Effects during Single-Field Inflation

Chapter 5

Quantum Origin of Cosmological Perturbations

"One of the basic rules of the universe is that nothing is perfect. Perfection simply doesnt exist. Without imperfection, neither you nor I would exist."

— Stephen Hawking

In the previous chapter, we have considered the perturbations around the FRW space-time as being classical.[1] However, we know – as far as the single-field scenario is concerned – that the perturbations are generated by quantum fluctuations in the inflaton field during inflation. That is, the primordial seeds for structure formation are manifestly quantum mechanical in their nature (see section 3.1). Hence, in order to study the perturbations generated during inflation properly, we should describe them in the context of quantum field theory. In this chapter, the quantum origin of cosmological perturbations will be studied in detail. In particular, the main aims of this chapter are to derive the power spectrum of scalar and tensor (gravitational wave) inflationary perturbations ($\mathcal{P}_{\mathcal{R}}$ and $\mathcal{P}_E$) when they are frozen in. That is, we will compute $\mathcal{P}_{\mathcal{R}}$ and $\mathcal{P}_E$ a few e-foldings of expansion after horizon exit.

This chapter is organized as follows, first, we will use the evolution for the Gravitational potential in section 5.1 to find a direct relationship between the comoving curvature perturbation $\mathcal{R}$ and the inflaton fluctuation $\delta\phi$. Then, in section 5.2, we will derive the Mukhanov-Sasaki which described the dynamics of the rescaled inflation fluctuation $f = a\delta\phi$. Subsequently, in the next section, we will quantize the field f by means of canonical quantization. Using the constructed quantum theory of inflationary perturbations $\delta\phi$ we will then compute the power spectrum $\mathcal{P}_{\mathcal{R}}$ in section 5.4 (by using the relation between $\mathcal{R}$ and $\delta\phi$ obtained in section 5.1). We will briefly discuss the quantum to classical transition occurring outside the horizon in section 5.5. Finally, we compute the power spectrum for tensor modes in section 5.6.

5.1 Evolution of the Gravitational Potential

One of the primary tasks of this chapter will be to determine the power spectrum for the curvature perturbation $\mathcal{R}$ on super-horizon scales. However, the dynamical equation governing the evolution of the quantum fluctuations will be in terms of the inflaton fluctuation $\delta\phi$.[2]

[1] As we will show in this chapter, outside the horizon, inflationary perturbations may indeed be regarded as classical; see section 5.5.

[2] This dynamical equation is called the Mukhanov-Sasaki equation and will be derived in the next section (see Eq. 5.2.14). To be precise, the MS equation is written in terms of the rescaled field $f \equiv \delta\phi/a$.

Therefore, in order to compute the power spectrum of $\mathcal{R}$, we require a relation between $\delta\phi$ and $\mathcal{R}$ at horizon exit and beyond. In finding this relationship, we will find that the evolution of the gravitational potential plays an important role.

In terms of the gravitational potential Ψ and the inflaton fluctuation $\delta\phi$, the curvature perturbation can be written as (Eq. D.4.17):

$$\mathcal{R} = \Psi + \mathcal{H}\frac{\delta\phi}{\phi'}. \tag{5.1.1}$$

As we will compute the power spectrum on super-horizon scales, we require an expression for $\mathcal{R}$ in this limit. We will show that outside the horizon, we can write the gravitational potential as:

$$\Psi = \varepsilon\mathcal{H}\frac{\delta\phi}{\phi'}, \tag{5.1.2}$$

and hence, to zeroth order in ε, we find that:

$$\mathcal{R} = \mathcal{H}\frac{\delta\phi}{\phi'} + \mathcal{O}(\varepsilon). \tag{5.1.3}$$

To show this, we will need to prove that Ψ is constant on super-horizon scales, i.e. $\Psi' = 0$. This fact will be proven by first solving for the gravitational potential to first in Hubble flow parameters and then consider the super-horizon limit.

During the inflationary stage dominated by a single scalar field, the evolution equation can be written in Fourier space as:

$$\Psi_k'' + 2\left(\mathcal{H} - \frac{\phi''}{\phi'}\right)\Psi_k' + 2\left(\mathcal{H}' - \mathcal{H}\frac{\phi''}{\phi'}\right)\Psi_k + k^2\Psi_k = 0. \tag{5.1.4}$$

The above equation is difficult to solve in full generality as, apart from Ψ, the Hubble parameter $\mathcal{H}$ and inflaton field ϕ in principle also change as a function of time. To simplify the equation of motion, we recall that the background during the inflationary era is well approximated as a quasi De-Sitter space-time. Hence, it will be insightfull to rewrite the above evolution equation for Ψ in terms of the Hubble parameters ε and η. The resulting equation, which will be derived below, reads:

$$\Psi_k'' + 2\mathcal{H}(\eta - \varepsilon)\Psi_k' + 2\mathcal{H}^2(\eta - 2\varepsilon)\Psi_k + k^2\Psi_k = 0. \tag{5.1.5}$$

Derivation: Evolution Equation Gravitational Potential

Here, we will derive Eqs. 5.1.4 and 5.1.5 using the perturbed field equations for a single scalar field (the inflaton) as given in section 4.9.1. The relevant equations are:

$$\begin{aligned} 3\mathcal{H}(\Psi' + \mathcal{H}\Psi) - \partial^2\Psi &= -(2M_{\rm pl}^2)^{-1}\big(\phi'\delta\phi + a^2V_\phi\delta\phi - \Psi\phi'^2\big), \\ \Psi'' + 3\mathcal{H}\Psi' + (2\mathcal{H}' + \mathcal{H}^2)\Psi &= (2M_{\rm pl}^2)^{-1}\big(\phi'\delta\phi - a^2V_\phi\delta\phi - \Psi\phi'^2\big), \\ \Psi' + \mathcal{H}\Psi &= (2M_{\rm pl}^2)^{-1}\phi'\delta\phi. \end{aligned} \tag{5.1.6}$$

Since we are dealing with a scalar field that satisfies the conditions of a perfect fluid, anisotropic stress is absent and the two gravitational potentials are equal: $\Phi = \Psi$. Adding the first two of the above equations yields:

$$\Psi'' + 6\mathcal{H}\Psi' + 2\mathcal{H}'\Psi + 4\mathcal{H}^2\Psi - \partial^2\Psi = -2\Delta(a)V_\phi\delta\phi. \tag{5.1.7}$$

The r.h.s. can be written in terms of graviational potentials as well by invoking the third field equations, which can be rewritten as:

$$\delta\phi = \frac{2M_{\rm pl}^2}{\phi'}(\Psi' + \mathcal{H}\Psi). \tag{5.1.8}$$

The potential derivative V_ϕ can be eliminated in favor of ϕ by means of the Klein-Gordon equation in conformal time:

$$\phi'' + 2\mathcal{H}\phi' = -a^2 V_\phi. \tag{5.1.9}$$

The r.h.s. of Eq. 5.1.7 then becomes:

$$-2\Delta(a)V_\phi\delta\phi = 2\left(\frac{\phi''}{\phi'} + 2\mathcal{H}\right)\Psi' + 2\mathcal{H}\left(\frac{\phi''}{\phi'} + 2\mathcal{H}\right)\Psi. \tag{5.1.10}$$

After moving all terms to the l.h.s. and simplifying we obtain the advocated evolution equation for the gravitational potential:

$$\Psi_k'' + 2\left(\mathcal{H} - \frac{\phi''}{\phi'}\right)\Psi_k' + 2\left(\mathcal{H}' - \mathcal{H}\frac{\phi''}{\phi'}\right)\Psi_k + k^2\Psi_k = 0, \tag{5.1.11}$$

where we moved to Fourier space, which amounts to replacing $\partial^2 \to -k^2$. Now we proceed and rewrite the above equation in terms of the Hubble flow parameters, which are defined in terms of the inflation field as:

$$\varepsilon = \frac{\phi'^2}{2M_{\rm pl}^2\mathcal{H}^2}, \qquad \delta = -\frac{\phi''}{\mathcal{H}\phi'}, \qquad \eta = 2(\varepsilon - \delta), \qquad \eta - \varepsilon = 1 - \frac{\phi''}{\mathcal{H}\phi'}. \tag{5.1.12}$$

The prefactors of Ψ_k' and Ψ_k can be obtained using the above relations as:

$$\mathcal{H} - \frac{\phi''}{\phi'} = \mathcal{H}(\eta - \varepsilon), \qquad \mathcal{H}' - \mathcal{H}\frac{\phi''}{\phi'} = \mathcal{H}^2(\eta - 2\varepsilon). \tag{5.1.13}$$

Substituting these results yields the second advocated equation (Eq. 5.1.5) for Ψ in terms of the Hubble flow parameters:

$$\Psi_k'' + 2\mathcal{H}(\eta - \varepsilon)\Psi_k' + 2\mathcal{H}^2(\eta - 2\varepsilon)\Psi_k + k^2\Psi_k = 0. \tag{5.1.14}$$

5.1.1 Solution to the Gravitational Potential

To find an approximate solution to the evolution of the gravitational potential Ψ, we solve Eq. 5.1.5 by assuming the Hubble parameters are small (relative to unity) and time-independent:

$$\varepsilon, \eta \ll 1, \qquad \varepsilon' = \eta' = 0, \tag{5.1.15}$$

in accordance with the slow-roll approximation (see section 2.7). The analysis here will be performed to first order in Hubble parameters, so that terms of order $\mathcal{O}(\varepsilon^2, \eta^2)$ can be neglected. As the prefactors in Eq. are already first order, it suffices to express $\mathcal{H}$ as function

of conformal time to zeroth order. Then, $\mathcal{H}$ is simply given by:

$$\mathcal{H} = -\frac{1}{\tau} + \mathcal{O}(\varepsilon, \eta), \tag{5.1.16}$$

during the inflationary stage, note furthermore that during inflation τ runs over *negative* values. The equation of motion for Ψ can then be written as follows:

$$\Psi_k'' - \frac{2}{\tau}(\eta - \varepsilon)\Psi_k' + \frac{2}{\tau^2}(\eta - 2\varepsilon)\Psi_k + k^2\Psi_k = 0. \tag{5.1.17}$$

The solution to this equation can be written in terms of Hankel functions of the first and second kind:

$$H_\mu^{(1)}(x) \equiv J_\mu(x) + iY_\mu(x), \qquad H_\mu^{(2)}(x) \equiv J_\mu(x) - iY_\mu(x), \tag{5.1.18}$$

where $J_\mu(x)$ and $Y_\mu(x)$ are Bessel functions of the first and second kind, respectively, and for real arguments $H_\mu^{(2)}(x) = H_\mu^{(1)}(x)^*$. Explicitly, the solution reads:

$$\Psi_k(\tau) = \sqrt{-\tau}\left[c_1 H_\mu^{(1)}(-k\tau) + c_2 H_\mu^{(2)}(-k\tau)\right], \tag{5.1.19}$$

where $c_{1,2}$ are constants to be determined by initial conditions at very early times (corresponding to the limit $\tau \to -\infty$) and μ is a function of the Hubble flow parameters. To first order:

$$\mu(\varepsilon, \eta) = \frac{1}{2}\sqrt{1 + 12\varepsilon - 4\eta} + \mathcal{O}(\varepsilon^2, \eta^2). \tag{5.1.20}$$

In order to determine the constants $c_{1,2}$, we consider the $\tau \to -\infty$ limit of the equation of motion (Eq. 5.1.5), yielding the differential equation for a simple harmonic oscillator with solution:

$$\tilde{\Psi}_k'' + k^2\tilde{\Psi}_k = 0, \qquad \tilde{\Psi} = \mathcal{C}_1 e^{ik\tau} + \mathcal{C}_2 e^{-ik\tau}. \tag{5.1.21}$$

The tilde expresses the fact that we consider the early time limit of the gravitational field, that is:

$$\tilde{\Psi} \equiv \lim_{\tau \to -\infty} \Psi(\tau). \tag{5.1.22}$$

In the context of QFT, the constants $\mathcal{C}_{1,2}$ are usually determined by constructing the quantum vacuum state of the field. Assuming the field is in its vacuum state then provides the expressions for $\mathcal{C}_{1,2}$.[3] However, the analysis to be performed will semi-classical, since the fluctuations in the inflaton field will be quantized whereas the gravitational background, as well as its metric perturbations (such as Ψ and Φ), will remain classical.[4]

Despite the fact that there is no trivial way to determine way the constants $\mathcal{C}_{1,2}$, we will still be able to show that Ψ_k does not evolve outside the horizon. In order to do so, we first connect the constants $\mathcal{C}_{1,2}$ and $c_{1,2}$, respectively. In the limit $\tau \to -\infty$, the asymptotic behavior of the first Hankel function is given by [57]:

$$H_\mu^{(1)} = \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{-k\tau}}e^{+ik\tau}\Delta^{(-)}, \tag{5.1.23}$$

[3] This is exactly the procedure to be applied in finding the initial conditions of the inflaton fluctuation field f and will be discussed in great detail in sections 5.3 and 5.4 when deriving the power spectrum of inflationary fluctuations.

[4] We regard the metric perturbations (e.g. Ψ) as classical entities throughout the whole analysis since quantizing them would require a consistent quantum theory of gravity, which is not available yet.

with $\Delta^{(\pm)} \equiv \exp(i\pi(\mu + 1/2)/2)$ and the asymptotic behavior of the second Hankel function is given by complex conjugation. The generic solution can then be written as follows deep inside the horizon:

$$\Psi_k = \sqrt{\frac{2}{\pi k}} \left[c_1 \, e^{ik\tau} \Delta^{(-)} + c_2 \, e^{-ik\tau} \Delta^{(+)} \right]. \tag{5.1.24}$$

As this expression should coincide with the solution $\tilde{\Psi}$ (Eq. 5.1.21), we find that the two sets of constants are related as:

$$c_1 = \sqrt{\frac{\pi k}{2}} \, \mathcal{C}_1 \, \Delta^{(+)}, \qquad c_2 = \sqrt{\frac{\pi k}{2}} \, \mathcal{C}_2 \, \Delta^{(-)}. \tag{5.1.25}$$

Therefore, in terms of early time constants $\mathcal{C}_{1,2}$, the generic solution to Eq. 5.1.5 reads:

$$\Psi_k(\tau) = \sqrt{\frac{\pi}{2}} \sqrt{-k\tau} \left[\mathcal{C}_1 \, \Delta^{(+)} H_\mu^{(1)}(-k\tau) + \mathcal{C}_2 \, \Delta^{(-)} H_\mu^{(2)}(-k\tau) \right], \tag{5.1.26}$$

to first order in Hubble flow parameters.

5.1.2 Super-Horizon Limit

To proceed, we take the super-horizon limit of the gravitational potential by considering the asymptotic behavior of the (first) Hankel function in the limit $\tau \to 0$:

$$H_\mu^{(1)} = \sqrt{\frac{2}{\pi}} \, e^{-i\pi/2} \, 2^{\mu - 3/2} \frac{\Gamma(\mu)}{\Gamma(3/2)} (-k\tau)^{-\mu}. \tag{5.1.27}$$

Assuming slow-roll, so that $\varepsilon, \eta \ll 1$, we can approximate $\mu \simeq 1/2$ and hence the asymptotic behavior becomes:

$$H_\mu^{(1)} = \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{-k\tau}} \, e^{-i\pi/2}, \tag{5.1.28}$$

and the second Hankel function is again given by complex conjugation. Using the fact that $\Delta^{\pm} = e^{\pm i\pi/2}$ in the limit $\mu \to 1/2$ yields (Eq. 5.1.26):

$$\Psi_k = \mathcal{C}_1 + \mathcal{C}_2, \qquad (\tau \to 0) \tag{5.1.29}$$

where $\mathcal{C}_{1,2}$ are time-independent constants. Hence, taking the conformal time derivative of the gravitational potential outside the horizon yields zero:

$$\Psi' = 0. \tag{5.1.30}$$

This formal analysis confirms the rough argument[5] given by Riotto [74] to show that Ψ does not evolve outside the horizon.

[5] Riotto argues that the time derivative of the gravitational potential is proportional to a function of Hubble flow parameters $f(\varepsilon, \eta)$ and can therefore be neglected: $\Psi \propto f(\varepsilon, \eta) \to 0$.

5.1.3 Relation between $\mathcal{R}$ and $\delta\phi$

On account of the result $\Psi' = 0$ outside the horizon, we can use the perturbed field equation (see the previous derivation box):

$$\Psi' + \mathcal{H}\Psi = \frac{1}{2M_{\rm pl}^2}\phi'\delta\phi = \varepsilon\mathcal{H}^2\frac{\delta\phi}{\phi'}, \tag{5.1.31}$$

where in the last equality we invoked $\phi'^2 = 2\varepsilon M_{\rm pl}^2\mathcal{H}^2$. Since outside the horizon the gravitational potential does not evolve, we conclude that:

$$\Psi = \varepsilon\mathcal{H}\frac{\delta\phi}{\phi'}. \tag{5.1.32}$$

Now substituting this result into Eq. D.4.17 gives, to zeroth order in ε, the advocated result:

$$\mathcal{R} = \mathcal{H}\frac{\delta\phi}{\phi'} + \mathcal{O}(\varepsilon). \tag{5.1.33}$$

In terms of the rescaled perturbation $f \equiv a\delta\phi$, we obtain:

$$\mathcal{R} = \frac{f}{z} + \mathcal{O}(\varepsilon), \qquad z \equiv a\phi'/\mathcal{H}. \tag{5.1.34}$$

This relation will be of significant importance deriving the power spectrum in section 5.4.

5.2 Mukhanov-Sasaki Equation

In the previous chapter, the perturbed Klein-Gordon (KG) equation for single field inflationary models described by a potential $V(\phi)$ is derived in conformal time as (Eq. 4.10.6):

$$4\Phi'\phi' - 2\mathcal{H}\delta\phi' - \delta\phi'' + \partial_i\partial^i\delta\phi = a^2\left(\delta\phi V_{\phi\phi} + 2\Phi V_\phi\right), \tag{5.2.1}$$

where V_ϕ and $V_{\phi\phi}$ denote the first and second derivative of the potential with respect to the inflaton field and we have set $\Phi = \Psi$. From this equation, the Mukhanov-Sasaki equation can be derived. In order to this, it proves convenient to first transform the perturbed KG equation to cosmic time.

For a generic quantity $g(\tau)$, the first and second conformal time derivatives are related to cosmic time derivatives as follows:

$$g' = a\dot{g}, \qquad g'' = a^2\ddot{g} + a^2H\dot{g}. \tag{5.2.2}$$

Therefore, the perturbed KG equation can be rewritten in cosmic time as:

$$4a^2\dot{\Phi}\dot{\phi} - 2a^2H\delta\dot{\phi} - a^2\delta\ddot{\phi} - a^2H\delta\dot{\phi} + \partial_i\partial^i\delta\phi = a^2\left(\delta\phi V_{\phi\phi} + 2\Phi V_\phi\right) \tag{5.2.3}$$

Now moving to the k-space representation by means of a Fourier transformation yields:

$$\boxed{\delta\ddot{\phi}_k + 3H\delta\dot{\phi}_k + \frac{k^2}{a^2}\delta\phi_k + V_{\phi\phi}\delta\phi_k = 4\dot{\Phi}_k\dot{\phi}_k - 2\Phi_k V_\phi.} \tag{5.2.4}$$

To proceed, the above equation can be simplified using the results of the previous section. In the previous section, we found that for slow roll inflation $\dot{\Phi}_k = 0$ outside the horizon.

Furthermore, we found in the previous section that the gravitational potential can be written in terms of the inflaton fluctuation as follows (using $\Phi = \Psi$):

$$\Psi_k = \Phi_k = \varepsilon H \frac{\delta\phi_k}{\dot{\phi}}. \tag{5.2.5}$$

Hence, using that in the slow-roll approximation $V_\phi = -3H\dot{\phi}$, the product $\Phi_k V_\phi$ can be written as:

$$\Phi_k V_\phi = \varepsilon H V_\phi \frac{\delta\phi_k}{\dot{\phi}} = -3\varepsilon H^2 \delta\phi_k. \tag{5.2.6}$$

The perturbed KG equation can then be written as:

$$\delta\ddot{\phi}_k + 3H\delta\dot{\phi}_k + \left[\frac{k^2}{a^2} + V_{\phi\phi} - 6\varepsilon H^2\right]\delta\phi_k = 0. \tag{5.2.7}$$

This expression can simplified even further by using that during slow-roll inflation $\varepsilon \ll 1$ (equivalently, one may use that $\varepsilon H^2 = -\dot{H} \simeq 0$ during inflation) and therefore the last term between brackets can be neglected safely. In order to find the evolution equation for the rescaled fluctuation in conformal time, we could now transform the above equation of motion back to f as a function of τ, which would yield the Mukhanov-Sasaki equation.[6]

Nevertheless, we will take a more clear and direct approach by expanding the inflaton-gravity action (Eq. 2.4.17) directly to second order in perturbations and identifying the Mukhanov-Sasaki action from this expansion. That is, we will expand the inflaton-gravity action:

$$S = S_{\text{EH}} + S_\phi = \int d^4x\sqrt{-g}\left[\frac{1}{2}M_{\text{pl}}^2 R - \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right], \tag{5.2.10}$$

in terms of the rescaled fluctuation $\phi(\boldsymbol{x},\tau) = \bar{\phi}(t) + f(\boldsymbol{x},\tau)/a(\tau)$. Isolating the second order terms in f gives:[7]

$$S^{(2)} = \frac{1}{2}\int d\tau\, d^3x\left[(f')^2 - (\partial f)^2 - 2\mathcal{H}ff' + (\mathcal{H}^2 V_{\phi\phi})\right]. \tag{5.2.11}$$

To simplify this result we will integrate the term proportional to $ff' = (f')^2/2$ by parts, yielding:

$$S^{(2)} = \frac{1}{2}\int d\tau\, d^3x\left[(f')^2 - (\partial f)^2 + \left(\frac{a''}{a} - a^2 V_{\phi\phi}\right)f^2\right], \tag{5.2.12}$$

where we have used the relation $a''/a = \mathcal{H}^2 + \mathcal{H}'$. During slow-roll inflation the ratio of the second field derivative of the potential over the Hubble parameter satisfies $V_{\phi\phi}/H^2 = 3\eta_V \ll 1$

[6]In this procedure, on has to compute the tedious transformation rules between the second order time derivatives of f in cosmic and conformal time, reading:

$$\delta\dot{\phi} = \frac{f'}{a^2} - \frac{a'f}{a^3}, \tag{5.2.8}$$

$$\delta\ddot{\phi} = \frac{f''}{a^3} - 2\frac{a'f'}{a^4} - \frac{a''f_k}{a^4} - \frac{a'f'}{a^4} + 3\frac{a'^2 f}{a^5}. \tag{5.2.9}$$

Using those relations, one could in principle find the correct Mukhanov-Sasaki equation.

[7]Expanding the action solely to first order in f, i.e. considering $S^{(1)}$ would give rise to the Klein-Gordon equation for the background field $\bar{\phi}$.

and $a''/a \simeq 2a^2H^2 \gg a^2V_{\phi\phi}$. Therefore, we can neglect the term proportional to $V_{\phi\phi}$ in the above action $S^{(2)}$, so the action reads:

$$S[f, f'] = \frac{1}{2}\int d\tau\, d^3x \left[(f')^2 - (\partial f)^2 - \frac{a''}{a}f^2\right]. \tag{5.2.13}$$

By the variational principle, this action gives rise to the Mukhanov-Sasaki equation. Working in momentum space, the Mukhanov-Sasaki equation reads:

$$\boxed{f_k'' + \left[k^2 - \frac{a''}{a} + a^2V_{\phi\phi}\right]f_k = 0,} \tag{5.2.14}$$

which describes the evolution fo the rescaled inflation fluctuation f. In particular, quantization of this action enables us to predict the power spectrum of inflaton fluctuations (see next sections).

5.3 Quantum Field Theory of Inflationary Perturbations

Up till now, the inflaton fluctuation is considered as a classical quantity, while has a manifestly quantum mechanical origin. In this section, we will quantize the theory of inflationary perturbations to examine the consequences of f being a quantum field, rather than a classical field. In particular, we will derive that the 2-point correlation function $\langle 0||\hat{f}|^2|0\rangle$ is non-vanishing, reflecting on the fact that the variance of the quantum fluctuations is non-zero. In close relationship to the 2-point correlation function, we will introduce the power spectrum of inflationary perturbations, which plays an essential role in relating inflationary predictions to late time observables.

The theory as given by the above action will be quantized according to the method of canonical quantization. The approach taken here is very similar to the method presented in [82] to canonically quantize a scalar field. First, the conjugate momentum of f is given by:

$$\pi \equiv \frac{\partial\mathcal{L}(f, f')}{\partial f'} = f'. \tag{5.3.1}$$

Now, we promote the classical fields f and π to quantum operator fields, denoted with hats:

$$f \longrightarrow \hat{f}, \quad \pi \longrightarrow \hat{\pi}. \tag{5.3.2}$$

We impose equal time canonical commutation relations (CCR) on the fields by means of the commutator:

$$[\hat{f}(\tau, \boldsymbol{x}), \hat{\pi}(\tau, \boldsymbol{y})] \equiv i\delta^{(3)}(\boldsymbol{x} - \boldsymbol{y}). \tag{5.3.3}$$

The above commutator reflects on *locality*: modes at different spatial locations ($\boldsymbol{x}$ and $\boldsymbol{y}$) evolve independently and hence the corresponding operators commute. In Fourier space, the CCR condition becomes:

$$[\hat{f}_{\boldsymbol{k}}(\tau), \hat{\pi}_{\boldsymbol{k}'}(\tau)] = i\delta^{(3)}(\boldsymbol{k} + \boldsymbol{k}'). \tag{5.3.4}$$

Derivation: CCR Condition in Fourier Space

We will now derive the canonical commutation relation in Fourier space explicitly:

$$\begin{aligned}[\hat{f}_{\boldsymbol{k}}(\tau), \hat{\pi}_{\boldsymbol{k}'}(\tau)] &= \int d^3\boldsymbol{x} \int d^3\boldsymbol{y}\, [\hat{f}(\tau, \boldsymbol{x}), \hat{\pi}(\tau, \boldsymbol{y})]\, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} e^{-i\boldsymbol{k}'\cdot\boldsymbol{y}} \\ &= i \int d^3\boldsymbol{x} \int d^3\boldsymbol{y}\, \delta^{(3)}(\boldsymbol{x}-\boldsymbol{y}) e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} e^{-i\boldsymbol{k}'\cdot\boldsymbol{y}} \\ &= i \int d^3\boldsymbol{x}\, e^{-i(\boldsymbol{k}+\boldsymbol{k}')\cdot\boldsymbol{x}} \equiv i\delta^{(3)}(\boldsymbol{k}+\boldsymbol{k}'). \end{aligned} \tag{5.3.5}$$

The operator $\hat{f}_k$ in Fourier space can be expanded in terms of a single time-independent operator $\hat{a}_{\boldsymbol{k}}$, its Hermitian conjugate and the complex mode function $f_k(\tau)$ as follows [14]:

$$\hat{f}_{\boldsymbol{k}} = f_k(\tau)\hat{a}_{\boldsymbol{k}} + f_k^*(\tau)\hat{a}_{\boldsymbol{k}}^\dagger. \tag{5.3.6}$$

Here, the operators $\hat{a}_{\boldsymbol{k}}$ and $\hat{a}_{\boldsymbol{k}}^\dagger$ may be regarded as annihilation and creation operators. They are defined to act on the vacuum state of the field, denoted by $|0\rangle$, in the following way [14]:

$$\hat{a}_{\boldsymbol{k}}|0\rangle = \langle 0|\hat{a}_{\boldsymbol{k}}^\dagger = 0, \tag{5.3.7}$$

and excited states are generated by multiple applications of the raising operator [14]:

$$|m_{\boldsymbol{k}_1}, n_{\boldsymbol{k}_2,\ldots}\rangle = \frac{1}{\sqrt{m!n!\cdots}} \left[(\hat{a}_{\boldsymbol{k}_1}^\dagger)^m (\hat{a}_{\boldsymbol{k}_2}^\dagger)^n \cdots\right] |0\rangle, \tag{5.3.8}$$

where $m_{\boldsymbol{k}_1}$ denotes the number of particles with momentum $\boldsymbol{k}_1$. The square root prefactor accounts for proper normalization. The creation and annihilation operators satisfy the following commutation relations:

$$[\hat{a}_{\boldsymbol{k}}, \hat{a}_{\boldsymbol{q}}] = [\hat{a}_{\boldsymbol{k}}^\dagger, \hat{a}_{\boldsymbol{q}}^\dagger] = 0, \qquad [\hat{a}_{\boldsymbol{k}}, \hat{a}_{\boldsymbol{q}}^\dagger] = [\hat{a}_{\boldsymbol{q}}, \hat{a}_{\boldsymbol{k}}^\dagger] = (2\pi)^3\delta^{(3)}(\boldsymbol{k}+\boldsymbol{q}), \tag{5.3.9}$$

where the factor of $(2\pi)^3$ is included due to the obeyed Fourier convention. In terms of the creation and annihilation operators ($\hat{a}_{\boldsymbol{k}}$, $\hat{a}_{\boldsymbol{k}}^\dagger$), the field $\hat{f}$ and its conjugate momentum $\hat{\pi}$ read:

$$\hat{f}(\tau, \boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \left[f_k(\tau)\hat{a}_{\boldsymbol{k}} + f_k^*(\tau)\hat{a}_{\boldsymbol{k}}^\dagger\right] e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{5.3.10}$$

$$\hat{\pi}(\tau, \boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \left[f_k'(\tau)\hat{a}_{\boldsymbol{k}} + (f_k^*(\tau))'\hat{a}_{\boldsymbol{k}}^\dagger\right] e^{i\boldsymbol{k}\cdot\boldsymbol{x}}. \tag{5.3.11}$$

To proceed, we will examine what constraint the canonical commutation relations will impose on the modes functions $f_k(\tau)$ and its temporal derivative $f_k' \equiv \partial_\tau f_k$. Substitution of the above expressions for $\hat{f}$ and $\hat{\pi}$ into the commutator $[\hat{f}(\tau, \boldsymbol{x}), \hat{\pi}(\tau, \boldsymbol{y})]$ and imposing the CCR condition to hold (Eq. 5.3.3), gives rise to the Wroskian normalization condition on the mode functions:

$$W[f_k, f_k^*] = (-i)\left[f_k(f_{-k}^*)' - f_k^* f_{-k}'\right] \equiv 1. \tag{5.3.12}$$

The derivation of this normalization condition on $W[f_k, f_k^*]$ is shown in the box below.

Derivation: Wroskian Normalization Condition

We start out by explicitly evaluating the commutator $[\hat{f}(\tau,\boldsymbol{x}),\hat{\pi}(\tau,\boldsymbol{y})]$ as follows:

$$[\hat{f}(\tau,\boldsymbol{x}),\hat{\pi}(\tau,\boldsymbol{y})] = \hat{f}(\tau,\boldsymbol{x})\hat{\pi}(\tau,\boldsymbol{y}) - \hat{\pi}(\tau,\boldsymbol{x})\hat{f}(\tau,\boldsymbol{y}) \equiv (*) - (**). \tag{5.3.13}$$

In terms of momentum integrals, creation/annihilation operators and the mode function, the two terms, $(*)$ and $(**)$, can be expanded as:

$$\begin{aligned}
(*) &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\int \frac{d^3\boldsymbol{q}}{(2\pi)^3}\left[f_k\hat{a}_{\boldsymbol{k}} + f_k^*\hat{a}_{\boldsymbol{k}}^\dagger\right]\left[f_q'\hat{a}_{\boldsymbol{q}} + (f_q^*)'\hat{a}_{\boldsymbol{q}}^\dagger\right]\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\, e^{i\boldsymbol{q}\cdot\boldsymbol{y}} \\
&= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\int \frac{d^3\boldsymbol{q}}{(2\pi)^3}\Big[f_kf_q'\hat{a}_{\boldsymbol{k}}\hat{a}_{\boldsymbol{q}} + f_k(f_q^*)'\hat{a}_{\boldsymbol{k}}\hat{a}_{\boldsymbol{q}}^\dagger \\
&\qquad + f_k^*f_q'\hat{a}_{\boldsymbol{k}}^\dagger\hat{a}_{\boldsymbol{q}} + f_k^*(f_q^*)'\hat{a}_{\boldsymbol{k}}^\dagger\hat{a}_{\boldsymbol{q}}^\dagger\Big]\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\, e^{i\boldsymbol{q}\cdot\boldsymbol{y}}.
\end{aligned} \tag{5.3.14}$$

$$\begin{aligned}
(**) &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\int \frac{d^3\boldsymbol{q}}{(2\pi)^3}\Big[f_kf_q'\hat{a}_{\boldsymbol{q}}\hat{a}_{\boldsymbol{k}} + f_k(f_q^*)'\hat{a}_{\boldsymbol{q}}^\dagger\hat{a}_{\boldsymbol{k}} \\
&\qquad + f_k^*f_q'\hat{a}_{\boldsymbol{q}}\hat{a}_{\boldsymbol{k}}^\dagger + f_k^*(f_q^*)'\hat{a}_{\boldsymbol{q}}^\dagger\hat{a}_{\boldsymbol{k}}^\dagger\Big]\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\, e^{i\boldsymbol{q}\cdot\boldsymbol{y}}.
\end{aligned} \tag{5.3.15}$$

Combining the two obtained expressions above yields:

$$\begin{aligned}
[\hat{f}(\tau,\boldsymbol{x}),\hat{\pi}(\tau,\boldsymbol{y})] &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\int \frac{d^3\boldsymbol{q}}{(2\pi)^3}\Big[[\hat{a}_{\boldsymbol{k}},\hat{a}_{\boldsymbol{q}}^\dagger]f_k(f_q^*)' + [\hat{a}_{\boldsymbol{k}}^\dagger,\hat{a}_{\boldsymbol{q}}]f_k^*f_q'\Big]\, e^{i(\boldsymbol{k}\cdot\boldsymbol{x}+\boldsymbol{q}\cdot\boldsymbol{y})} \\
&= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\int \frac{d^3\boldsymbol{q}}{(2\pi)^3}\Big[[\hat{a}_{\boldsymbol{k}},\hat{a}_{\boldsymbol{q}}^\dagger]\left(f_k(f_q^*)' - f_k^*f_q'\right)\Big]\, e^{i(\boldsymbol{k}\cdot\boldsymbol{x}+\boldsymbol{q}\cdot\boldsymbol{y})} \\
&= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\int \frac{d^3\boldsymbol{q}}{(2\pi)^3}(2\pi)^3\delta^{(3)}(\boldsymbol{k}+\boldsymbol{q})\left(f_k(f_q^*)' - f_k^*f_q'\right)\, e^{i(\boldsymbol{k}\cdot\boldsymbol{x}+\boldsymbol{q}\cdot\boldsymbol{y})} \\
&= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\Big[f_k(f_{-k}^*)' - f_k^*f_{-k}'\Big]\, e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{y})}.
\end{aligned} \tag{5.3.16}$$

To obtain the second line, we used $[\hat{a}_{\boldsymbol{k}}^\dagger,\hat{a}_{\boldsymbol{q}}] = -[\hat{a}_{\boldsymbol{q}},\hat{a}_{\boldsymbol{k}}^\dagger] = -[\hat{a}_{\boldsymbol{k}},\hat{a}_{\boldsymbol{q}}^\dagger]$. In the third line, the commutation relation is substituted for $[\hat{a}_{\boldsymbol{k}},\hat{a}_{\boldsymbol{q}}^\dagger]$. In order to get to the final result, integration is performed over the $\boldsymbol{q}$-momentum to obtain $\boldsymbol{k} = -\boldsymbol{q}$ on account of the Dirac delta function. For the commutator to satisfy the canonical commutation relation:

$$[\hat{f}(\tau,\boldsymbol{x}),\hat{\pi}(\tau,\boldsymbol{y})] = i\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}(-i)\Big[f_k(f_{-k}^*)' - f_k^*f_{-k}'\Big]\, e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{y})} \equiv i\delta^{(3)}(\boldsymbol{x}-\boldsymbol{y}), \tag{5.3.17}$$

we conclude that the expression:

$$W[f_k,f_k^*] = (-i)\left[f_k(f_{-k}^*)' - f_k^*f_{-k}'\right], \tag{5.3.18}$$

must equal unity: $W[f_k,f_k^*] \equiv 1$. The function $W[f_k,f_k^*]$ is called the Wroskian.

Correlation Function and Power Spectrum

Now that we have developed the quantum field theory of inflationary perturbations, we can predict the statistics of the inflaton quantum fluctuation field $\hat{f}$. In section 3.1, it was argued that the mean – i.e. the one-point correlation function $\langle 0|\hat{f}|0\rangle$ – of the quantum fluctuations vanishes due to their random nature. In contrast to the mean, the variance or two-point correlation function $\langle 0||\hat{f}|^2|0\rangle$ is non-zero. Here, these statements will be proven for single-field inflationary models and it will be shown that the power spectrum (Eq. 3.2.34) arises naturally from the computation of the two-point correlation function.

First, we consider the average $\langle \hat{f}\rangle$ in the vacuum state $|0\rangle$. Using the fact that $\hat{a}_{\boldsymbol{k}}$ annihilates the vacuum, it is straightforward to show that the average vanishes:

$$\boxed{\langle \hat{f}\rangle \equiv \langle 0|\hat{f}|0\rangle = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\langle 0|\left[f_k(\tau)\hat{a}_{\boldsymbol{k}} + f_k^*(\tau)\hat{a}^\dagger_{\boldsymbol{k}}\right]|0\rangle\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = 0,} \tag{5.3.19}$$

where the contractions show explicitly how the terms vanish: $\hat{a}^\dagger_{\boldsymbol{k}}$ annihilates $\langle 0|$ and $\hat{a}_{\boldsymbol{k}}$ has the same effect but on $|0\rangle$. Hence, we have shown that the 1-point correlation function vanishes.

Now we consider the 2-point correlation function:

$$\langle |\hat{f}|^2\rangle \equiv \langle 0|\hat{f}^\dagger(\tau,\boldsymbol{x})\hat{f}(\tau,\boldsymbol{x})|0\rangle. \tag{5.3.20}$$

In terms of creation and annihilation operators, the 2-point correlation function can be written as:

$$\begin{aligned}\langle |\hat{f}|^2\rangle &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\int \frac{d^3\boldsymbol{q}}{(2\pi)^3}\langle 0|(f_k^*\hat{a}^\dagger_{\boldsymbol{k}} + f_k\hat{a}_{\boldsymbol{k}})(f_q\hat{a}_{\boldsymbol{q}} + f_q^*\hat{a}^\dagger_{\boldsymbol{q}})|0\rangle \\ &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\int \frac{d^3\boldsymbol{q}}{(2\pi)^3}\langle 0|f_k f_q^*\hat{a}_{\boldsymbol{k}}\hat{a}^\dagger_{\boldsymbol{q}}|0\rangle \\ &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\int \frac{d^3\boldsymbol{q}}{(2\pi)^3} f_k f_q^*\langle 0|\hat{a}_{\boldsymbol{k}}\hat{a}^\dagger_{\boldsymbol{q}}|0\rangle,\end{aligned} \tag{5.3.21}$$

where the contractions show which terms vanish. To proceed, the commutation relation for the creation and annihilation operators sqeezed between two vacuum states will be used to rewrite the expectation value $\langle 0|\hat{a}_{\boldsymbol{k}}\hat{a}^\dagger_{\boldsymbol{q}}|0\rangle$:

$$\langle 0|[\hat{a}_{\boldsymbol{k}},\hat{a}^\dagger_{\boldsymbol{q}}]|0\rangle = \langle 0|\hat{a}_{\boldsymbol{k}}\hat{a}^\dagger_{\boldsymbol{q}}|0\rangle - \langle 0|\hat{a}^\dagger_{\boldsymbol{q}}\hat{a}_{\boldsymbol{k}}|0\rangle = \langle 0|\hat{a}_{\boldsymbol{k}}\hat{a}^\dagger_{\boldsymbol{q}}|0\rangle. \tag{5.3.22}$$

Notice that the second term after the first equality sign vanishes due to the contractions indicated. On account of this result, the 2-point correlation function can be written as follows:

$$\begin{aligned}\langle |\hat{f}|^2\rangle &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\int \frac{d^3\boldsymbol{q}}{(2\pi)^3} f_k f_q^*\langle 0|[\hat{a}_{\boldsymbol{k}},\hat{a}^\dagger_{\boldsymbol{q}}]|0\rangle \\ &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\int \frac{d^3\boldsymbol{q}}{(2\pi)^3} f_k f_q^* \times (2\pi)^3\delta^{(3)}(\boldsymbol{k}+\boldsymbol{q}) \\ &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}|f_k|^2 = \int d\ln k\, \mathcal{P}_f.\end{aligned} \tag{5.3.23}$$

In going from the first to the second line, we used the expression for the commutator (Eq. 5.3.9) in terms of the 3-momentum delta function. In going to the third, the delta function

$\delta^{(3)}(\boldsymbol{k}+\boldsymbol{q})$ is used to evaluate the $\boldsymbol{q}$-integral. From the last line it follows that the 2-point correlation function – or variance – of the inflaton fluctuations is proportional to the square of the Fourier mode function. In the last equality, the power spectrum for f is defined as:

$$\boxed{\mathcal{P}_f \equiv \frac{k^3}{2\pi^2}|f_k|^2,} \tag{5.3.24}$$

in accordance with result presented in Eq. 3.2.34.

5.4 Power Spectrum for Single Field Slow Roll Inflation

Recall that the aim of this chapter is to compute the power spectrum of quantum fluctuations generated during inflation at horizon exit. In the previous section, we found that the power spectrum is proportional to the square of the classical solution for the mode function $f_k(\tau)$ (see Eq. 5.3.24). From that perspective, the approach is quite straightforward: we first find the classical solution for the mode function, evaluate it at horizon exit and obtain the power spectrum. However, the evolution of the mode function is governed by the Mukhanov-Sasaki equation, which is difficult to solve in full generality as it depends on the background dynamics of for instance the inflaton and scale factor. To nevertheless obtain a solution to the mode function, we expand the MS Equation to first order in Hubble flow parameters. For this form of the MS equation, an algebraic solution for the mode function exists.

Subsequently, the solution to the mode function should be supplied with appropriate boundary conditions in order to completely fix the dynamics of the mode function. As we will see, finding the appropriate initial condition for the dynamics for $f_k(\tau)$ is closely related to constructing the vacuum state $|0\rangle$ of the field f. However, in a time-dependent background – such as the quasi de Sitter background for inflation – constructing the vacuum state is not a trivial task. Fortunately, for inflation we will find that there exists a preferred way to construct the vacuum. At early times ($|k\tau| \gg 1$) all modes of cosmological interest are far inside the horizon. In this limit, the MS equation for f_k reduces to the equation of motion for a simple harmonic oscillator in Minkowski space. For this system, the vacuum can be constructed unambiguously. Constructing this vacuum comes with a condition on the mode function in the early time limit. This requirement then forms the initial condition needed to completely fix the solution to mode function.

Expansion in Hubble Flow Parameters

The starting point will be the Mukhanov-Sasaki equation in Fourier space as derived in the first section of this chapter:

$$\boxed{f_k'' + \left(k^2 - \frac{z''}{z}\right) f_k = 0,} \tag{5.4.1}$$

where $z \equiv \dot{\phi}/aH = \sqrt{2\varepsilon}a$. Expanding this equation into the Hubble flow parameters comes down to writing the term z''/z in these variables. We define the Hubble flow parameters as:

$$\varepsilon \equiv -\frac{\dot{H}}{H^2}, \qquad \eta \equiv \frac{\dot{\varepsilon}}{\varepsilon H}, \qquad \kappa \equiv \frac{\dot{\eta}}{\eta H}, \tag{5.4.2}$$

and assume them all to be much smaller than unity, so that we can make a perturbative expansion to first order of z''/z in terms of them.

First, z'/z and z''/z can be expressed in terms of ε, η and κ as:

$$\frac{z'}{z} = \mathcal{H}\left[1 + \frac{1}{2}\eta\right], \quad \frac{z''}{z} = \mathcal{H}^2\left[1 - \varepsilon + \frac{3}{2}\eta - \frac{1}{2}\varepsilon\eta + \frac{1}{4}\eta^2 + \eta\kappa\right], \tag{5.4.3}$$

where $\mathcal{H} = aH$ is the Hubble parameter in conformal time. Notice that the above two expansions are still exact. In addition, we can write $\mathcal{H}$ explicitly is terms of conformal time by integrating the definition of the first Hubble flow parameter:

$$\partial_\tau \mathcal{H}^{-1} = \varepsilon - 1. \tag{5.4.4}$$

Assuming no time dependence in ε, i.e. $\varepsilon \neq \varepsilon(\tau)$, the following expression can be obtained for $\mathcal{H}$:

$$\boxed{\mathcal{H} = -\frac{1}{\tau}(1+\varepsilon),} \tag{5.4.5}$$

squaring this equation to first order in ε yields $\mathcal{H}^2 = (1+2\varepsilon)/\tau^2$.

Writing the expansion z''/z $(\varepsilon, \eta, \kappa)$ to first order in the flow parameters and substituting in the expression for $\mathcal{H}^2$ gives:

$$\frac{z''}{z} = \frac{1}{\tau^2}(1+2\varepsilon)\left[2 - \varepsilon + \frac{3}{2}\eta\right] = \frac{1}{\tau^2}\left[2 + 3\varepsilon + \frac{3}{2}\eta\right]. \tag{5.4.6}$$

For later convenience, all dependence on the flow parameters in the above expression will be absorbed into the variable $\nu \equiv 3/2 + \varepsilon + \eta/2$, in terms of which z''/z becomes:

$$\boxed{\frac{z''}{z} \equiv \frac{1}{\tau^2}\left[\nu^2 - \frac{1}{4}\right].} \tag{5.4.7}$$

Mukhanov-Sasaki Equation in Hubble Parameters

Using this form for z''/z, the MS equation can be rewritten as follows:

$$f_k'' + \omega_k^2(\tau) f_k = 0, \quad \text{where: } \omega_k^2(\tau) \equiv k^2 - \frac{\nu^2 - 1/4}{\tau^2}. \tag{5.4.8}$$

This equation of motion for $f_k(\tau)$ has an exact solution in terms of Hankel functions of the first and second kind [60]:

$$\boxed{f_k(\tau) = \sqrt{-\tau}\left[\alpha H_\nu^{(1)}(-k\tau) + \beta H_\nu^{(2)}(-k\tau)\right],} \tag{5.4.9}$$

where $H_\nu^{(1,2)}$ are the Hankel functions of the first and second kind. Hankel functions are two linearly independent combinations of the Bessel functions of the first and second kind [1]:

$$H_\nu^{(1)}(x) \equiv J_\nu(x) + iY_\nu(x), \qquad H_\nu^{(2)}(x) \equiv J_\nu(x) - iY_\nu(x), \tag{5.4.10}$$

where $J_\nu(x)$ and $Y_\nu(x)$ are the Bessel functions of the first and second kind, respectively. The constants α and β are to be determined by the initial conditions of the system. As we will show now, finding the appropriate initial condition for $f_k(\tau)$ is closely related to constructing the vacuum state of the system.

Bunch-Davies Vacuum and Initial Condition

In a time-*in*dependent background, fixing the vacuum state is rather straightforward. One simply imposes those boundary conditions on the mode functions that minimize the expectation value of the Hamiltonian in the quantum vacuum state $|0\rangle$. However, for a time-dependent background – such as the quasi de Sitter space-time in the case of inflation – the vacuum state is ill-defined, as the vacuum state expectation value of the Hamiltonian $\langle 0|\hat{H}|0\rangle$ changes with time because mode functions involve time-dependent frequencies $\omega_k(\tau)$. Hence, the minimum-energy vacuum state depends on the time at which it is defined. In other words, constructing the vacuum state $|0\rangle(\tau_0)$ at time τ_0 does not guarantee that $|0\rangle(\tau_0)$ is still the lowest energy state at any later time τ_2, that is $|0\rangle(\tau_1) \neq |0\rangle(\tau_2)$.

Nevertheless, for inflation there exists a preferred choice for the vacuum. At sufficiently early times $k/\mathcal{H} \sim |k\tau| \gg 1$, i.e. in the limit of large negative conformal time, all modes of cosmological interest are deep inside the horizon. In this limit, the frequencies $\omega_k(\tau)$ of the k-modes become time-independent:

$$\omega_k^2(k,\tau) = k^2\left[1 - \frac{\nu^2 - 1/4}{(k\tau)^2}\right] \quad \xrightarrow{|k\tau| \gg 1} \quad k^2. \tag{5.4.11}$$

Therefore, the MS equation reduces to the equation of motion for a simple harmonic oscillator in a time-independent Minkowski background:

$$f_k'' + k^2 f_k = 0. \tag{5.4.12}$$

For this system, the vacuum state can be constructed unambiguously. Constructing this vacuum state, which is called the *Bunch-Davies* vacuum, will give rise to a boundary condition for the mode functions in the early time limit. This boundary condition can then be used to determine the constants α and β and in this way, the solution for the mode function can be fixed completely.

Constructing the Vacuum

In order to construct the vacuum state, we first promote the Hamiltonian of the system to an operator $\hat{H}$ and compute the vacuum-state expectation value to find the lowest energy state:

$$E_0 \equiv \langle 0|\hat{H}|0\rangle. \tag{5.4.13}$$

The vacuum state $|0\rangle$ is defined via the following conditions on the creation and annihilation operators:

$$\hat{a}_{\boldsymbol{k}}\,|0\rangle = 0, \qquad \langle 0|\,\hat{a}_{\boldsymbol{k}}^\dagger = 0. \tag{5.4.14}$$

In the early time limit, $z''/z \to 0$ because $\tau \to -\infty$ and the effective potential $V(\hat{f})$ vanishes, so that the Hamiltonian operator can be written as:

$$\hat{H} = \int d^3x\,\hat{\mathcal{H}} = \frac{1}{2}\int d^3x\left[\hat{\pi}^2 + (\partial\hat{f})^2\right]. \tag{5.4.15}$$

Substituting the definitions of the quantized fields $\hat{f}$ and $\hat{\pi}$ in terms of the creation and annihilation operators in the Hamiltonian operator $\hat{H}$ yields:

$$\boxed{\hat{H} = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\left[\hat{a}_{\boldsymbol{k}}\hat{a}_{-\boldsymbol{k}}F_k + \hat{a}_{\boldsymbol{k}}^\dagger\hat{a}_{-\boldsymbol{k}}^\dagger F_k^* + \left(2\hat{a}_{\boldsymbol{k}}^\dagger\hat{a}_{\boldsymbol{k}} + (2\pi)^3\,\delta^{(3)}(0)\times E_k\right)\right],} \tag{5.4.16}$$

here the quantities E_k and F_k are defined as $E_k \equiv |f_k'|^2 + k^2|f_k|^2$ and $F_k \equiv f'^2_k + k^2 f_k^2$.

The vacuum expectation value then becomes:

$$\begin{aligned} E_0 \equiv \langle 0|\hat{H}|0\rangle &= \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\langle 0| \left[\hat{a}_{\boldsymbol{k}}\hat{a}_{-\boldsymbol{k}}F_k + \hat{a}^\dagger_{\boldsymbol{k}}\hat{a}^\dagger_{-\boldsymbol{k}}F_k^* + \left(2\hat{a}^\dagger_{\boldsymbol{k}}\hat{a}_{\boldsymbol{k}} + (2\pi)^3\,\delta^{(3)}(0)\right)\right]|0\rangle \\ &= \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\langle 0| \left[(2\pi)^3\delta^{(3)}(0)\times E_k\right]|0\rangle = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}E_k \times (2\pi)^3\delta^{(3)}(0). \end{aligned} \tag{5.4.17}$$

To go from the first to the second line we used the fact that acting with creation and annihilation operators on $\langle 0|$ and $|0\rangle$, respectively, yields zero. Therefore, all terms but the one containing the delta function vanish. Furthermore, we used normalization condition on the vacuum states: $\langle 0|0\rangle \equiv 1$. The integral expression for the vacuum state expectation value of the Hamiltonian is given by:

$$\boxed{E_0 = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\left(|f_k'|^2 + k^2|f_k|^2\right)\times(2\pi)^3\delta^{(3)}(0).} \tag{5.4.18}$$

Straightforward integration of the above equation for E_0 would yield infinity for two reasons. First, the presence of the delta function factor $(2\pi)^3\delta^{(3)}(0)$ makes the integral divergent. However, the delta function has an intuitive origin in this case: it arises because we integrate over all of space, which yields an infinite volume. This becomes apparent when we consider the system inside a box with lengths L and let the lengths tend to infinity. The delta function can then be written as:

$$(2\pi)^3\delta^{(3)}(0) = \lim_{L\to\infty}\int_{-L/2}^{L/2} d^3x\; e^{i\boldsymbol{x}\cdot\boldsymbol{p}}\Big|_{\boldsymbol{p}\,=\,0} = \lim_{L\to\infty}\int_{-L/2}^{L/2} d^3x \equiv \lim_{L\to\infty} V \longrightarrow \infty. \tag{5.4.19}$$

In other words, the delta function arises because now the total vacuum-state energy is computed, instead of the energy density. Hence, from now on we will consider the vacuum-state energy density ρ_0 rather than the total energy by factoring out the delta function.

Although perhaps less apparent, even by factoring out the delta function, the momentum integral for the energy density will still diverge. The reason is that the integral is not supplied with a high momentum cutoff. In other words, the theory is assumed to be valid to arbitrarily high momentum (or small length scales). However, the quantum field theory of inflation should be considered as an effective theory, valid up to a certain momentum limit, as quantified by the maximum $|k|$-value, denoted as Λ_*. Taking the above considerations into account, the vacuum energy density can be written as:

$$\boxed{\rho_0 = \frac{1}{2}\int^{\Lambda_*} \frac{d^3k}{(2\pi)^3}\left(|f_k'|^2 + k^2|f_k|^2\right).} \tag{5.4.20}$$

Minimizing the Vacuum State Energy

Now, we will find the mode function corresponding to the minimized the energy density ρ_0. For convenience, the (complex) mode function will be written in the following exponential form: $f_k \equiv r_k e^{i\alpha_k}$, where r_k and α_k are real. Using the Wroskian normalization condition on the mode functions,

$$W[f_k, f_k^*] = (-i)\left[f_k(f_k^*)' - f_k^* f_k'\right] = 1, \tag{5.4.21}$$

yields the constraint $\alpha_k' r_k^2 = -1/2$. The quantity E_k becomes:

$$E_k = r_k'^{\,2} + r_k^2 \alpha_k'^{\,2} + k^2 r_k^2, \tag{5.4.22}$$

which is minimized by choosing $r_k' = 0$ and $r_k = 1/\sqrt{2k}$. In addition, we need an explicit form for the phase factor α_k. By using the obtained form for r_k, the constraint $\alpha_k' r_k^2 = -1/2$ becomes $\alpha_k' = -k$. Integration of this equation gives $\alpha_k = -k\tau$.[8]

Now the vacuum state mode function can be written explicitly in terms of the mode k and conformal time τ as we found the expressions for r_k and α_k. However, notice that the explicit form is derived in the early time limit $|k\tau| \gg 1$. Hence, the obtained expression for f_k is merely an initial condition for the generic mode function as given by Eq. 5.4.9 and can be written as the following early time limit:

$$\boxed{\lim_{|k\tau|\gg 1} f_k(\tau) = \frac{1}{\sqrt{2k}} e^{-ik\tau}.} \tag{5.4.23}$$

Solution to the Mode Function

Now the above initial conditions can be used to determine the constants α and β in Eq. 5.4.9. Consider the $|k\tau| \gg 1$ limit of the Hankel functions $H_\nu^{(1,2)}(-k\tau)$:

$$\lim_{|k\tau|\gg 1} H_\nu^{(1)} = \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{-k\tau}} e^{-ik\tau} \times \Delta^{(-)}, \qquad \lim_{|k\tau|\gg 1} H_\nu^{(2)} = \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{-k\tau}} e^{+ik\tau} \times \Delta^{(+)}. \tag{5.4.24}$$

Here, we defined $\Delta^{(\pm)} \equiv \exp(\pm i\pi(\nu + 1/2)/2$ and $\Delta^{(+)}\Delta^{(-)} = 1$. In this limit the mode function can then be written as:

$$\lim_{|k\tau|\gg 1} f_k(\tau) = \sqrt{-\tau} \left[\alpha \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{-k\tau}} e^{-ik\tau} \times \Delta^{(-)} + \beta \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{-k\tau}} e^{+ik\tau} \times \Delta^{(+)} \right]. \tag{5.4.25}$$

In order to bring the above equation in agreement with the initial condition as given by Eq. 5.4.23, the constants are set to:

$$\alpha \equiv \frac{\sqrt{\pi}}{2} \times \Delta^{(+)}, \qquad \beta \equiv 0. \tag{5.4.26}$$

Therefore, the appropriate mode function f_k in line with the initial condition reads:

$$\boxed{f_k(\tau) = \sqrt{-\tau} \left[\frac{\sqrt{\pi}}{2} \Delta^{(+)} \right] H_\nu^{(1)}(-k\tau).} \tag{5.4.27}$$

The final form for the mode function as given by the above equation is plotted in Fig. 5.1 (in the slightly rescaled form $\tilde{f}_k = k^{1/2} f_k$). Notice that $|\tilde{f}_k|$ is constant or growing for all negative values of $k\tau$. This implies that, irrespective of the considered k-mode, the quantum fluctuation in the inflaton $\delta\phi_k \equiv a f_k$ grows as time evolves until $k\tau = 0$. During inflation, $a \simeq e^{\mathcal{H}\tau}$ and hence the quantum fluctuation grows exponentially fast. Therefore, as advocated, quantum fluctuations are indeed stretched to cosmological scales very rapidly.

[8] We set the integration constant to zero. This can be done safely as observables, such as the power spectrum, do only depend on the absolute square of the mode function, in which the complex exponentials are not present anymore.

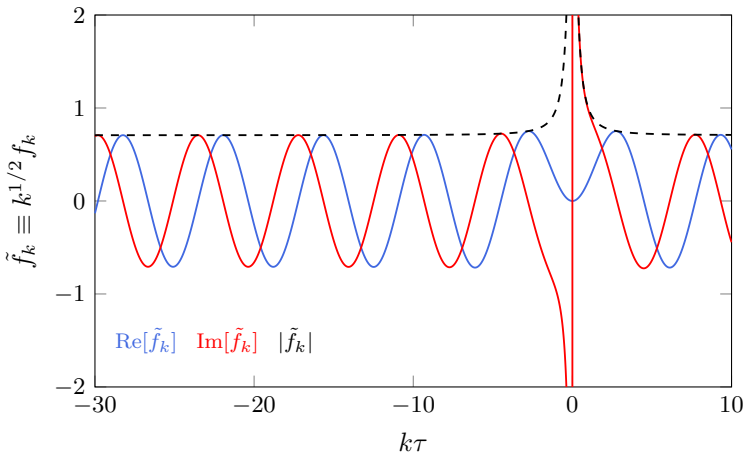


Figure 5.1: Plot of the final solution for the mode function as given by Eq. 5.4.27. The real, imaginary and absolute parts are shows separately in different colors. The unimportant phase factor $\Delta^{(+)}$ is dropped for all plotted functions.

Power Spectrum

Now that the mode function is known, the power spectrum can be computed. We will compute $\mathcal{P}_\mathcal{R}$ at horizon exit, that is for momenta $k = aH$. Then, since the curvature perturbation $\mathcal{R}$ is constant outside the horizon (see next chapter), the power spectrum computed at horizon exit can be taken directly to horizon re-entry. Horizon exit corresponds to the limit $|k\tau| \to 0$. In this limit the first Hankel function becomes:

$$\lim_{|k\tau|\to 0} H_\nu^{(1)}(-k\tau) = \sqrt{\frac{2}{\pi}} e^{-i\pi/2} 2^{\nu-3/2} \frac{\Gamma(\nu)}{\Gamma(3/2)} (-k\tau)^{-\nu}. \tag{5.4.28}$$

Using this result, the mode function reads:

$$\lim_{|k\tau|\to 0} f_k(\tau) = 2^{\nu-3/2} \Delta^{(+)} e^{-i\pi/2} \frac{\Gamma(\nu)}{\Gamma(3/2)} \frac{1}{\sqrt{2k}} (-k\tau)^{1/2-\nu} \tag{5.4.29}$$

Taking the absolute value of the above limit and using the (zeroth order) relation between the comoving Hubble radius and conformal time during inflation ($\tau = -1/aH$) yields:

$$\lim_{|k\tau|\to 0} |f_k(\tau)| = \frac{C(\nu)}{\sqrt{2k}} \left(\frac{k}{aH}\right)^{1/2-\nu}, \qquad C(\nu) \equiv 2^{\nu-3/2} \frac{\Gamma(\nu)}{\Gamma(3/2)}. \tag{5.4.30}$$

Notice that in the limit $\nu \to 3/2$, the constant $C(\nu)$ tends to unity.

Now, we can finally compute the power spectrum for $\mathcal{R}$, which is related to f via the relation $\mathcal{R} = f/z$ where $z \equiv a\dot{\phi}/H$ (see Eq. 5.1.34). In the following, we assume $\nu \to 3/2$ so $C(\nu)$ can be set to unity. By the definition of the power spectrum (Eq. 3.2.34):

$$\begin{aligned} \mathcal{P}_\mathcal{R} &= \frac{k^3}{2\pi^2} |\mathcal{R}_k|^2 = \frac{k^3}{2\pi^2} \left|\frac{f_k}{z}\right|^2 \\ &= \frac{k^3}{2\pi^2 z^2} \frac{1}{2k} \left(\frac{k}{aH}\right)^{3-2\nu} = \frac{k^2}{a^2\dot{\phi}^2} \left(\frac{H}{2\pi}\right)^2 \left(\frac{k}{aH}\right)^{1-2\nu}. \end{aligned} \tag{5.4.31}$$

By definition of the parameter ε, the time derivative of the inflaton field is given by $\dot{\phi}^2 = 2\varepsilon M_{\rm pl}^2 H^2$ and the power spectrum can be written as:

$$\boxed{\mathcal{P}_{\mathcal{R}} = \frac{1}{2\varepsilon M_{\rm pl}^2}\left(\frac{H}{2\pi}\right)^2\left(\frac{k}{aH}\right)^{3-2\nu} \equiv A_{\rm S}^2\left(\frac{k}{aH}\right)^{n_{\rm s}-1}.} \tag{5.4.32}$$

The spectral index $n_{\rm s}$, governing the scale dependence of the power spectrum, is defined as:

$$n_{\rm s} - 1 \equiv \frac{d\ln\mathcal{P}_{\mathcal{R}}}{d\ln k} = 3 - 2\nu = -2\varepsilon - \eta = 2\eta_{\rm V} - 6\varepsilon_{\rm V}. \tag{5.4.33}$$

Here, we used the definition of ν and the relations between the Hubble flow parameters (ε, η) and the potential slow roll parameters $(\varepsilon_{\rm V}, \eta_{\rm V})$, which are given by $\varepsilon = \varepsilon_{\rm V}$ and $\eta_{\rm V} = 2\varepsilon - \eta/2$.

5.5 Quantum to Classical Transition

As modes exit the horizon, they lose their quantum nature and can be regarded as classical quantities. To be more specific, on super-horizon scales the quantum fluctuation field $f \equiv a\delta\phi$ can be regarded as a classical stochastic field: this transition is known as the quantum-to-classical transition. In this section, we will discuss this transition in a quantitative way.

The signature of classical rather then quantum modes are commuting quantum operators, i.e. the commutators of the relevant quantum operators vanish. For the quantum field theory of inflaton fluctuations, we constructed the operators $\hat{f}$ and $\hat{\pi}$ as follows:

$$\hat{f}(\tau, \boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\left[f_k(\tau)\hat{a}_{\boldsymbol{k}} + f_k^*(\tau)\hat{a}_{\boldsymbol{k}}^\dagger\right]e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{5.5.1}$$

$$\hat{\pi}(\tau, \boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\left[f_k'(\tau)\hat{a}_{\boldsymbol{k}} + (f_k^*)'(\tau)\hat{a}_{\boldsymbol{k}}^\dagger\right]e^{i\boldsymbol{k}\cdot\boldsymbol{x}}. \tag{5.5.2}$$

In the limit $|k\tau| \to 0$, corresponding to horizon exit, the mode function and its first time derivative become:

$$f_k = -\frac{1}{\sqrt{2k^3}}\frac{i}{\tau}, \qquad f_k' = \frac{1}{\sqrt{2k^3}}\frac{i}{\tau^2}, \tag{5.5.3}$$

where the value of ν is set to $3/2$. The operators can then be written as:

$$\hat{f}(\tau, \boldsymbol{x}) = -\frac{i}{\tau}\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\frac{1}{\sqrt{2k^3}}\left[\hat{a}_{\boldsymbol{k}} - \hat{a}_{\boldsymbol{k}}^\dagger\right]e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{5.5.4}$$

$$\hat{\pi}(\tau, \boldsymbol{x}) = \frac{i}{\tau^2}\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\frac{1}{\sqrt{2k^3}}\left[\hat{a}_{\boldsymbol{k}} - \hat{a}_{\boldsymbol{k}}^\dagger\right]e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = -\frac{1}{\tau}\hat{f}(\tau, \boldsymbol{x}). \tag{5.5.5}$$

Notice that after horizon exit, the two operators are proportional to each other and hence commute on super-horizon scales:

$$\boxed{\left[\hat{f}(\tau, \boldsymbol{x}), \hat{\pi}(\tau, \boldsymbol{x})\right] \xrightarrow{|k\tau|\to 0} 0.} \tag{5.5.6}$$

The vanishing commutator implies that on super-horizon scales the field $\hat{f}(\tau, \boldsymbol{x})$ loses its quantum nature and can be identified with a classical field.

5.6 Gravitational Waves from Single-Field Inflation

In addition to the scalar metric perturbations induced by single-field inflation, tensor perturbation in the metric are also generated, which are referred to as primordial gravitational waves [23, 48, 61, 87]. According to the SVT decomposition, those metric perturbations decouple from scalar (and vector) perturbations at linear order and can therefore be studied independently. The perturbed line element for (spatial) metric fluctuations can be written as:

$$ds^2 = a^2(\tau)\big[-d\tau^2 + (\delta_{ij} + 2\hat{E}_{ij})dx^i dx^j\big]. \tag{5.6.1}$$

The tensorial perturbation $\hat{E}_{ij}$ is symmetric, transverse ($\partial^i \hat{E}_{ij} = 0$) and traceless ($\hat{E}^i_i = 0$). The perturbed Christoffel symbols for this line element are given by (see also [56]):

$$\delta\Gamma^0_{ij} = 2\mathcal{H}\hat{E}_{ij} + \hat{E}'_{ij}, \tag{5.6.2}$$

$$\delta\Gamma^i_{0j} = \hat{E}'_{ij}, \tag{5.6.3}$$

$$\delta\Gamma^i_{jk} = \partial_k \hat{E}^i_j + \partial_j \hat{E}^i_k - \partial^i \hat{E}_{jk}. \tag{5.6.4}$$

For notational convenience, we will drop the hats on the tensor fluctuation from now on and write E_{ij} instead. The perturbed Ricci tensor can now be written in terms of the perturbed Christoffel symbols as follows:

$$\delta R_{ij} = 2(\mathcal{H}' - 2\mathcal{H}^2)E_{ij} + 2\mathcal{H}E'_{ij} + E''_{ij} - \partial^2 E_{ij}. \tag{5.6.5}$$

The mixed variant of the perturbed Ricci tensor, denoted as δR^i_j, equals the perturbed Einstein tensor since tensorial perturbations do not contribute to variations in the Ricci scalar, yielding [56]:

$$\delta G^i_j = \delta R^i_j = \frac{1}{a^2}\big(E^{i\prime\prime}_j + 2\mathcal{H}E^{i\prime}_j - \partial^2 E^i_j\big). \tag{5.6.6}$$

This result can be equated to the purely spatial part of the perturbed energy-momentum tensor for a perfect fluid (along with the dimensional factor $8\pi G$):

$$\delta T^i_j = \delta P \delta^i_j + \delta\Sigma^i_j, \tag{5.6.7}$$

see Eq. 6.4.43. However, the tensorial perturbation E^i_j is traceless and hence δG^i_j vanishes in case $i = j$. Therefore, we can solely consider the anistropic stress contribution $\delta\Sigma^i_j$, so that the equation of motion for E^i_j, corresponding to the purely spatial part of the perturbed EFE's, reads:

$$E^{i\prime\prime}_j + 2\mathcal{H}E^{i\prime}_j - \partial^2 E^i_j = 2\Delta(a)\delta\Sigma^i_j, \tag{5.6.8}$$

here we defined as before $\Delta(a) \equiv a^2/(2M^2_{\rm pl}) = 4\pi G a^2$. In momentum space, the equation of motion becomes:

$$\boxed{E^{i\prime\prime}_j(\boldsymbol{k}) + 2\mathcal{H}E^{i\prime}_j(\boldsymbol{k}) + k^2 E^i_j(\boldsymbol{k}) = 2\Delta(a)\delta\Sigma^i_j(\boldsymbol{k}),} \tag{5.6.9}$$

where we denote the momentum dependence of the Fourier mode by means of $E^i_j(\boldsymbol{k})$ instead of the usual notation $(E^i_j)_{\boldsymbol{k}}$, in order to avoid notational clutter.

In full generality, the above equation of motion for the tensor perturbation (written in momentum space) mimics that of a driven harmonic oscillator with driving term proportional to $\delta\Sigma^i_j(\boldsymbol{k})$ and friction term $2\mathcal{H}E^{i\prime}_j(\boldsymbol{k})$. For single-field inflation, however, we know that no

anisotropic stress contribution is generated and hence equation of motion reduces to that of a damped undriven oscillator:

$$E''_{ij} + 2\mathcal{H}E'_{ij} - \partial^2 E_{ij} = 0. \tag{5.6.10}$$

As the tensor perturbation to the metric satisfies a wave equation, it is evident that E^i_j follows a wave behavior and are hence referred to as gravitational waves. The above equation of motion can be obtained from the following action:

$$S_E = \frac{M_{\rm pl}^2}{2} \int d^3\boldsymbol{x}\, d\tau \left(E'_{ij} E^{ij\prime} - \partial_i E_{jk} \partial^i E^{jk} \right), \tag{5.6.11}$$

where the factor of $M_{\rm pl}^2/2$ is included to make E_{ij} manifestly dimensionless. Note that in the literature the prefactor $M_{\rm pl}^2/8$ is used frequently as well, corresponding to a perturbed spatial metric of the form $g_{ij} = a^2(\delta_{ij} + E_{ij})$, omitting the additional factor of two used in our definition.

Decomposing the gravitational wave in terms of its momentum modes yields:

$$E_{ij}(\tau, \boldsymbol{x}) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\, E_{ij}(\tau, \boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}. \tag{5.6.12}$$

In Fourier space, the symmetry, tracelessness and transversality constraints become:

$$E_{ij}(\tau, \boldsymbol{k}) = E_{ji}(\tau, \boldsymbol{k}), \qquad E_{ii}(\tau, \boldsymbol{k}) = 0, \qquad k^i E_{ij}(\tau, \boldsymbol{k}) = 0. \tag{5.6.13}$$

In general, the tensor E_{ij} constitutes nine independent components. However, due to the three constraints the number of independent components reduces to two, as there arise three constraints from both symmetry and tranversality in addition to the one coming from tracelessness.

The two remaining independent degrees of freedom correspond to the two (linear) polarization modes of the gravitational wave. On account of rotational invariance of the background, we can choose $\boldsymbol{k}$ to point in the z-direction and hence the non-vanishing components of the tensor $E_{ij}(\boldsymbol{k})$ become:

$$E_{(1)} \equiv E_{11} = -E_{22}, \qquad E_{(2)} \equiv E_{12} = E_{21}. \tag{5.6.14}$$

Using these amplitudes and the corresponding polarization tensors, the Fourier mode of the gravitational wave $E_{ij}(\boldsymbol{k})$ can be expanded as the sum of those two polarization modes:

$$E_{ij}(\boldsymbol{k}) = \sum_{(\gamma=1,2)} E_{(\gamma)}(\tau, \boldsymbol{k}) e^{(\gamma)}_{ij}(\hat{\boldsymbol{k}}), \tag{5.6.15}$$

where $E_{(\gamma)}$ and $e^{(\gamma)}_{ij}$ correspond to the amplitude and polarization tensor of the mode (γ), respectively. The polarization tensors for the modes can be written as:

$$e^{(1)}_{ij} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad e^{(2)}_{ij} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{5.6.16}$$

and are subject to the orthogonality and reality constraints:

$$e^{(\gamma)}_{ij} e^{(\gamma')*}_{ij} \equiv 2\delta_{(\gamma,\gamma')}, \qquad e^{(\gamma)}_{ij}(\hat{\boldsymbol{k}})^* = e^{(\gamma)}_{ij}(-\hat{\boldsymbol{k}}). \tag{5.6.17}$$

Often it is convenient to work with polarization tensors defined in the circular basis, instead of the linear polarization tensors given above. In terms of the linear polarization tensor, the circular ones are defined as:

$$e_{ij}^{(+,\times)} \equiv e_{ij}^{(1)} \pm i e_{ij}^{(2)}, \tag{5.6.18}$$

where the plus corresponds to the (+) (right-handed) polarization.

By making a field redefinition, the action S_E (Eq. 5.6.11) can be recasted in a form very similar to the Mukhanov-Sasaki action for scalar perturbations (Eq. 5.2.13). The required field redefinition is given by:

$$f_{ij} \equiv \frac{a M_{\rm pl}^2}{\sqrt{2}} E_{ij}, \tag{5.6.19}$$

and induces the action to change to:

$$S_E = \frac{M_{\rm pl}^2}{2} \int d\tau \, d^3\boldsymbol{x} \Big[f'_{ij} f^{ij\prime} - \partial_i f_{jk} \partial^i f^{jk} + \frac{a''}{a} f_{ij} f^{ij} \Big]. \tag{5.6.20}$$

The corresponding equation of motion in Fourier space then reads:

$$f''_{ij}(\boldsymbol{k}) + \Big[k^2 - \frac{a''}{a} \Big] f_{ij}(\boldsymbol{k}) = 0. \tag{5.6.21}$$

Note that the polarization tensors posses no time-dependence and hence they can be factored out to give two equation of motions identical to the MS equation (Eq. 5.2.14), one for each of the two polarization modes:

$$f''_{(\gamma)}(\boldsymbol{k}) + \Big[k^2 - \frac{a''}{a} \Big] f_{(\gamma)}(\boldsymbol{k}) = 0. \tag{5.6.22}$$

Notice that this equation is equivalent to the massless limit ($m_\chi \to 0$) of the equation of motion for the Toy model field σ (Eq. 3.5.11). In other words, the amplitude of the gravitational wave modes behaves as fluctuations in a massless scalar field. Hence, for each polarization mode we can take over the results from section 3.5, with the understanding that the parameter $\eta_\chi \propto m_\chi$ vanishes. To be more specific, let us define the parameter $\mu \equiv 3/2 + \varepsilon$ which is equivalent to the parameter ν_χ defined in section 3.5 for the case $\eta_\chi = 0$. The term a''/a in the above equation of motion can then be expanded in terms μ as follows:

$$\frac{a''}{a} = \mathcal{H}^2 (2 - \varepsilon) = \frac{1}{\tau^2} (\mu^2 - 1/4). \tag{5.6.23}$$

The equation of motion than becomes:

$$\boxed{f''_{(\gamma)}(\boldsymbol{k}) + \Big[k^2 - \frac{\mu^2 - 1/4}{\tau^2} \Big] f_{(\gamma)}(\boldsymbol{k}) = 0,} \tag{5.6.24}$$

which is equivalent to Eq. 3.5.18 under the replacement $\nu_\chi \to \mu$.

Therefore, we can take over the results of the super-horizon evolution of the field modes as given in Eq. 3.5.49 under the replacement, yielding:

$$\lim_{|k\tau| \to 0} |f_{(\gamma)}(k)| = \frac{C(\mu)}{\sqrt{2k}} \left(\frac{k}{aH} \right)^{1/2 - \mu}, \qquad C(\mu) \equiv 2^{\mu - 3/2} \frac{\Gamma(\mu)}{\Gamma(3/2)} \to 1, \tag{5.6.25}$$

where we used the fact that $C(\mu)$ approaches unity in the limit $\mu \to 3/2$. The power spectrum for gravitational waves can now be evaluated after horizon crossing. By definition, the power spectrum is given by:

$$\mathcal{P}_E = \frac{k^3}{2\pi^2} E^*_{ij} E^{ij} = \frac{k^3}{2\pi^2} \sum_{(\gamma,\gamma')} E_{(\gamma)} E_{(\gamma')} e^{(\gamma)}_{ij} e^{(\gamma')}_{ij} = \frac{k^3}{2\pi^2} \left(2|E_{(+)}|^2 + 2|E_{(\times)}|^2\right), \tag{5.6.26}$$

where we used the orthogonality between the two polarization tensors $e^{(\gamma)}_{ij} e^{(\gamma')*}_{ij} = 2\delta_{(\gamma,\gamma')}$. As both polarization amplitudes satisfy the same equation of motion, they can combined to yield:

$$\mathcal{P}_E = \frac{k^3}{2\pi^2}\, 4|E|^2, \tag{5.6.27}$$

i.e. we omit the polarization labels $(+,\times)$ on the amplitudes and write E instead. Via the transformation factor introduced in Eq. 5.6.19, the amplitude E can be written in terms of f, which allows us to use the results form section 3.5 to evaluate the power spectrum of gravitational waves. Transforming E into f and substituting the super-horizon result for the mode function f according to Eq. 5.6.25 yields:

$$\boxed{\mathcal{P}_E = \frac{k^3}{2\pi^2} \frac{8}{a^2 M_{\rm pl}^2} |f|^2 = \frac{8}{M_{\rm pl}^2} \left(\frac{H}{2\pi}\right)^2 \left(\frac{k}{aH}\right)^{3-2\mu} \equiv A_{\rm T}^2 \left(\frac{k}{aH}\right)^{n_{\rm t}}.} \tag{5.6.28}$$

In the last equality we defined the tensor amplitude $A_{\rm T}$ and the tensor spectral index $n_{\rm t}$ as follows:

$$n_{\rm t} \equiv \frac{d\ln \mathcal{P}_E}{d\ln k} = 3 - 2\mu = -2\varepsilon. \tag{5.6.29}$$

In the case of slow-roll, the tensor spectrum becomes scale invariant. Contrary to the scalar spectrum (Eq. 5.4.32), which depends on both H and ε, the tensor power spectrum solely depends on the Hubble parameter. Hence, in the slow-roll approximation, the Hubble parameter at horizon exit (i.e. the end of inflation) uniquely determines $\mathcal{P}_E$. Therefore, if the gravitational wave power spectrum could be measured empirically, it would allow us to estimate the energy scale at which inflation occurred in a model-independent sense,[9] since the Hubble parameter is directly related to the energy density via the Friedmann equation.

Since primordial gravitational waves are not observed yet, only provide an (upper) bound on the correlation between the amplitudes of the scalar and tensor spectra can be obtained observationally. The correlation is measured via the ratio r of the scalar and tensor amplitudes (i.e. the power spectra at the pivot scale, see Eq. 3.4.6):

$$r \equiv \frac{A_{\rm T}^2}{A_{\rm S}^2}, \tag{5.6.30}$$

Since we know that the scalar and tensor spectral amplitudes are given by:

$$A_{\rm S}^2 = \frac{1}{2\varepsilon M_{\rm pl}^2} \left(\frac{H}{2\pi}\right)^2, \qquad A_{\rm T}^2 = \frac{8}{M_{\rm pl}^2} \left(\frac{H}{2\pi}\right)^2, \tag{5.6.31}$$

we find that r can be written as follows:

$$r = 16\varepsilon \simeq 16\epsilon_{\rm V}, \tag{5.6.32}$$

where we used the slow-roll approximation in the last equality to replace ε by $\epsilon_{\rm V}$.

[9] Model-independent, in this case, means independent of the choice of inflaton potential $V(\phi)$.

Chapter 6

Evolution Outside the Horizon

"By looking far out into space we are also looking far back into time, back toward the horizon of the universe, back toward the epoch of the Big Bang."

— Carl Sagan

In order to relate inflationary predictions to late time observables such as the temperature anisotropies in the CMB, we have to show that the curvature perturbation $\mathcal{R}$ is constant on super-horizon scales. The condition of $\dot{\mathcal{R}} = 0$ on super-horizon scales is required because, as the modes of cosmological interest are outside the horizon, the universe is believed to go through the era of reheating. This era essentially forms the transition from the inflationary era to the conventional radiation and matter dominated eras, as described by conventional Big Bang cosmology. However, the physics during the stage of reheating is very uncertain, even the equations governing the evolution of the perturbations are not well-known. By encompassing the perturbations in a quantity – the comoving curvature perturbation $\mathcal{R}$ – that is constant on super-horizon scales, one completely avoids the issues concerned with lack of knowledge during the stage of reheating.

In this chapter, we will therefore focus in on the (non-)evolution of perturbations on super-horizon scales.[1] First, we will define more quantitatively what is meant by the super-horizon limit or a Fourier mode k being outside the horizon. Recall that during inflation, the comoving Hubble sphere ($(aH)^{-1}$) decreases, while the comoving wavelength k^{-1} of a specific Fourier mode is constant (Fig. 3.2). Hence, after a sufficient number of e-foldings, the comoving Hubble radius becomes substantially smaller than the mode wavelength:

$$\frac{k}{aH} \ll 1,$$

and the mode k is said to be outside the horizon. Typically, this condition can be replaced by the limit $k \to 0$, so that spatial derivative ∂_i and gradient terms ∂^2, which are proportional to k respectively k^2 in Fourier space, can be neglected. However, the condition $(k/aH) \ll 1$ is not equivalent to the limit $k \to 0$, and in some cases the replacement of the first by the second leads to difficulties [6]. However, for most inflationary models the limit $k \to 0$ can be made safely and this will be done in the rest of this chapter.

In this chapter, we will provide three different proofs presented in the literature for the constancy of the comoving curvature perturbation $\mathcal{R}$ outside the horizon. In this respect, this chapter summarizes some different possible approaches taken in the literature to show that

[1] For the gravitational potential Ψ, we already concluded that it is constant on outside the horizon in section 5.1.

$\mathcal{R}$ does not evolve on super-horizon scales. However, before we give these different proofs, we will derive the important statement that outside the horizon the curvature perturbation on slices of uniform energy density is equal the comoving curvature perturbation in section 6.1, that is:

$$\zeta = \mathcal{R}.$$

This result will be at the core of the different proofs as it allows us to show that $\dot{\zeta} = 0$, which is often more convenient, and then conclude that $\mathcal{R}$ is conserved as well based on the above relation. The first proof will be given in section 6.2 and is completely based on the perturbed field equations and therefore only applies to Einstein gravity. The second proof (section 6.3) relies on the conservation of energy and momentum, but is independent of the considered theory of gravitation and therefore more general. However, as we will show, both of these proofs rely on the assumption that perturbations are adiabatic (section 4.6). The third and final proof to be given, coming from Weinberg [90], does not rely on the assumption that perturbations are adiabatic, but instead *shows* that there are always[2] two adiabatic solutions for $\mathcal{R}$ outside horizon, of which one is non-zero and constant and the other is a decaying mode.[3]

6.1 Equality of ζ and $\mathcal{R}$ Outside the Horizon

Outside the horizon, i.e. in the limit $k \to 0$, one can show that the comoving curvature perturbation $\mathcal{R}$ coincides with the comoving curvature perturbation on slices of uniform energy density ζ. We will show two approaches to arrive at this result. The first applies to a scalar field (e.g. the inflaton) and solely employs its energy-momentum tensor, i.e. it is independent of the theory of gravitation. The second approach is based on the perturbed field equations for an arbitrary perfect fluid, but constrains to Einstein gravity, whereas the first approach does not.

6.1.1 Energy-Momentum Approach: Uniform Density Gauge

We will show here that $\mathcal{R}$ and the comoving curvature perturbation on slices of uniform energy density ζ are equal to each other in the uniform density gauge, by using solely the perturbed energy-momentum tensor for a scalar field. To show this, we invoke the results for the purely temporal and spatial components of the perturbed energy-momentum tensor (Eq. 4.8.31):

$$\delta T^0_0 = -\delta\rho = \Phi\phi'^2 - \delta\phi'\ \phi' - \delta\phi V_\phi a^2, \tag{6.1.1}$$

$$\delta T^i_j = \delta P \delta^i_j = \left(-\Phi\phi'^2 + \delta\phi'\ \phi' - \delta\phi V_\phi a^2\right)\delta^i_j. \tag{6.1.2}$$

Combining the above equations gives the following result for the difference between the density and pressure perturbations:

$$\delta\rho - \delta P = 2a^2\delta\phi V_\phi. \tag{6.1.3}$$

For single-field inflation the perturbations are adiabatic and hence the pressure and density perturbations are related via the sound speed c_s^2 as $\delta P_{(\mathrm{ad})} = c_s^2\delta\rho$. Hence, in going to the

[2] That is, irrespective of the constituents of the universe.

[3] It should be mentioned that the last proof is more precise but therefore also more convoluted and detailed than the preceding two.

uniform density gauge by setting $\delta\rho \equiv 0$, the pressure perturbation vanishes as well and the l.h.s. of the above equation is zero. Assuming $V_\phi \neq 0$, we find that $\delta\phi = 0$ as well so that:

$$\delta\rho = \delta\phi = 0, \tag{6.1.4}$$

in the uniform density gauge. By definition of ζ and $\mathcal{R}$:

$$\zeta = \Phi + \mathcal{H}\frac{\delta\rho}{\rho'}, \qquad \mathcal{R} = \Phi + \mathcal{H}\frac{\delta\phi}{\phi'}, \tag{6.1.5}$$

they are equal to each other for $\delta\rho = \delta\phi = 0$. Therefore, we have shown that ζ and $\mathcal{R}$ are equal to each other in the uniform density guage and hence the conservation equation (Eq. 6.3.11) for adiabatic perturbations holds for $\mathcal{R}$ as well.

6.1.2 Field Equations Approach: Newtonian Gauge

Now, we will show the equivalence of ζ and R outside the horizon for a generic fluid (i.e. possibly with non-anisotropic stress at first order) by means of the field equations in Newtonian gauge. As described in Appendix D.4, by construction ζ coincides with Ψ on slices of constant energy density ($\delta\rho \equiv 0$) and it is related to arbitrary gauge by:

$$\zeta \equiv \Psi_{\rm UD} = \Psi + \mathcal{H}\frac{\delta\rho}{\rho'} = \Psi - \frac{\delta\rho}{3(\rho + P)}, \tag{6.1.6}$$

where in the last equality we used the continuity equation. As the r.h.s. applies to arbitrary gauge, we can specify to the comoving gauge, in which case $\Psi_{\rm C} \equiv \mathcal{R}$ (Appendix D.4) and we obtain:

$$\zeta = \mathcal{R} - \frac{\delta\rho_{\rm C}}{3(\rho + P)}. \tag{6.1.7}$$

In order for ζ and $\mathcal{R}$ to coincide outside the horizon, we thus have to show that term proportional to $\delta\rho_{\rm C}$ vanishes in the super-horizon limit $k \to 0$. We will do this by using the perturbed field equations (Eqs. 4.9.16–4.9.19) for a non-perfect fluid at first order (i.e. we allow for anisotropic perturbations). However, Eqs. 4.9.16–4.9.19 are defined in Newtonian gauge and we require an expression for $\delta\rho$ in the comoving gauge. To find a relation between $\delta\rho_{\rm C}$ and variables in the Newtonian gauge, we use the scalar gauge transformation rule (Eq. 4.3.39) applied to $\delta\rho$, yielding:

$$\delta\rho_{\rm C} = \delta\rho_{\rm N} - \rho'\xi^0, \tag{6.1.8}$$

where the temporal shift is the one to go needed to go from the Newtonian gauge to the comoving orthogonal gauge, i.e. $\xi_0 = -v$.[4] Hence, using the continuity equation we obtain:

$$\delta\rho_{\rm C} = \delta\rho_{\rm N} - 3\mathcal{H}(\rho + P)v_{\rm N}. \tag{6.1.10}$$

Now we can re-express the above expression in terms the gravitational potential Ψ by using Eqs. 4.9.16 and 4.9.17:

$$\partial^2\Psi - 3\mathcal{H}(\Psi' + \mathcal{H}\Phi) = \Delta(a)\delta\rho, \tag{6.1.11}$$

$$-\partial_i(\Psi' + \mathcal{H}\Phi) = \Delta(a)(\rho + P)v_i. \tag{6.1.12}$$

[4]Recall that the temporal shift required to form arbitrary gauge to the comoving orthogonal gauge is:

$$\xi^0 = -(v + B), \tag{6.1.9}$$

on account of Eq. D.4.5. However, in Newtonian gauge $B = 0$ so we get $\xi^0 = -v$ as the shift to go to the comoving orthogonal gauge.

Multiplying the second by $3\mathcal{H}$ and adding it to the first gives the constraint equation:

$$\delta\rho_{\rm N} - 3\mathcal{H}(\rho + P)v_{\rm N} = \frac{\partial^2\Psi}{\Delta(a)} = \delta\rho_{\rm C}. \tag{6.1.13}$$

Substituting this result for $\delta\rho_{\rm C}$ into Eq. 6.1.7 gives:

$$\zeta = \mathcal{R} - \frac{\partial^2\Psi}{3\Delta(a)(\rho + P)}. \tag{6.1.14}$$

Outside the horizon ($k \to 0$), the gradient term $\partial^2\Psi$ vanishes as it is proportional to k^2 in Fourier space. Hence, we find that ζ and $\mathcal{R}$ indeed coincide outside the horizon.

6.2 Field Equations Approach

Here, we will show that the comoving curvature perturbation $\mathcal{R}$ is constant outside the horizon for adiabatic perturbations, based on the perturbed field equations in Newtonian gauge. By explicitly using the field equations, the validity of the proof is thus limited to Einstein gravity.[5] In the Newtonian gauge, $\mathcal{R}$ is defined as:

$$\mathcal{R} = \Psi - \mathcal{H}v. \tag{6.2.1}$$

Using the field equations for a generic fluid in Newtonian gauge (Eqs. 4.9.16–4.9.19), we can eliminate the velocity potential in favor of the gravitational potentials:

$$v = \frac{2}{3\mathcal{H}^2(1+w)}(\Psi + \mathcal{H}\Phi), \tag{6.2.2}$$

where the equation of state $w = P/\rho$ is evaluated at background level. Substituting this result in the expression for $\mathcal{R}$, and taking the conformal time derivative we obtain, after some algebraic manipulations:

$$\mathcal{R}' = \Psi' - \frac{2w'}{3(1+w)^2}(\mathcal{H}^{-1}\Psi' + \Phi) - \frac{2}{3(1+w)}\left(-\mathcal{H}^{-1}\Psi'' + \frac{\mathcal{H}'}{\mathcal{H}^2}\Psi' - \Phi'\right). \tag{6.2.3}$$

In Appendix E.1 it is shown that, using the field equations in Newtonian gauge, the above equation can be rewritten as:

$$\boxed{\mathcal{R}' = -\frac{2\mathcal{H}}{3(1+w)}\left(\frac{k}{\mathcal{H}}\right)^2\left[c_s^2\Psi + \frac{1}{2}(\Psi - \Phi)\right] - 3\mathcal{H}c_s^2\mathcal{S}.} \tag{6.2.4}$$

In the super-horizon limit, corresponding to $k/\mathcal{H} \to 0$, we find that the evolution of the comoving curvature perturbations is solely sourced by the isocurvature perturbation $\mathcal{S}$, that is:

$$\lim_{k/\mathcal{H}\to 0} \mathcal{R}' = -3\mathcal{H}c_s^2\mathcal{S}. \tag{6.2.5}$$

Therefore, for adiabatic perturbations, we find that the comoving curvature perturbation does not evolve outside the horizon, since we can set the isocurvature perturbation to zero: $\mathcal{S} = 0$. As single-field inflation produces adiabatic perturbations, we conclude that $\mathcal{R}$ is conserved in the single-field scenario.

[5] In the next section, however, we will derive the evolution of $\mathcal{R}$ outside the horizon from energy-momentum conservation (i.e. independent of the considered metric theory of gravity) and show that the result is the same.

6.3 Energy-Momentum Approach

In this section, we will prove that $\mathcal{R}$ does not evolve on super-horizon scales, based on energy-momentum conservation (i.e. independent of the considered theory of gravitation). In other words, the proof is independent of the equations governing the evolution of the perturbations, in contrast to the proof given in the previous section. The proof presented here will be performed at linear order in perturbations, and is partially based on [74].[6] Energy and momentum conservation can be expressed as the following condition on the energy-momentum tensor:

$$\nabla_\mu T^{\mu\nu} = 0. \tag{6.3.1}$$

For a perturbed energy-momentum tensor $T^\mu_\nu = \bar{T}^\mu_\nu + \delta T^\mu_\nu$, the above energy-momentum conservation equation for the perturbed part can be written as:

$$\nabla_\mu \delta T^\mu_\nu = 0. \tag{6.3.2}$$

Assuming a scalar perturbed perfect fluid quantified by an energy density ρ, pressure P, 3-velocity potential v and vanishing anisotropic stress tensor $\delta\Sigma = 0$, the perturbed components δT^μ_ν can be written as (Eq. 4.8.24):

$$\delta T^0_0 = -\delta\rho, \qquad \delta T^i_0 = -(\rho + P)v^i, \qquad \delta T^i_j = \delta P \delta^i_j. \tag{6.3.3}$$

On account of the Helmholtz theorem, the velocity can decomposed into scalar and vector components $v_i = \partial_i v + \hat{v}_i$. Considering only scalar perturbations, the δT^i_0 perturbation can be written as:

$$\delta T^i_0 = -(\rho + P)\partial_i v, \tag{6.3.4}$$

where v is called the velocity potential. Evaluating Eq. 6.3.2 outside the horizon for a scalar perturbed metric in arbitrary gauge, yields:

$$\delta\rho' = 3\Psi'(\rho + P) - 3\mathcal{H}(\delta\rho + \delta P). \tag{6.3.5}$$

In Appendix E.2, we will discuss in more detail how this equation is derived.

Now, in order to proceed, it proves convenient to go to the uniform density gauge, defined by the constraint $\delta\rho \equiv 0$. In this gauge, the curvature perturbation on slices of uniform energy density can be written as:

$$\boxed{\zeta \equiv \Psi + \mathcal{H}\frac{\delta\rho}{\rho'} \quad \xrightarrow{\delta\rho \equiv 0} \quad \zeta = \Psi,} \tag{6.3.6}$$

by the very definition of ζ and hence $\zeta' = \Psi'$. The equation for $\delta\rho'$ can then be rewritten as:

$$\zeta'(\rho + P) = \mathcal{H}\delta P. \tag{6.3.7}$$

The generic pressure perturbation δP can be decomposed into adiabatic and non-adiabatic parts:

$$\delta P = \delta P_{\rm ad} + \delta P_{\rm nad}. \tag{6.3.8}$$

[6] In the context of quantum mechanics, the conservation of $\mathcal{R}$ on super-horizon scales is proved in [10] and [80] as a quantum operator statement for $\mathcal{R}$, which means that the proof is valid at all orders.

On account of Eq. 4.6.13, we recognize:

$$\delta P_{\rm ad} \equiv c_s^2 \delta\rho, \qquad \delta P_{\rm nad} = -3c_s^2(\rho + P)\mathcal{S}. \tag{6.3.9}$$

where c_s^2 is the sound speed and $\mathcal{S}$ is the isocurvature perturbation. That is, for the adiabatic part there exists a direct relationship between the pressure and energy density since then $\mathcal{S} = 0$. In the uniform density gauge, the adiabatic pressure perturbation will therefore vanish per definition:

$$\delta P_{\rm ad} \xrightarrow{\delta\rho \equiv 0} 0. \tag{6.3.10}$$

In contrast to the adiabatic pressure perturbation, the non-adiabatic part will survive in the uniform density gauge.

Therefore, we can now obtain a first order equation for the time evolution of ζ on super-horizon scales in the uniform density gauge, where $\delta P = \delta P_{\rm nad}$:

$$\boxed{\zeta' = \mathcal{R}' = \mathcal{H}\frac{\delta P_{\rm nad}}{(\rho + P)} = -3c_s^2\mathcal{H}\mathcal{S},} \tag{6.3.11}$$

where we have used the fact that ζ and $\mathcal{R}$ are equal outside the horizon. Notice that, dispite of the completely different approach taken here compared to the previous section, we obtain exactly the same result for the super-horizon evolution of $\mathcal{R}$. From the above equation, we conclude that if the pressure perturbation is adiabatic (i.e. $\delta P_{\rm nad} = 0$ or equivalently $\mathcal{S} = 0$), the curvature perturbation $\zeta = \mathcal{R}$ is constant outside the horizon. For single-field inflation, we already argued that the perturbations are of the adiabatic type.

6.4 Weinberg's Proof

Given its significance in the connection to late observables, a generic and rigorous proof for the constancy of the curvature perturbation on super-horizon scales is very important. However, the two proofs given before still rely on specific and arguably non-trivial assumptions and are therefore not completely rigorous. Here, we will review a proof by Weinberg [90] for the non-evolution of $\mathcal{R}$ on super-horizon scales. Before we review the proof, we will first give an explanation as to motivate why the preceding two proofs are not completely rigorous.

As the modes of cosmological interest are on super-horizon scales, the universe is believed to go through the era of reheating, in which the degrees of freedom during inflation decayed into the degrees of freedom of the standard model of particle physics. However, this process is not well known and even the dynamical equation governing the evolution of the perturbations are not well known. Notice that this observation cannot be reconciled with the assumptions underlying the proof based on the perturbed field equations. The second proof, based on the assumption that energy-momentum conservation for the perturbed stress-tensor applies, does not rely on the details of the perturbation equations. Nevertheless, we had to impose that the perturbations are of the adiabatic type and hence $\delta P_{\rm nad} = 0$ (which is indeed the case in the single-field scenario).

Weinberg's proof does not rely on such assumptions and shows that there exists always an adiabatic mode that freezes on super-horizon scales. To be more precise, Weinberg shows that, irrespective of the constituents of the universe, there always exists a pair of solutions in the super-horizon limit ($k \to 0$) of the following form:

1. A constant non-vanishing mode: $\mathcal{R} \neq 0$ and $\dot{\mathcal{R}} = 0$,

2. A decaying mode: $\mathcal{R} = 0$ and $\dot{\mathcal{R}} = 0$.[7]

In order to show the existence of such a pair of solutions irrespective of the constituents of the universe, i.e. without specifying the perturbations (e.g. $\delta\rho$, δP and v), we need to take a different approach in which we need not specify the perturbations explicitly.

We will follow Weinberg and consider a spatially homogeneous universe in which the perturbations only obey temporal dependence. In this case, it can be shown that the Newtonian gauge field equations in the super-horizon limit ($k \to 0$) are invariant under specific gauge transformations that are not symmetries of the unperturbed metric. Notice the similarity to Goldstone's theorem in QFT in this respect: since the metric satisfies the field equations before and after the induced gauge transformation, the change in the metric (as induced by the transformation) must also be a solution to the field equations. The constant mode $\mathcal{R}$ on super-horizon scales takes the place of the Goldstone mode which becomes a free particle in the limit of long-wavelength [92].

However, notice that in the limit $k \to 0$, the advocated solutions are just gauge modes, which are not necessarily physical. In order for them to physical, they should be extendable to the $k \neq 0$ regime, i.e. they should be the $k \to 0$ limit to solutions of the field equations for arbitrary wavenumber. Extending the gauge modes to the $k \neq 0$ regime results in a number of constraints on the solutions as imposed by the field equations for arbitrary wavelength.

6.4.1 Evolution of $\mathcal{R}$ on Super-Horizon Scales

To start, we will derive an equation governing the evolution of the curvature perturbation on super-horizon scales in the Newtonian gauge. We will start from Eq. 6.3.5, which is the super-horizon limit ($k/aH \ll 1$) of the perturbed energy-momentum conservation equation (Eq. 6.3.2):

$$\delta\dot{\rho} + 3H(\delta\rho + \delta P) = 3(\rho + P)\dot{\Psi}. \tag{6.4.2}$$

The curvature perturbation on slices of uniform energy density can be written as:

$$\zeta = \Psi + H\frac{\delta\rho}{\dot{\rho}} = \Psi - \frac{\delta\rho}{3(\rho + P)}, \tag{6.4.3}$$

where we used the background energy conservation equation $\dot{\rho} = -3H(\rho + P)$. The time derivative of the curvature perturbation on slices of uniform energy density can be written as:

$$\dot{\zeta} = \dot{\Psi} - \frac{\delta\dot{\rho}}{3(\rho + P)} + \frac{\delta\rho(\dot{\rho} + \dot{P})}{3(\rho + P)^2} = \frac{\dot{P}\delta\rho - \dot{\rho}\,\delta P}{3(\rho + P)^2} \equiv X, \tag{6.4.4}$$

where in the second equality we substituted the super-horizon result for $\delta\dot{\rho}$ from Eq. 6.3.5 and hence X is only equal to the time derivative of ζ outside the horizon. Since $\mathcal{R} = \zeta$ outside the horizon, we find that on those scales:

$$\dot{\mathcal{R}} = \dot{\zeta} = \frac{\dot{P}\delta\rho - \dot{\rho}\,\delta P}{3(\rho + P)^2} \equiv X. \tag{6.4.5}$$

[7] In the next subsection, we will show that outside the horizon $\dot{\mathcal{R}} = X$ with:

$$X \equiv \frac{\dot{P}\delta\rho - \dot{\rho}\,\delta P}{3(\rho + P)^2}. \tag{6.4.1}$$

In terms of X, the theorem thus states that there always exists a pair of solutions for which the first mode is constant and the second mode decays towards zero, while both modes satisfy $X \to 0$ (no evolution).

Hence, in terms of X the task is to show that outside the horizon, whatever the constituents of the universe are, there always exists a pair of solutions of the form:

1. A constant non-vanishing mode: $\mathcal{R} \neq 0$ and $X \to 0$,
2. A decaying mode: $\mathcal{R} = 0$ and $X \to 0$.

6.4.2 Gauge Transformations, Modes and Lie derivatives

Following Weinberg, we will consider a spatially homogeneous perturbed universe and show in the next subsection that outside the horizon ($k \to 0$), the Newtonian gauge allows for the existence of gauge modes.[8] Here, we will first review gauge transformations and the associated Lie derivatives. Then, we will discuss the notion of gauge modes. The more general concepts developed here will be applied subsequently to show the existence of the gauge modes in a spatially homogeneous perturbed FRW universe.

Consider the coordinate transformation:

$$x^\mu \to \tilde{x}^\mu = x^\mu + \xi^\mu(t, \boldsymbol{x}), \tag{6.4.6}$$

where we assume the space-time dependent shift ξ^μ to be infinitesimal. Under such a coordinate transformation, the metric changes as:

$$g_{\mu\nu}(x) \to \tilde{g}_{\mu\nu}(\tilde{x}) = g_{\lambda\kappa}(x)\frac{\partial x^\lambda}{\partial \tilde{x}^\mu}\frac{\partial x^\kappa}{\partial \tilde{x}^\nu}. \tag{6.4.7}$$

Here we denote the space-time dependence by means of x between parentheses. Notice that such a coordinate transformation affects the coordinates, background fields and the perturbations. Therefore, it is more convenient to work with *gauge* transformations instead, which act only on the perturbations. Furthermore, gauge transformations posses the property that they leave the (perturbed) Einstein field equations (EFE's) invariant. This property will exploited when discussing gauge modes below.

There exists a general prescription to construct gauge transformations from coordinate transformations (such as the one above), which we will review here first (based on [92]). Gauge transformations are obtained, after performing the coordinate transformation, by relabelling the coordinates by dropping the tilde on the coordinate argument (i.e. $\tilde{x} \to x$) and attributing the whole change in the metric $g_{\mu\nu} = \bar{g}_{\mu\nu} + \delta g_{\mu\nu}$ (as induced by the coordinate transformation) to a change only in the perturbation $\delta g_{\mu\nu}$. The change in the perturbation, denoted as $\Delta(\delta g_{\mu\nu})$, can then be written as:

$$\Delta(\delta g_{\mu\nu})(x) \equiv \tilde{g}(x) - g_{\mu\nu}(x) = \delta\tilde{g}_{\mu\nu}(x) - \delta g_{\mu\nu}(x), \tag{6.4.8}$$

where the second equality follows from the fact that the background metric is the same in both frames, $\tilde{\bar{g}}_{\mu\nu} = \bar{g}_{\mu\nu}$. In Appendix E.3, we will show that $\Delta(\delta g_{\mu\nu})$ can be written to first order in the form:

$$\Delta(\delta g_{\mu\nu})(x) = -\bar{g}_{\mu\kappa}\partial_\nu\xi^\kappa - \bar{g}_{\kappa\nu}\partial_\mu\xi^\kappa - \partial_\kappa\bar{g}_{\mu\nu}\xi^\kappa. \tag{6.4.9}$$

The right hand side of the above expression is known as the Lie derivative of the metric, denoted as:

$$\Delta(\delta g_{\mu\nu})(x) \equiv \Delta_\xi\, g_{\mu\nu}(x), \tag{6.4.10}$$

[8] As mentioned above, to show that these gauge modes are physical, we will subsequently extend them to the $k \neq 0$ regime.

see e.g. section 10.9 of [89].

The EFE's will be invariant solely if the same gauge transformation is also applied to all other tensors present, and in particular the energy-momentum tensor. Hence, we have to transform the perturbation $\delta T_{\mu\nu}$ in a similar manner:

$$\delta T_{\mu\nu} \to \delta T_{\mu\nu} + \Delta(\delta T_{\mu\nu}). \tag{6.4.11}$$

The explicit expression for $\Delta(\delta T_{\mu\nu})$ can be written as follows and equals the Lie derivative of the energy-momentum tensor:

$$\Delta(\delta T_{\mu\nu}) = -\bar{T}_{\lambda\mu}\partial_\nu\xi^\lambda - \bar{T}_{\lambda\nu}\partial_\mu\xi^\lambda - \partial_\lambda\bar{T}_{\mu\nu}\xi^\lambda \equiv \Delta_\xi\, T_{\mu\nu}(x). \tag{6.4.12}$$

Motivated by the fact that the proof to be given has to hold for any (perhaps unknown) matter constituent, ware are not to explicitly specify the form of the energy-momentum tensor. However, based on the Cosmological Principle (see sections 1.3.1 and 4.8.1), we expect the zeroth order energy-momentum tensor for any constituent to take the form of that for a perfect fluid:

$$\bar{T}_{\mu\nu} = P g_{\mu\nu} + (\rho + P) U_\mu U_\nu, \tag{6.4.13}$$

with $U_0 = -1$ and $U_i = 0$ (i.e. we work in the local rest frame). The explicit non-vanishing components are $\bar{T}_{00} = \rho$ and $\bar{T}_{ij} = a^2\delta_{ij}P$. We expect possible deviations from the perfect fluid description to arise only at first (or higher) order in perturbations.

Gauge Modes

With the notions of gauge transformation and Lie derivatives at our disposal, we will give a more precise definition of gauge modes. Observe that the linearly perturbed EFE's are linear differential equations, which are generically solved by the perturbed metric $\delta g_{\mu\nu}$ and the corresponding energy-momentum tensor $\delta T_{\mu\nu}$. Collectively, we may denote the solution to the EFE's as:

$$\{\delta g_{\mu\nu}, \delta T_{\mu\nu}\}. \tag{6.4.14}$$

As mentioned above, gauge transformations leave the field equations invariant. Hence, the gauge transformed metric and energy-momentum tensor also satisfy the EFE's. That is, the EFE's are also solved by:

$$\{\delta g_{\mu\nu} + \Delta_\xi g_{\mu\nu}, \delta T_{\mu\nu} + \Delta_\xi T_{\mu\nu}\}. \tag{6.4.15}$$

This solution governs the exact same physical content as the first set of solutions.

Now, since the above two sets solve the perturbed field equations, their difference will also solve the equations, since they are linear differential equations. That is, a third solution is given by:

$$\{\Delta_\xi g_{\mu\nu}, \Delta_\xi T_{\mu\nu}\}. \tag{6.4.16}$$

Solutions of this sort are referred to as gauge modes, since they can be regarded as being generated by the gauge transformation. To summarize, since $\{\delta g_{\mu\nu}, \delta T_{\mu\nu}\}$ and $\{\delta g_{\mu\nu} + \Delta_\xi\, g_{\mu\nu}, \delta T_{\mu\nu} + \Delta_\xi\, T_{\mu\nu}\}$ are solutions, their difference $\{\Delta_\xi\, g_{\mu\nu}, \Delta_\xi\, T_{\mu\nu}\}$ must also be a solution. Although this subtle observation might appear trivial, we will find that it constitutes a key ingredient to the proof.

Application to Perturbed FRW Space-Time

We will now evaluate the expressions for $\Delta(\delta g_{\mu\nu})$ and $\Delta(\delta T_{\mu\nu})$ in case of the flat FRW metric in cosmic time, for which $\bar{g}_{00} = -1$ and $\bar{g}_{ij} = a^2\delta_{ij}$. In other words, we will construct the gauge modes for a flat FRW universe. In cosmic time, the perturbed metric components are:

$$\begin{aligned}\delta g_{00} &= -2\Phi,\\ \delta g_{0i} &= aB_i,\\ \delta g_{ij} &= -2a^2(\Psi\delta_{ij} - E_{ij}),\end{aligned} \tag{6.4.17}$$

different from the expressions given in conformal time in section 4.3. The explicit components of $\Delta(\delta g_{\mu\nu})$ can be written as:

$$\begin{aligned}\Delta(\delta g_{00}) &= -2\dot{\xi}_0,\\ \Delta(\delta g_{0i}) &= -\partial_i\xi_0 - \dot{\xi}_i + 2H\xi_i,\\ \Delta(\delta g_{ij}) &= -2\partial_{(i}\xi_{j)} + 2a\dot{a}\xi_0\delta_{ij}.\end{aligned} \tag{6.4.18}$$

In the above, the index on ξ is raised and lowered using the FRW metric, so that $\xi_0 = -\xi^0$ and $\xi_i = a^2\xi^i$.

In order to examine the effect of the gauge transformation on the scalar, vector and tensor (SVT) components in the perturbed metric, we decompose the spatial part of the shift vector ξ^i as:

$$\xi_i = \partial_i\xi + \hat{\xi}_i, \tag{6.4.19}$$

where $\partial^i\hat{\xi}_i = 0$. Under this decomposition, the mixed and spatial components become:

$$\begin{aligned}\Delta\delta g_{0i} &= -\partial_i\xi_0 - \partial_i\dot{\xi} - \dot{\hat{\xi}}_i + 2H(\partial_i\xi + \hat{\xi}_i),\\ \Delta\delta g_{ij} &= -2\partial_i\partial_j\xi - 2\partial_{(i}\hat{\xi}_{j)} + 2a\dot{a}\xi_0\delta_{ij},\end{aligned} \tag{6.4.20}$$

and the purely temporal component does not change.

We will now consider how these expressions correspond to changes in the perturbations as induced by the gauge transformation. In the SVT-decomposed form of the perturbations, the expression for $\Delta(\delta g_{\mu\nu})$ is given by:

$$\begin{aligned}\Delta(\delta g_{00}) &= -2\Delta\Phi,\\ \Delta(\delta g_{0i}) &= a(\partial_i(\Delta B) + \Delta\hat{B}_i),\\ \Delta(\delta g_{ij}) &= -2a^2(\Delta\Psi\delta_{ij} - \Delta E_{ij}),\end{aligned} \tag{6.4.21}$$

and the tensor perturbation is decomposed in the usual way:

$$\Delta E_{ij} = \partial_i\partial_j\Delta E + \partial_{(i}\Delta\hat{E}_{j)} + \Delta\hat{E}_{ij}. \tag{6.4.22}$$

Equating the first equation in Eqs. 6.4.18 with the first in Eqs. 6.4.21 yields the following relation between $\Delta\Phi$ and ξ_0:

$$\Delta\Phi = \dot{\xi}_0. \tag{6.4.23}$$

Similarly, equating the remaining equations gives the following expressions for the individual SVT-components:

$$\begin{aligned}&\Delta B = \frac{1}{a}(-\xi_0 - \dot{\xi} + 2H\xi), && \Delta\hat{B}_i = \frac{1}{a}(-\dot{\hat{\xi}}_i + 2H\hat{\xi}_i),\\ &\Delta\Psi = -H\xi_0, \quad \Delta E = -\xi/a^2, && \Delta\hat{E}_i = -\xi_i/a^2, \quad \Delta\hat{E}_{ij} = 0.\end{aligned} \tag{6.4.24}$$

Notice that the purely tensorial perturbation $\hat{E}_{ij}$ is not affected by the gauge transformation: in line with the conclusion drawn in section 4.3.2.

We will now proceed by considering the energy-momentum tensor. Perturbing the background energy-momentum tensor (Eq. 6.4.13) to first order and allowing for anisotropic stress contributions in the form of $\delta\Sigma_{\mu\nu}$, we obtain:

$$\delta T_{\mu\nu} = \delta P g_{\mu\nu} + P\delta g_{\mu\nu} + (\delta\rho + \delta P)U_\mu U_\nu + (\rho + P)(\delta U_\mu U_\nu + U_\mu \delta U_\nu) + \delta\Sigma_{\mu\nu}. \tag{6.4.25}$$

As discussed before, the anisotropic contribution can be taken purely spatial. Using cosmic as the evolution variable,[9] and invoking the results $\delta U^0 = \delta U_0 = \delta g_{00}/2$ and $\delta U^i \equiv v^i$ from section 4.8.1 gives:

$$\begin{aligned}
\delta T_{00} &= \delta\rho - \rho\delta g_{00}, \\
\delta T_{i0} &= P\delta g_{i0} - (\rho + P)v_i, \\
\delta T_{ij} &= P\delta g_{ij} + a^2(\delta_{ij}\delta P + \delta\Sigma_{ij}).
\end{aligned} \tag{6.4.26}$$

For later convenience, we decompose the above expressions into SVT parts, which amounts to making the substitutions:

$$v_i = \partial_i v + \hat{v}_i, \qquad \delta\Sigma_{ij} = \partial_i\partial_j\delta\Sigma + \partial_{(i}\delta\hat{\Sigma}_{j)} + \delta\hat{\Sigma}_{ij}. \tag{6.4.27}$$

The purely vectorial and tensorial perturbations introduced above satisfy the constraints:

$$\partial_i\hat{v}_i = \partial_i\delta\hat{\Sigma}_i = \partial_i\delta\hat{\Sigma}_{ij} = \delta\hat{\Sigma}_{ii} = 0, \tag{6.4.28}$$

in analogy with the divergenceless 3-vector $\hat{\beta}_i$ and 3-tensor $\hat{\gamma}_{ij}$ introduced in section 4.2. The mixed and purely spatial components then become:

$$\begin{aligned}
\delta T_{i0} &= P\delta g_{i0} - (\rho + P)(\partial_i v + \hat{v}_i), \\
\delta T_{ij} &= P\delta g_{ij} + a^2(\delta_{ij}\delta P + \partial_i\partial_j\delta\Sigma + \partial_{(i}\delta\hat{\Sigma}_{j)} + \delta\hat{\Sigma}_{ij}).
\end{aligned} \tag{6.4.29}$$

Under the gauge transformation, the perturbed elements of the energy-momentum tensor change according to Eq. 6.4.12 as follows:

$$\begin{aligned}
\Delta(\delta T_{00}) &= 2\rho\dot{\xi}_0 + \dot{\rho}\xi_0, \\
\Delta(\delta T_{i0}) &= -P\dot{\xi}_i + \rho\partial_i\xi_0 + 2HP\xi_i, \\
\Delta(\delta T_{ij}) &= -2P_{(i}\xi_{j)} + (2a\dot{a} + a^2\dot{P})\xi_0\delta_{ij}.
\end{aligned} \tag{6.4.30}$$

After substituting the scalar-vector decomposition for the spatial shift $\xi_i = \partial_i\xi + \hat{\xi}_i$ and the velocity $v_i = \partial_i v + \hat{v}_i$, we can identify the transformations of all perturbations by equating the above equations to the corresponding expressions for $\Delta(\delta T_{\mu\nu})$, which can be obtained from Eqs. 6.4.26 as:

$$\begin{aligned}
\Delta(\delta T_{00}) &= \Delta(\delta\rho) - \rho\Delta(\delta g_{00}), \\
\Delta(\delta T_{i0}) &= P\Delta(\delta g_{i0}) - (\rho + P)(\partial_i\Delta v + \Delta\hat{v}_i), \\
\Delta(\delta T_{ij}) &= P\Delta(\delta g_{ij}) + a^2\left(\delta_{ij}\Delta(\delta P) + \Delta(\delta\Sigma_{ij})\right).
\end{aligned} \tag{6.4.31}$$

[9] This is contrary to section 4.8, in which we used conformal time $d\tau = dt/a$, so the expressions stated here will be different.

The transformations of the perturbations $\delta\rho$, δP and the velocity potential v can be extracted as:

$$\Delta(\delta\rho) = \dot{\rho}\xi_0, \qquad \Delta(\delta P) = \dot{P}\xi_0, \qquad \Delta v = -\xi_0. \tag{6.4.32}$$

The other ingredients of the perturbed energy-momentum tensor are invariant under the gauge transformation. In particular:

$$\Delta\hat{v}_i = \Delta(\delta\Sigma) = \Delta(\delta\hat{\Sigma}_i) = \Delta(\delta\hat{\Sigma}_{ij}) = 0. \tag{6.4.33}$$

6.4.3 Gauge Modes

As mentioned in the previous section, we will now derive that a spatially homogeneous perturbed universe allows for gauge modes in the Newtonian gauge in the limit $k \to 0$. In a spatially homogeneous Newtonian gauge, the scalar and tensor metric perturbations can be written as:

$$\begin{aligned} \delta g_{00} &= -2\Phi(t), \\ \delta g_{0i} &= aB_i(t), \\ \delta g_{ij} &= -2a^2(\Psi(t)\delta_{ij} - E_{ij}(t)), \end{aligned} \tag{6.4.34}$$

Now we consider a generic gauge transformation (Eq. 6.4.6) and examine what constraints are needed on the perturbations to stay in the spatially homogeneous Newtonian gauge after the transformation.

▷ *Temporal Component.*—First we consider δg_{00} and use the first equation of Eqs. 6.4.18 obtain the required constraint. For δg_{00} to remain spatially homogeneous, we conclude that the temporal shift ξ_0 must of the form:

$$\xi_0(\boldsymbol{x}, t) = \bar{\xi}(t) + \chi(\boldsymbol{x}). \tag{6.4.35}$$

Then the change in the perturbation, $\Delta(\delta g_{00})$ can be expressed as:

$$\Delta(\delta g_{00}) = -2\dot{\xi}_0 = -2(\dot{\bar{\xi}} + \dot{\chi}) = -2\dot{\bar{\xi}} = -2\Delta\Phi(t), \tag{6.4.36}$$

where the last equality is obtained from Eq. 6.4.23. Hence, $\Delta\Phi$ can be extracted to be:

$$\Delta\Phi = \dot{\bar{\xi}}. \tag{6.4.37}$$

▷ *Mixed Component.*—In order for the perturbation δg_{0i} to preserve the Newtonian gauge after the transformation, we require $\Delta(\delta g_{0i}) \equiv 0$ since δg_{0i} vanishes by definition in the Newtonian gauge. Explicitly:

$$\Delta(\delta g_{0i}) - \partial\xi_0 - \dot{\xi}_i + 2H\xi_i = -\partial_i\chi(\boldsymbol{x}) - \dot{\xi}_i + 2H\xi_i \equiv 0. \tag{6.4.38}$$

This constraint can be used to find a differential equation for the spatial shift:

$$\dot{\xi}_i = -\partial_i\chi(\boldsymbol{x}) + 2H\xi_i. \tag{6.4.39}$$

Integrating the above equation gives the solution:

$$\xi_i(t, \boldsymbol{x}) = a^2(t)f_i(\boldsymbol{x}) - a^2(t)\partial_i\chi(\boldsymbol{x})\int\frac{dt}{a^2(t)}, \tag{6.4.40}$$

with $f_i(\boldsymbol{x})$ an arbitrary 3-vector with spatial dependence.

▷ *Spatial Component.*—Inserting the above expression for $\xi_i(t, \boldsymbol{x})$ into the expression for $\Delta(\delta g_{ij})$ yields:

$$\Delta(\delta g_{ij}) = -2a^2\partial_{(i}f_{j)} - 2\partial_i\partial_j\chi(\boldsymbol{x})a^2(t)\int\frac{dt}{a^2(t)} + 2a\dot{a}(\chi(\boldsymbol{x}) + \bar{\xi}(t)). \tag{6.4.41}$$

In order not to introduce any spatial-dependence in $\Delta(\delta g_{ij})$, we conclude that χ must be a constant without spatial dependence. In that case, in can be absorbed into the definition of $\bar{\xi}$ so that $\chi \equiv 0$ and the second term in the above expression vanishes. By similar reasoning, the vector $f_i(\boldsymbol{x})$ must take the form $f_i(\boldsymbol{x}) = \omega_{ij}x^j$, with ω_{ij} a constant matrix since then:

$$2\partial_{(i}f_{j)} = (\omega_{ij} + \omega_{ji}). \tag{6.4.42}$$

Now, $\Delta(\delta g_{ij})$ can be written as follows:

$$\Delta(\delta g_{ij}) = -a^2(\omega_{ij} + \omega_{ji}) + 2a\dot{a}\bar{\xi}(t). \tag{6.4.43}$$

To proceed and identify the contributions in the above expression with the changes in the perturbations $\Delta\Psi$ and ΔE_{ij}, we first split the generic rank two tensor ω_{ij} into the irreducible representations of the rotation group $SO(3)$. The decomposition reads [22]:

$$\omega_{ij} = \frac{1}{3}\omega_{kk}\delta_{ij} + \frac{1}{2}(\omega_{ij} - \omega_{ji}) + \frac{1}{2}\left(\omega_{ij} + \omega_{ji} - \frac{2}{3}\omega_{kk}\delta_{ij}\right), \tag{6.4.44}$$

where the first term in the above decomposition is invariant under spatial rotations. The second term is the antisymmetric part whereas the third is symmetric and hence they do not couple under $SO(3)$ rotations. Notice that antisymmetric contribution can be omitted, as the unperturbed FRW metric is invariant under spatial rotations [92] and hence considering an antisymmetric contribution to ω_{ij} has no effect, i.e. we take $\omega_{ij} = \omega_{ji}$. The decomposition then simplifies to:

$$\omega_{ij} = \frac{1}{3}\omega_{kk}\delta_{ij} + \left(\omega_{ij} - \frac{1}{3}\omega_{kk}\delta_{ij}\right). \tag{6.4.45}$$

In the above and final expression for ω_{ij}, the first term transforms as a scalar under rotations (i.e. it does not transform) and the terms between parentheses constitute a traceless symmetric tensor and transforms as an ordinary rank 2 tensor. Substituting the above expression for ω_{ij} into Eq. 6.4.43 and setting the resulting expression equal to Eq. 6.4.21, we obtain:

$$\Delta\Psi = \frac{1}{3}\omega_{kk} - H\bar{\xi}(t), \qquad \Delta E_{ij} = \frac{1}{3}\omega_{kk}\delta_{ij} - \omega_{ij}. \tag{6.4.46}$$

▷ *Ingredients of the Energy-Momentum Tensor.*—Spatial homogeneity also implies that the ingredients of the perturbed energy-momentum tensor are functions only of time. In particular, for scalar perturbations (such as the velocity potential v and perturbations $\delta\rho$ and δP) we have:

$$\Delta(\delta P) = -\dot{P}\bar{\xi}, \qquad \Delta(\delta\rho) = -\dot{\rho}\bar{\xi}, \qquad \Delta v = -\bar{\xi}, \qquad \Delta(\delta\Sigma) = 0, \tag{6.4.47}$$

and for the vectorial and tensorial contributions:

$$\Delta\hat{v}_i = \Delta(\delta\Sigma) = \Delta(\delta\hat{\Sigma}_i) = \Delta(\delta\hat{\Sigma}_{ij}) = 0. \tag{6.4.48}$$

Now we invoke the trivial but powerful observation that since $\{\delta g_{\mu\nu}, \delta T_{\mu\nu}\}$ and $\{\delta g_{\mu\nu} + \Delta_\xi\, g_{\mu\nu}, \delta T_{\mu\nu} + \Delta_\xi\, T_{\mu\nu}\}$ are solutions to the field equations, their difference $\{\Delta_\xi\, g_{\mu\nu}, \Delta_\xi\, T_{\mu\nu}\}$ must also be a solution. That is, we conclude that the scalar perturbations:

$$\begin{aligned} \Psi &= H\bar{\xi}(t) - \frac{1}{3}\omega_{kk}, & \Phi &= -\dot{\bar{\xi}}, \\ \delta P &= -\dot{P}\bar{\xi}, \qquad \delta\rho = -\dot{\rho}\bar{\xi}, & v &= \bar{\xi}, \qquad \delta\Sigma = 0, \end{aligned} \tag{6.4.49}$$

always satisfy the $k \to 0$ field equations in the spatially homogenous Newtonian gauge.[10] For tensor modes, the solutions become:

$$E_{ij} = \frac{1}{3}\omega_{kk}\delta_{ij} - \omega_{ij}, \qquad \delta\hat{\Sigma}_{ij} = 0. \tag{6.4.51}$$

In conclusion, Eqs. 6.4.49 and 6.4.51 are solutions for the Fourier transformations of the scalar en tensor perturbations of zero wavenumber, respectively.

6.4.4 Promoting Gauge Modes to Arbitrary Wavenumber

So far, the solutions (Eqs. 6.4.49 and 6.4.51) for the scalar and tensor perturbations are merely gauge modes, applicable only at zero wavenumber. For those modes to be physical, we have to extend them to the regime of arbitrary wavenumber, resulting in a number of constraints on the solutions. That is, the $k = 0$ solutions should be the $k \to 0$ limit of solutions to the field equations in the $k \neq 0$ regime.

We will consider the perturbed field equations (Eq. 4.9.16–4.9.19) including anisotropic stress, examine which of them disappear in the limit $k \to 0$ and derive the resulting constraints. In Fourier space, vanishing wavenumber implies vanishing partial derivatives, as $\partial_i \leftrightarrow ik_i$. Hence, we consider those field equations which include spatial derivatives and therefore disappear in the zero wavenumber limit. The equations which disappear are only Eqs. 4.9.17 and 4.9.19, which read:

$$\begin{aligned} -\partial_i(\Psi' + \mathcal{H}\Phi) &= \Delta(a)(\rho + P)v_i, \\ \partial_j\partial^i(\Psi - \Phi) &= 2\Delta(a)\delta\Sigma^i_j. \end{aligned} \tag{6.4.52}$$

For scalar perturbations in the limit $k \to 0$, the l.h.s. of the first equation vanishes and the r.h.s. can be rewritten by setting $v_i = \partial_i v$ to obtain:

$$0 = \Delta(a)(\rho + P)\partial_i v. \tag{6.4.53}$$

Since $v = \bar{\xi}(t)$ possesses no spatial dependence, $\partial_i v$ vanishes and the above equation is trivially satisfied. Therefore, although the first equation vanishes for zero wavenumber, no additional constraints have to be imposed. The second equation can be rewritten for scalar perturbations by substituting $\delta\Sigma^i_j = \delta\Sigma\, \delta^i_j$. We know that $\delta\Sigma = 0$ is a gauge mode solution. However, note that this solution implies:

$$\Phi = \Psi, \tag{6.4.54}$$

in the second equation of Eqs. 6.4.52. Hence, for the scalar gauge modes to be extended to non-zero wavenumber, we find the constraint that the gravitational potentials must be equal.

[10] Here we consider a single fluid. For a set of perfect fluids labelled by (α), the results would be [90, 92]:

$$\delta P_{(\alpha)} = -\dot{P}_{(\alpha)}\bar{\xi}, \qquad \delta\rho_f = -\dot{\rho}_{(\alpha)}\bar{\xi}, \qquad \delta\Sigma_{(\alpha)} = 0. \tag{6.4.50}$$

6.4.5 Existence of Two Adiabatic Modes

Now we will combine the results obtained above to show the existence of two adiabatic modes for the curvature perturbation on large scales, of which one is a constant mode and the other decays over time.

Constant Mode

We will start by proving the existence of the constant mode. Using the constraint and Eqs. 6.4.49, we can derive the following evolution equation for the function $\bar{\xi}(t)$:

$$\dot{\bar{\xi}} = \omega_{kk}/3 - H\bar{\xi}. \tag{6.4.55}$$

In addition, by definition of the curvature perturbation in the Newtonian gauge:

$$\mathcal{R} = \Psi - Hv, \tag{6.4.56}$$

we find using Eqs. 6.4.49 that there exists a constant mode for the curvature perturbation, denoted by $\mathcal{R}^0$, which reads:

$$\mathcal{R}^0 \equiv -\omega_{kk}/3. \tag{6.4.57}$$

The differential equation for $\dot{\bar{\xi}}$ can then be written as:

$$\dot{\bar{\xi}} = -\mathcal{R}^0 - H\bar{\xi}, \tag{6.4.58}$$

whose solution to this differential equation is given by:

$$\boxed{\bar{\xi}(t) = -\frac{\mathcal{R}^0}{a(t)} \int_{\mathcal{T}}^{t} a(\tilde{t})\, d\tilde{t},} \tag{6.4.59}$$

where the lower integration limit $\mathcal{T}$ is arbitrary. On account of the constraint $\Phi = \Psi$ and the expression for Ψ in Eqs. 6.4.49, the solutions to the gravitational potentials become:

$$\Phi = \Psi = \mathcal{R}^0 \left[1 - \frac{H(t)}{a(t)} \int_{\mathcal{T}}^{t} a(\tilde{t})\, d\tilde{t} \right]. \tag{6.4.60}$$

Furthermore, now that we have a solution for $\bar{\xi}(t)$, we can write energy-momentum tensor ingredients ($\delta P/\dot{P}$, $\delta\rho/\dot{\rho}$ and v) as:

$$\frac{\delta P}{\dot{P}} = \frac{\delta\rho}{\dot{\rho}} = -v = \frac{\mathcal{R}^0}{a(t)} \int_{\mathcal{T}}^{t} a(\tilde{t})\, d\tilde{t}. \tag{6.4.61}$$

Decaying Mode

Since Eq. 6.4.59 is a solution for $\bar{\xi}$ with any lower integration constant $\mathcal{T}$, the difference between two such solutions with $\mathcal{T}_2 \neq \mathcal{T}_1$ is also a solution. However, for this solution $\mathcal{R}^0 = 0$ and so Eq. 6.4.58 reduces to:

$$\dot{\bar{\xi}} = -H\bar{\xi}, \tag{6.4.62}$$

which has the solution:

$$\boxed{\bar{\xi}(t) = \frac{\mathcal{C}}{a(t)},} \tag{6.4.63}$$

for an arbitrary constant $\mathcal{C}$. As this solution decays away over time it can be ignored at late times (large values of a) and is therefore the decaying mode. The solutions to the gravitational potentials and the energy-momentum tensor perturbations become:

$$\Phi = \Psi = \frac{H(t)\mathcal{C}}{a(t)}, \qquad \frac{\delta P}{\dot{P}} = \frac{\delta\rho}{\dot{\rho}} = -v = -\frac{\mathcal{C}}{a(t)}. \tag{6.4.64}$$

Notice that as $H(t)$ and $a(t)$ decrease and increase over time, respectively, the gravitational potentials are indeed also decaying modes.

Adiabicity

In conclusion, we have thus shown that, regardless of the constituents of the universe, the pair of solutions advocated in the introduction to this section indeed exists. The constant mode corresponds to:

$$\mathcal{R}^0 = -\omega_{kk}/3, \qquad X \to 0, \tag{6.4.65}$$

with ω_{kk} time-independent and therefore $X \equiv \dot{\mathcal{R}}$ (Eq.) tends to zero. The corresponding solutions to the gravitational potentials and the ingredients of the energy-momentum tensor are given by Eqs. 6.4.60 and 6.4.61. Moreover, we found a decaying mode for which:

$$\mathcal{R}^0 = 0, \qquad X \to 0, \tag{6.4.66}$$

and the solutions to the associated perturbations are given by Eqs. 6.4.64.

Notice that for both modes, we have the resulting condition that:

$$\frac{\delta P}{\dot{P}} = \frac{\delta\rho}{\dot{\rho}} = \frac{\delta P_{(\alpha)}}{\dot{P}_{(\alpha)}} = \frac{\delta\rho_{(\alpha)}}{\dot{\rho}_{(\alpha)}}, \tag{6.4.67}$$

where the last two equalities hold in case of multiple fluids labelled by (α). The above result coincides with the defining condition for adiabatic perturbations (section 4.6, Eq. 4.6.6). Hence, both obtained solutions are said to be of the adiabatic type. Therefore, we arrive at an important statement: if the perturbed field equations have no more than two independent solutions, then any perturbation must be adiabatic.

For single-field inflation, we recall that the curvature perturbation is given by:

$$\mathcal{R}_k = f_k/z. \tag{6.4.68}$$

Hence, we have only one $\mathcal{R}$ mode in the case of single-field inflation, which is necessarily adiabatic by the above theorem. Therefore, perturbations generated by single-field inflation are adiabatic. To summarize, quantum fluctuations in the inflaton excites solely adiabatic modes in the comoving curvature perturbation and those remain adiabatic as long as they are outside the horizon.

6.5 Adiabicity after Single-Field Inflation

In the previous section, we have proven that there always exists a pair of adiabatic solutions for the comoving curvature perturbation $\mathcal{R}$ outside the horizon. Assuming only the inflation makes a contribution to the energy-momentum tensor during inflation, we have shown that

quantum fluctuations during inflation excite the adiabatic $\mathcal{R}$-modes. Hence, after single-field inflation, modes will remain adiabatic as long as they are outside the horizon. This statement remains valid when the modes enter the reheating stage after inflation (provided they are still outside the horizon), since the proof for the existence of the adiabatic modes is valid irrespective of the constituents of the universe.[11] Therefore, as long as modes are outside the horizon, they will remain adiabatic during reheating after single-field inflation.

For the adiabatic mode, the comoving curvature perturbations is a time-independent constant $\mathcal{R}^0$. The gravitational potentials can be written as:

$$\Phi = \Psi = \mathcal{R}^0(1 + H\mathcal{I}), \qquad \mathcal{I}(t) \equiv -\frac{H(t)}{a(t)} \int_{\mathcal{T}}^{t} a(\tilde{t}) d\tilde{t}. \tag{6.5.1}$$

In addition, in terms of $\mathcal{I}(t)$, the density and pressure perturbations become:

$$\delta\rho = -\dot{\rho}\mathcal{I}, \qquad \delta P = -\dot{P}\mathcal{I}, \tag{6.5.2}$$

or more generically for any four-scalar S, such as the inflaton ϕ, the perturbation reads:

$$\delta S = -\dot{S}\mathcal{I}. \tag{6.5.3}$$

In case of a system of fluids labelled by (α), the above relations hold for perturbations in each individual fluid.

The above reasoning contains one weak assumption, namely that the only relevant degrees of freedom during inflation are the inflaton and gravitational fields. For reheating to occur, there must have been other fields[12] besides the inflaton and gravitational fields during inflation, which accommodate the production of matter particles during reheating. Those additional degrees of freedom did not suddenly come into existence at the onset of the reheating stage. In other words, the interaction between the inflaton and matter fields cannot be completely absent during the inflationary stage. Hence, the assumption that the inflaton and gravitional fields are the only degrees of freedom during inflation cannot be true, although it can still be a reasonable assumption in case the energy density in the matter fields is sufficiently small during the inflationary stage. Therefore, the question is whether perturbations in the matter fields can induce non-adiabatic contributions to $\mathcal{R}$ during the reheating stage.

This question is answered by another theorem due to Weinberg [91], which states that even if the decay of the inflaton ϕ during inflation produces perturbations in the matter fields φ_I are initially not at all adiabatic, the departure from adiabicity will have decayed exponentially when the energy density in the matter fields will become large during the reheating stage. Since the adiabatic density perturbation in the matter fields can be written as $\delta\rho_\varphi = -\dot{\rho}_\varphi\mathcal{I}$, the departure of generic $\delta\rho_\varphi$ from adiabicity can be measured by the quantity N, defined as:[13]

$$N \equiv \left| \frac{\Delta_{\mathrm{AD},\varphi}}{\rho_\varphi + P_\varphi} \right|, \qquad \Delta_{\mathrm{AD},\varphi} \equiv \delta\rho_\varphi + \dot{\rho}_\varphi\mathcal{I}. \tag{6.5.4}$$

[11] Which are, up till now, still speculative during the era of reheating.

[12] We will call those additional degrees of freedom *matter* fields, although the fields may also describe radiation species.

[13] For notational convenience, we will adopt the convention that any quantity A which can be written as the sum over all matter fields as $A_\varphi \equiv \sum_I A_{\varphi_I}$, e.g. $\delta\rho_\varphi \equiv \sum_I \delta\rho_{\varphi_I}$.

We will derive that the time evolution of N over time can be expressed as:

$$\frac{d}{dt}\ln N = -\frac{(1+c_\varphi^2)\hat{Q}_\varphi}{(1+w_\varphi)\rho_\varphi}, \tag{6.5.5}$$

here the source term $\hat{Q}_\varphi$ quantifies the energy transfer from the inflaton field to the matter fields φ_I, c_φ^2 is the sound speed of the matter fields together and w_φ is the combined equation of state (explicit expressions will be given below). Using this equation, one concludes that after some time during the reheating stage, when most of the energy in the inflaton ρ_φ is transferred to the matter fields ($\rho_\phi \ll \rho_\varphi$), N has decayed exponentially:[14]

$$N \to 0, \qquad (\rho_\varphi \gg \rho_\phi). \tag{6.5.6}$$

Therefore, even in case perturbations in the matter fields are initially not adiabatic, they will become adiabatic eventually become adiabatic. In conclusion, during reheating after single-field inflation non-adiabatic perturbations will not be produced as long as the modes are outside the horizon.

6.5.1 Combined Inflaton-Matter Fluid

To prove the above theorem, we consider the inflaton scalar field ϕ accompanied with whatever matter fields are present during inflation, denoted as φ_I. The combined system being the inflaton-matter fluid. The total energy-momentum tensor can be written as:

$$T^\mu_\nu = \sum_{(\alpha)} T^\mu_{\nu(\alpha)}, \tag{6.5.7}$$

where the label $(\alpha) \equiv (\phi, \varphi_I)$ runs over all fields in the combined fluid. The total energy-momentum tensor is covariantly conserved, but we allow for energy transfer between the different fields in the combined fluid by introducing a source term $Q_{\nu(\alpha)}$ as follows:

$$\nabla_\mu T^\mu_{\nu(\alpha)} = Q_{\nu(\alpha)}. \tag{6.5.8}$$

From the conservation of the total energy-momentum tensor, we obtain the following constraint on the individual source terms:

$$\sum_{(\alpha)} Q_{\nu(\alpha)} = 0. \tag{6.5.9}$$

In principle, interactions between all the fields are allowed, so that the source term for field φ_I is a generic function of all the other fields (and possibly their time derivatives):

$$Q_{\nu(I)} = \hat{Q}_{\nu(I)}(\phi, \varphi_J, \Phi, \Psi), \tag{6.5.10}$$

where $J \neq I$ since we consider energy transfer between different fields; for $J = I$ the energy remains in field φ_I. However, we will argue below that at early times the source terms depend solely on the inflaton and gravitational potentials.

[14] As we will show, it is not sufficient to show that $\Delta_{\rm AD}$ decays during reheating, as this can also be achieved for both $\delta\rho_\phi$ and $-\rho_\varphi\mathcal{I}$ decaying separately. By construction, the result $N \to 0$ shows that the combination $\delta\rho_\varphi + \rho_\varphi\mathcal{I}$ tends to zero and hence that perturbations become adiabatic.

For convenience, we decompose the source term $Q_{\nu(\alpha)}$ in the following way:

$$Q_{\nu(\alpha)} = \hat{Q}_{(\alpha)} U_\nu + f_{\nu(\alpha)}, \tag{6.5.11}$$

where f_ν is defined to be orthogonal to the 4-velocity, i.e. $U^\nu f_\nu = 0$. The 4-velocity satisfies the constraint $U_\nu U^\nu = -1$ and can be written at background level as $U^\nu = \delta^\nu_0$ and $U_\nu = -\delta^0_\nu$.[15] Hence, we can write:

$$-U^\nu \nabla_\mu T^\mu_{\nu(\alpha)} = \hat{Q}_{(\alpha)}. \tag{6.5.12}$$

At zeroth and first order in perturbations, the $\nu = 0$ equation gives gives:

$$\dot{\rho}_{(\alpha)} + 3H(\rho_{(\alpha)} + P_{(\alpha)}) = \hat{Q}_{(\alpha)}, \tag{6.5.13}$$

$$\delta\dot{\rho}_{(\alpha)} + 3H(\delta\rho_{(\alpha)} + \delta P_{(\alpha)}) - 3(\rho_{(\alpha)} + P_{(\alpha)})\dot{\Psi} = \delta\hat{Q}_{(\alpha)} + \Phi\hat{Q}_{(\alpha)}. \tag{6.5.14}$$

Derivation: Sourced Energy-Momentum Equations

At background level, the $\nu = 0$ equation gives the sourced continuity equation:

$$\dot{\rho}_{(\alpha)} + 3H(\rho_{(\alpha)} + P_{(\alpha)}) = \hat{Q}_{(\alpha)}, \tag{6.5.15}$$

on account of the fact that $\nabla_\mu \bar{T}^\mu_0$ is given by:

$$\nabla_\mu \bar{T}^\mu_{0(\alpha)} = -\dot{\rho}_{(\alpha)} - 3H(\rho_{(\alpha)} + P_{(\alpha)}). \tag{6.5.16}$$

At first order in perturbations, we obtain:

$$\delta U^0 \nabla_\mu T^\mu_0 + U^0 \nabla_\mu \delta T^\mu_0 = -\delta\hat{Q}, \tag{6.5.17}$$

Outside the horizon, where we can neglect spatial derivatives, we have already obtained an explicit expression for $\nabla_\mu \delta T^\mu_0$ in Appendix E.2, reading:

$$\nabla_\mu \delta T^\mu_0 = -\delta\dot{\rho} - 3H(\delta\rho + \delta P) + 3(\rho + P)\dot{\Psi}. \tag{6.5.18}$$

On account of this result, the first order equation becomes:

$$\delta\dot{\rho}_{(\alpha)} + 3H(\delta\rho_{(\alpha)} + \delta P_{(\alpha)}) - 3(\rho_{(\alpha)} + P_{(\alpha)})\dot{\Psi} = \delta\hat{Q}_{(\alpha)} + \Phi\hat{Q}_{(\alpha)}. \tag{6.5.19}$$

6.5.2 Early Times

We assume that at early times during inflation energy density in the matter fields φ_I is small compared to the inflation density: $\rho_\phi \gg \rho_\varphi$. This assumption is reasonable, since the energy density of fermions and gauge fields as produced by quantum fluctuations is quadratic in those fluctuations [91]. Under these circumstances, the inflaton provides the dominant source to the gravitational fields. In that case, the fluctuations in the inflaton and metric are given by:

$$\delta\phi = -\dot{\phi}\mathcal{I}, \qquad \Phi = \Psi = \mathcal{R}^0 + H\mathcal{I} \tag{6.5.20}$$

[15] At first order, we obtain $\delta U^\nu = (-\Phi, v^i)$ and $\delta U^\nu = (-\Phi, -v_i)$, see section 4.8.

Since the predominant fraction of the total energy density is contained in the inflaton and gravitational fields, the energy transfer from the inflaton towards the matter fields will depend solely on the inflaton, gravitational potentials and possibly their time derivatives. Hence, we can write the perturbation to the source term $\delta\hat{Q}_{(\alpha)}$ in the adiabatic way:

$$\delta\hat{Q}_{(\alpha)} = -\hat{Q}_{(\alpha)}\mathcal{I}, \tag{6.5.21}$$

since $\hat{Q}$ is a four-scalar. Now, the right-hand-side of Eq. 6.5.14 can be written as:

$$\delta\hat{Q}_{(I)} + \Phi\hat{Q}_{(I)} = -\frac{d}{dt}(\hat{Q}_{(I)}\mathcal{I}), \tag{6.5.22}$$

for $(\alpha) = \varphi_I$, i.e. we consider the evolution of the matter fields.

6.5.3 Evolution of Departures from Adiabicity

Now we proceed by deriving an equation governing the time-evolution for the departure of matter perturbations from adiabicity. Since the adiabatic density perturbation in the matter fields is $\delta\rho_\varphi = -\dot{\rho}_\varphi\mathcal{I}$, the departure of generic $\delta\rho_\varphi$ from adiabicity for each matter field φ_I can be measured by the difference:

$$\Delta_{\mathrm{AD},\varphi_I} = \delta\rho_{\varphi_I} + \dot{\rho}_{\varphi_I}\mathcal{I}, \tag{6.5.23}$$

and the total non-adiabatic contribution, summed over all matter fields, becomes simply: $\Delta_{\mathrm{AD},\varphi} = \delta\rho_\varphi + \dot{\rho}_\varphi\mathcal{I}$. Thus, we require an equation for the time evolution of $\Delta_{\mathrm{AD},\varphi}$. This equation is obtained by adding the time derivative of Eq. 6.5.13 times $\mathcal{I}$ and Eq. 6.5.14 for $(\alpha) = \varphi$ (i.e. we sum over all matter field), yielding:

$$\frac{d}{dt}(\dot{\rho}_\varphi\mathcal{I}) = -3\dot{H}(\rho_\varphi + P_\varphi)\mathcal{I} + 3H(\delta\rho_\varphi + \delta P_\varphi) - 3H(\rho_\varphi + P_\varphi)\dot{\mathcal{I}} + \frac{d}{dt}(\hat{Q}_\varphi\mathcal{I}), \tag{6.5.24}$$

$$\frac{d}{dt}(\delta\rho_\varphi) = -3H(\delta\rho_\varphi + \delta P_\varphi) + 3(\rho_\varphi + P_\varphi)\dot{\Psi} - \frac{d}{dt}(\hat{Q}_\varphi\mathcal{I}). \tag{6.5.25}$$

Adding those two equations gives:

$$\dot{\Delta}_{\mathrm{AD},\varphi} = -3\dot{H}(\rho_\varphi + P_\varphi)\mathcal{I} - 3H(\rho_\varphi + P_\varphi)\dot{\mathcal{I}} + 3(\rho_\varphi + P_\varphi)\dot{\Psi}. \tag{6.5.26}$$

To simplify the above equation, we rewrite the time derivative of the gravitational potential by using:

$$\dot{\Psi} = \frac{d}{dt}(H\mathcal{I}) = -\ddot{\mathcal{I}}, \tag{6.5.27}$$

where the second equality is obtained by noting the $\mathcal{I}$ satisfies the differential equation:

$$\ddot{\mathcal{I}} + \frac{d}{dt}(H\mathcal{I}) = 0. \tag{6.5.28}$$

Using the above equation again and substituting for the pressure and density perturbations in the matter fields as $\delta\rho_\varphi = -\dot{\rho}_\varphi\mathcal{I}$ and $\delta P_\varphi = -\dot{P}_\varphi\mathcal{I}$, we can write the equation for $\dot{\Delta}_{\mathrm{AD},\varphi}$ as:

$$\begin{aligned}\dot{\Delta}_{\mathrm{AD},\varphi} &= -3H(\delta\rho_\varphi + \delta P_\varphi) - 3(\rho_\varphi + P_\varphi)\ddot{\mathcal{I}} - 3\frac{d}{dt}\left((\rho_\varphi + P_\varphi)H\mathcal{I}\right) \\ &= -3H(\delta\rho_\varphi + \delta P_\varphi + \dot{\rho}_\varphi\mathcal{I} + \dot{P}_\varphi\mathcal{I}).\end{aligned} \tag{6.5.29}$$

To proceed, we assume a direct relationship between the density and pressure for each matter field, as is the case for radiation of pressureless (non-relativistic) matter. This assumption is plausible, since in the decay of a single scalar field (the inflaton) typically no chemical potentials are produced.[16] The relation between density and pressure is provided by the sound speed $c^2_{\varphi_I}$, defined as:

$$c^2_{\varphi_I} = \frac{\dot{P}_{\varphi_I}}{\dot{\rho}_{\varphi_I}} = \frac{dP_{\varphi_I}}{d\rho_{\varphi_I}}. \tag{6.5.30}$$

For all matter fields together, the total sound speed can be written as:

$$c^2_\varphi = \frac{1}{\dot{\rho}_\varphi} \sum_I \rho_{\varphi_I} c^2_{\varphi_I}. \tag{6.5.31}$$

The equation of motion for $\Delta_{\mathrm{AD},\varphi}$ can now be written in the simple form:

$$\dot{\Delta}_{\mathrm{AD},\varphi} = -3H(1 + c^2_\varphi)\Delta_{\mathrm{AD}}. \tag{6.5.32}$$

The solution to this differential equation is given by:

$$\Delta_{\mathrm{AD},\varphi} = \mathcal{C}\ a^{-3(1+c^2_\varphi)}, \tag{6.5.33}$$

for some time-independent constant $\mathcal{C}$. Hence, $\Delta_{\mathrm{AD},\varphi}$ is a decaying mode, which decays at least as a^{-3} (since $c^2_\varphi > 0$). Therefore, we conclude that at late times $\Delta_{\mathrm{AD},\varphi} \to 0$.

6.5.4 Converging towards Adiabicity

Notice that the result $\Delta_{\mathrm{AD},\varphi} \to 0$ does not imply directly that $\delta\rho_\varphi$ approaches $-\dot{\rho_\varphi}\mathcal{I}$, i.e. $\delta\rho_\varphi$ becomes adiabatic. The reason is that the asymptotic behaviour $\Delta_{\mathrm{AD},\varphi} \to 0$ could also be achieved by $\delta\rho_\varphi$ and $\dot{\rho}_\varphi\mathcal{I}$ decaying separately. To show that $\delta\rho_\varphi$ indeed approaches $-\dot{\rho_\varphi}\mathcal{I}$ and hence that the perturbations to the matter fields truely become adiabatic, we introduce the dimensionless measure for the deviation from adiabicity:

$$N \equiv \left| \frac{\Delta_{\mathrm{AD},\varphi}}{\rho_\varphi + P_\varphi} \right|. \tag{6.5.34}$$

To obtain an evolution equation for N, we take the time derivative of $\ln N$, yielding:

$$\frac{d}{dt}\ln N = \frac{\dot{\Delta}_{\mathrm{AD},\varphi}}{\Delta_{\mathrm{AD},\varphi}} - \frac{\dot{\rho} + \dot{P}}{\rho + P} = -3H(1 + c^2_\varphi) - \frac{\dot{\rho}(1 + c^2_\varphi)}{\rho + P} = -\frac{(1 + c^2_\varphi)\hat{Q}_\varphi}{(1 + w_\varphi)\rho_\varphi}, \tag{6.5.35}$$

in the last equality, we used Eq. 6.5.14 and the combined equation of state for the matter fields $w_\varphi \equiv \rho^{-1}\sum_I \rho_{\varphi_I} w_{\varphi_I}$.

At some early time during inflation, the matter perturbation may be far from adiabatic, so that N is of the same order as the fractional density perturbation: $N = \mathcal{O}(\delta_\varphi)$. As time elapses during inflation and the subsequent reheating stage, the transfer of energy from the inflaton towards the matter might cause $\delta\rho_\varphi$ and $\dot{\rho_\varphi}$ large in addition to the energy density ρ_φ in the matter fields, which continuously grows over time. Therefore, at late times in the reheating stage, i.e. when $\rho_\varphi \gg 1$ and $\hat{Q}_\varphi \to 0$, the departure from adiabicity as measured by N will have decayed exponentially. Therefore, as long as modes are outside the horizon, non-adiabatic perturbations will not arise during reheating after single-field inflation.

[16] Note that this assumption is *not* made for the combined inflaton-matter fluid.

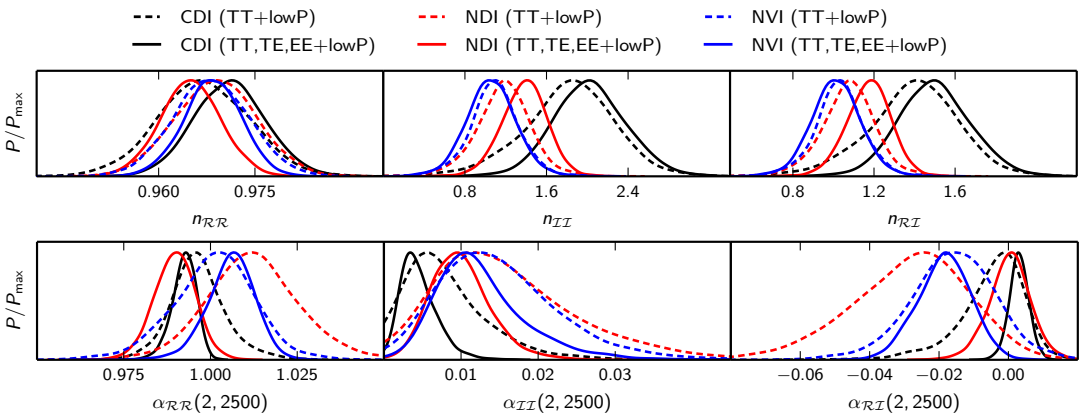


Figure 6.1: Normalized probability density functions $P/P_{\max}$ for n_{ab} and $\alpha_{ab}(2, 2500)$, where $a, b = \mathcal{R}, \mathcal{I}_i$. Mixtures of an adiabatic mode and an isocurvature mode are shown, where the considered isocurvature types are CDI, NDI and NVI. Figure taken from [3].

6.6 Observational Constraints on Adiabicity

In the previous section, we concluded that perturbations generated by single field inflation are adiabatic and will remain so during the stage of reheating, as long as the modes of interest are outside the horizon. Here we will discuss the observational constraints on the adiabicity condition, as found by the Planck collaboration [3] by studying temperature fluctuations in the CMB. In particular, we will constrain a simple model which, in addition to the adiabatic mode (ADI), allows for the presence of an isocurvature mode (ISO). That is, we will test their compatibility with the Planck data. The types of isocurvature modes considered are: (a) the cold dark matter isocurvature mode (CDI), (b) neutrino density isocurvature mode (NDI) and (c) neutrino velocity isocurvature mode (NVI).

A generic mixture of the ADI mode and an ISO mode is described by the individual power spectra of the two modes and the cross-correlation spectrum:

$$\mathcal{P}_{\mathcal{R}\mathcal{R}},\ \mathcal{P}_{\mathcal{I}_i\mathcal{I}_i},\ \mathcal{P}_{\mathcal{I}_i\mathcal{R}}, \tag{6.6.1}$$

where $\mathcal{R}$ and $\mathcal{I}_i$ (i = CDI, NDI, NVI) denote the adiabatic and type i isocurvature modes, respectively. The power spectrum $\mathcal{P}_{ab}$ ($a, b = \mathcal{R}, \mathcal{I}$) is defined at a low and high momentum scale ($k_1 < k_2$) as $\mathcal{P}_{ab}^{(1)} \equiv \mathcal{P}_{ab}(k_1)$ and $\mathcal{P}_{ab}^{(2)} \equiv \mathcal{P}_{ab}(k_2)$. For modes $k \in (k_1, k_2)$, the power spectrum is interpolated as [3]:

$$\mathcal{P}_{ab}(k) = \exp\left[\frac{\ln k - \ln k_2}{\ln k_1 - \ln k_2}\ln \mathcal{P}_{ab}^{(1)} + \frac{\ln k - \ln k_1}{\ln k_2 - \ln k_1}\ln \mathcal{P}_{ab}^{(2)}\right]. \tag{6.6.2}$$

In the analysis of Planck, the momentum limits are set to $k_1 = 0.002\,\mathrm{Mpc}^{-1}$ and $k_2 = 0.1\,\mathrm{Mpc}^{-1}$, so that the whole momentum range examined by Planck is taken into account. The corresponding spectral indices are denoted $n_{\mathcal{R}\mathcal{R}}$, $n_{\mathcal{I}\mathcal{I}}$ and $n_{\mathcal{R}\mathcal{I}} \equiv (n_{\mathcal{R}\mathcal{R}} + n_{\mathcal{I}\mathcal{I}})/2$.

The primordial isocurvature fraction β_{iso} (relative to the adiabitic mode), is defined as:

$$\beta_{\mathrm{iso}}(k) \equiv \frac{\mathcal{P}_{\mathcal{I}\mathcal{I}}(k)}{\mathcal{P}_{\mathcal{R}\mathcal{R}}(k) + \mathcal{P}_{\mathcal{I}\mathcal{I}}(k)}, \tag{6.6.3}$$

Table 6.1: Constraints on (non-)adiabicity in terms of $\beta_{\rm iso}$ and $\alpha_{\text{non-ad}}(2, 2500)$ as obtained by Planck [3] using the TT+lowP dataset. For all values or ranges the CL is 95%. The momentum scales are taken to be $(k_{\rm low}, k_{\rm mid}, k_{\rm high}) = (0.002, 0.05, 0.1)\ \text{Mpc}^{-1}$.

	Constraints on (Non-)Adiabicity			
Model	$100\beta_{\rm iso}(k_{\rm low})$	$100\beta_{\rm iso}(k_{\rm mid})$	$100\beta_{\rm iso}(k_{\rm high})$	$\alpha_{\text{non-ad}}(2, 2500)$
ADI+CDI	4.1	35.4	56.9	$(-1.5\%, 1.9\%)$
ADI+CDI	14.3	22.4	27.4	$(-4.0\%, 1.4\%)$
NVI+NVI	8.3	$[0.1 : 10.2]$	11.9	$(-2.3\%, 2.4\%)$

and will be evaluated at the following low, intermediate and high momentum scales $k = (k_1, 0.05\,\text{Mpc}^{-1}, k_2)$. Furthermore, we define the fractional contribution of the adiabatic $ab = \mathcal{R}\mathcal{R}$, isocurvature ($ab = \mathcal{I}_i\mathcal{I}_i$) and mixed ($ab = \mathcal{R}\mathcal{I}_i$) components to the total CMB temperature variance as:

$$\alpha_{ab}(\ell_{\rm min}, \ell_{\rm max}) \equiv \frac{(\Delta T)^2_{ab}\ (\ell_{\rm min}, \ell_{\rm max})}{(\Delta T)^2_{\rm tot}(\ell_{\rm min}, \ell_{\rm max})}. \tag{6.6.4}$$

In particular, the non-adiabatic contribution to the temperature variance, which can be calculated via:

$$\alpha_{\text{non-ad}}(2, 2500) = 1 - \alpha_{\mathcal{R}\mathcal{R}}(2, 2500), \tag{6.6.5}$$

is of interest. The normalized probability functions $P/P_{\rm max}$ for n_{ab} and $\alpha_{ab}(2, 2500)$ are presented in Fig. 6.1. The 95% CL constraints on $\beta_{\rm iso}$ and $\alpha_{\text{non-ad}}(2, 2500)$ based on Planck data [3] is presented in Tab. 6.1. Notice that for all mixed models (ADI+ISO) the non-adiabatic contribution to the temperature variation is very small, supporting the hypothesis of adiabicity. This statement is statistically supported by a Bayesian model comparison between purely adiabatic and mixed scenario's as discussed in detail in section 11.2 of [3]. To be more quantitative, odds greater than 1:100 are found by Planck for all mixed scenarios compared to the purely adiabatic scenario. That is, the purely adiabatic scenario is strongly favored over the mixed scenarios according to the Planck data.

Chapter 7

Non-Gaussianity and CMB Anisotropies

"It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong."

— Richard P. Feynman

One of the profound predictions of inflation is that density perturbations in the primordial plasma are generated by quantum fluctuations during inflation. In particular, in the preceding chapter we found that (at least in the single-field scenario) the density perturbations are characterized by the following properties:

- ▷ *Primordial Signatures.*—The density perturbations generated during inflation are primordial. That is, they were laid down outside the horizon and entered the horizon after the Big Bang.
- ▷ *Scale-Invariance.*—The Fourier modes of the density perturbations are nearly scale-invariant. The reason is during the 60 e-folds of inflationary expansion, all modes undergo a similar expansion as they are stretched to super-horizon scales. This prediction is most evident when considering the scale-dependence of the power spectrum $\mathcal{P}_{\mathcal{R}}$; we found that the spectral index $n_{\rm s}-1$ is very close to zero, supporting scale invariance.
- ▷ *Gaussianity.*—The density perturbations are approximately Gaussian. In the simplest models of inflation, the scalar inflaton field is freely propagating in the inflationary background to a first approximation.[1] As a consequence, quantum fluctuations in the inflaton can be treated as Gaussian perturbations.

Assuming the temperature fluctuations in the CMB are indeed caused by density perturbations induced during inflation, we can test the above predictions. We already discussed the observational support for scale-invariant perturbations back in section 3.4. However, this is not the case for the last prediction concerning the Gaussianity of density perturbations. This prediction can be tested observationally by measuring the amount of primordial non-gaussianity (NG) contained in the temperature anisotropies in the CMB. However, even if the detected NG is greater than the amount expected from the (standard) single field scenario, those distinctive NG signatures have to potential to guide us in how the standard picture of inflation should be adapted.

To this end, the approach is two-fold. From the observational perspective, the challenge is to extract the primordial NG from the data, since later in the evolution of the universe,

[1] Although we constrain ourselves here to the single-field scenario for simplicity, this is also true for most of the multi-field models.

numerous non-primordial sources may have contributed to the total signal of non-gaussianity. From the theoretical side, the amount of primordial NG generated in different models should be predicted in order compare with observations. The latter task will be the main objective of the remaining part of this thesis.

This chapter will give an introduction to the basic concepts concerned with (primordial) non-gaussianity, leaving the details of the theoretical formalism for subsequent chapters. In the first section, we will briefly reflect on the observational challenges, i.e. we will discuss the different non-primordial sources that can contribute to the NG in the detected CMB signal. In the subsequent sections, we will constrain ourselves to primordial NG and discuss in global terms how primordial NG can possibly be used as an observational window to constrain inflationary models. In this discussion, several concepts will be introduced that are key to the detailed formalism laid down in the next chapters.

7.1 Sources of Non-Gaussianity

As mentioned above, in order to extract the primordial NG signal from CMB data, the contributions due to secondary and late time effects must be filtered out. This is a challenging task both on the experimental as well as the theoretical side. Only if we understand the secondary effects well enough in the theoretical context we will be able to filter out these contributions and reliably extract the primordial signal. Below, we will classify different sources contributing to the NG signal in the CMB data [54].

- ▷ *Primordial Non-Gaussianity.*—The primordial NG in the comoving curvature perturbation $\mathcal{R}$ is generated by quantum fluctuations during inflation (or alternative mechanisms [24]). It is this source of NG that is of interest in constraining inflationary models.
- ▷ *Second-Order Non-Gaussianity.*—We measure NG contributions in $\mathcal{R}$ via temperature fluctuations $\delta T/T$ in the CMB, which are related via the so-called transfer function (see section 7.3). Non-primordial NG effects can be induced via this transfer function.
- ▷ *Secondary Non-Gaussianity.*—Secondary NG refers to contributions induced by late-time effects experienced by the CMB photons as they propagate towards us. Secondary NG, for instance, includes the influence of gravitational lensing.
- ▷ *Foreground Non-Gaussianity.*—These contributions to NG are caused by Galactic and inter-Galactic sources.

Having classified the different contributions to the NG signal in the CMB, we will mainly focus on primordial non-Gaussianity in a theoretical context from now.[2] That is, we assume the non-primordial contributions can be filtered away as to reliably extract the primordial signal.

7.2 Primordial Non-Gaussianity

If the primordial curvature perturbation $\mathcal{R}$ is sourced by non-gaussian contributions, additional statistical information is contained in higher-order correlation functions. In particular,

[2] The only exception will be section 7.3, where we will briefly reflect on how the primordial curvature perturbation is related to temperature anisotropies in the CMB and how we can extract NG from the data.

empirically constraining the 3-point correlation function,

$$\langle \mathcal{R}_{\boldsymbol{k}_1} \mathcal{R}_{\boldsymbol{k}_2} \mathcal{R}_{\boldsymbol{k}_3} \rangle \equiv B_{\mathcal{R}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3), \tag{7.2.1}$$

is relevant for testing NG as it is the lowest order statistic that can discriminate between Gaussian and non-gaussian perturbations. (For purely Gaussian perturbations it vanishes.)

Based on translational invariance of the background, the 3-point correlation function must obey 3-momentum conservation for the three considered $\boldsymbol{k}$-modes. This constraint is enforced using a delta function over the sum of the considered momenta, denoted as $\boldsymbol{K}$, that is:

$$B_{\mathcal{R}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \propto (2\pi)^3 \delta^{(3)}(\boldsymbol{K}), \tag{7.2.2}$$

where the factor $(2\pi)^3$ is included due to our Fourier convention and the total momentum vector is $\boldsymbol{K} \equiv \boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3$. Stated differently, the delta function ensures that the momenta form closed triangles in momentum space. However, notice that no restrictions are imposed on the shapes of the triangles. Therefore, NG is analyzed using different triangular shapes (e.g. equilateral folded or squeezed triangular configurations). In fact, as we will discuss below, analyzing NG for different triangular shapes can discriminate between different (classes of) inflationary models that are degenerate based on their spectral index and tensor-to-scalar ratio.

Apart from the momentum-conserving delta function, the 3-point correlation function does not the depend on the orientation of the momenta (this is a direct consequence of rotational invariance). Therefore, the 3-point function can be written as:

$$B_{\mathcal{R}}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = (2\pi)^3 \delta^{(3)}(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) B_{\mathcal{R}}(k_1, k_2, k_3), \tag{7.2.3}$$

where $B_{\mathcal{R}}(k_1, k_2, k_3)$ is referred to as the bispectrum. However, in the coming chapters, the terms bispectrum and three-point function can be used interchangeably. For different triangular configurations, the bispectrum assumes different forms. In general, the bispectrum is defined in terms of a non-linearity scalar f_{NL} as follows:

$$B_{\mathcal{R}}(k_1, k_2, k_3) \equiv f_{\mathrm{NL}} F(k_1, k_2, k_3). \tag{7.2.4}$$

The parameter f_{NL} is a measure for the deviation from Gaussianity and is the quantity that is constrained experimentally.

Single-field slow roll inflation predicts a small but calculable level of NG, resulting in a small non-linearity parameter f_{NL}, generically proportional to the slow roll parameters.[3] There is an intuitive reason why single-field slow roll models predict a small level of NG. In order to achieve accelerated expansion, the inflaton potential must be very flat (as quantified by $\varepsilon_V, \eta_V \ll 1$). Flatness of the potential can only be achieved by suppressing inflaton (self-)interactions and other sources of non-linearity, thereby minimizing possible sources of large NG and hence leaving only the weak coupling to gravity as a small source of NG. Therefore, detection of a large value for f_{NL} would be able to rule out the single-field scenario, making NG the most stringent test for single-field inflation to date.

On the other hand, even if constraints on NG would rule out the standard single-field scenario, the information contained in the NG would still provide us with strong indications

[3] Here, we restrict ourselves to the slow roll scenario for simplicity. However, as we will show in section 7.5, there is no need to restrict to the slow-roll approximation in order to show that single-field inflation predicts the non-linearity parameter to be small.

on what type of extensions to the standard single-field scenario is preferred by the data. We will now discuss briefly how NG data could provide a clear empirical distinction between different (classes of) extensions to the single-field scenario. Over the last decade, an important theoretical realization was made [2]: different classes of extensions predict a detectable level of NG for different triangular shapes of the 3-momenta. Therefore, measuring significant NG in a specific triangular configuration allows to discriminate between different competing extensions to the single-field scenario, which are often degenerate in terms of the spectral index and tensor-to-scalar ratio.

Below, we list in which triangular configuration different classes of extensions produce a detectable level of NG. The different triangular shapes are also represented schematically in terms of the rescaled momenta $x_2 \equiv k_2/k_1$ and $x_3 \equiv k_3/k_1$ in Fig. 7.1.

- ▷ *Squeezed Triangle (Local)* ($k_1 \ll k_2 \simeq k_3$).—Dominant mode for models with multiple light scalar fields, instead of only one scalar field. This scenario is commonly referred to as multi-field inflation [28]. This case is also referred to as the local mode, as the form of the bispectrum corresponding to this triangular configurations follows directly from a local parametrization of NG in real space (see Eq. 7.2.5).

- ▷ *Equilateral Triangle* ($k_1 = k_2 = k_3$).—This configuration is dominant in models with higher derivative interactions and non-trivial speeds of sound. These models include DBI inflation [33] and Ghost inflation [52].

- ▷ *Folded Triangle* ($k_1 = 2k_2 = 2k_3$).—For models with non-standard initial states (i.e. non Bunch-Davies vacua) generate maximized signals in the folded triangle configuration [50].

- ▷ *Elongated and Isosceles Triangles.*—For the sake of completeness, we also mention elongated ($k_1 = k_2 + k_3$) and isosceles ($k_1 > k_2 = k_3$) configurations, which are intermediate cases of the previous triangular shapes.

The Bispectrum

As we mentioned above, the bispectrum takes on different forms for the different triangular configurations. Here, the form of the bispectrum will be given for the squeezed, equilateral and folded triangles. Those different triangular shapes are also called templates. In particular, the bispectrum in the squeezed limit will be considered in detail, as this form is of significant importance in deriving a powerful consistency relation for single-field inflation (see section 7.5).

Local Template

One of the first phenomenological parametrizations of NG is via the gravitational potential Φ, first introduced by Komatsu and Spergel [55]:

$$\Phi(\boldsymbol{x}) = \Phi_g(\boldsymbol{x}) + f_{\rm NL}^{\rm local} \times \left(\Phi_g(\boldsymbol{x})^2 - \langle \Phi_g(\boldsymbol{x}) \rangle^2\right), \tag{7.2.5}$$

where Φ_g is Gaussian. This parametrization is said to be local, as it is locally defined in real space via the above equation. As the gravitational potential is related to the comoving

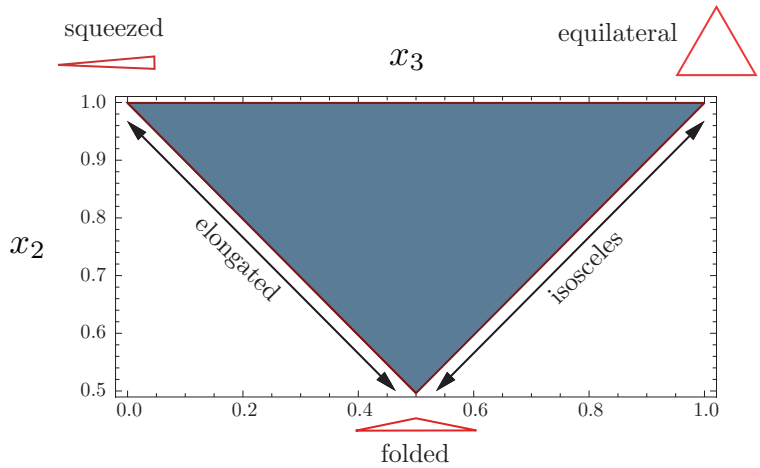


Figure 7.1: Different triangular shapes (i.e. squeezed, equilateral, folded, elongated and isosceles triangles) expressed in terms of the rescaled momenta $x_2 \equiv k_2/k_1$ and $x_3 \equiv k_3/k_1$. Note that the triangle inequality $x_2 + x_3 > 1$ is satisfied. Figure taken from [16].

curvature perturbation $\mathcal{R}$ by a factor of 3/5, the above definition of local NG can be rewritten in terms of $\mathcal{R}$ as follows:

$$\mathcal{R}(\boldsymbol{x}) = \mathcal{R}_g(\boldsymbol{x}) + \frac{3}{5} f_{\rm NL}^{\rm local} \times \left(\mathcal{R}_g(\boldsymbol{x})^2 - \langle \mathcal{R}_g(\boldsymbol{x})\rangle^2\right), \tag{7.2.6}$$

where $P_{\mathcal{R}}(k_1) \equiv |\mathcal{R}_{\boldsymbol{k}_1}|^2$.

From this parametrization, the bispectrum can be derived as (see the derivation box at the end of this section):

$$B_{\mathcal{R}}^{\rm local}(k_1, k_2, k_3) = \frac{6}{5} f_{\rm NL}^{\rm local} \times \Big[P_{\mathcal{R}}(k_1)P_{\mathcal{R}}(k_2) + 2 \text{ perm.}\Big]. \tag{7.2.7}$$

For a scale invariant power spectrum $P_{\mathcal{R}}(k) \equiv A_{\mathcal{R}}/k^3$, where $A_{\mathcal{R}}$ is an amplitude, and the local bispectrum can be rewritten in the following way:

$$B_{\mathcal{R}}^{\rm local}(k_1, k_2, k_3) = \frac{6}{5} f_{\rm NL}^{\rm local} \times A_{\mathcal{R}}^2 \left[\frac{1}{(k_1 k_2)^3} + 2 \text{ perm.}\right]. \tag{7.2.8}$$

In the squeezed limit, defined by $k_1 \to 0$ and $k_2 \simeq k_3$, the local bispectrum can be written as follows:

$$\lim_{k_1 \to 0} B_{\mathcal{R}}^{\rm local}(k_1, k_2, k_3) = \frac{12}{5} f_{\rm NL}^{\rm local} \times P_{\mathcal{R}}(k_1)P_{\mathcal{R}}(k_3). \tag{7.2.9}$$

This form of the local bispectrum corresponds to the squeezed triangular configuration $k_1 \to 0$.

Equilateral and Orthogonal Templates

For the sake of completeness, we will also mention the other two templates which are commonly used: the equilateral and orthogonal configurations. In the equilateral configuration,

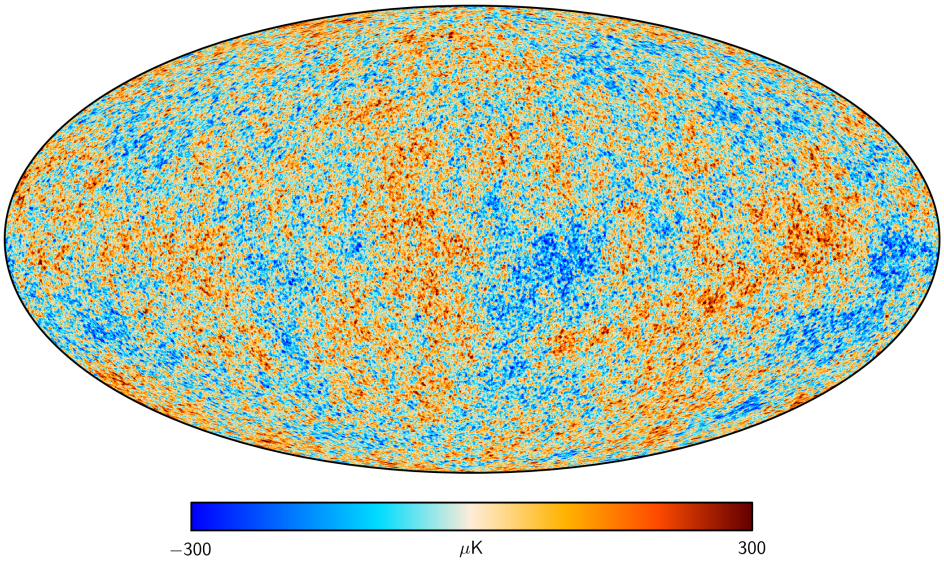


Figure 7.2: CMB temperature map as obtained by the Planck collaboration. The color-coding quantifies temperature fluctuations δT in the CMB around the background temperature of 2.72 K. Figure taken from the ESA website: https://www.esa.int.

the bispectrum becomes:

$$B_{\mathcal{R}}^{\text{equil}} = \frac{18}{5} f_{\text{NL}}^{\text{equil}} \times A_{\mathcal{R}}^2 \left[-\frac{1}{(k_1 k_2)^3} + 2 \text{ perm.} - \frac{2}{(k_1 k_2 k_3)^2} + \left(\frac{1}{k_1 k_2^2 k_3^3} + 5 \text{ perm.} \right) \right]. \tag{7.2.10}$$

The orthogonal template, so-called because it is orthogonal to both the local and equilateral bispectra, is given by:

$$B_{\mathcal{R}}^{\text{ortho}} = \frac{18}{5} f_{\text{NL}}^{\text{ortho}} \times A_{\mathcal{R}}^2 \left[-\frac{3}{(k_1 k_2)^3} + 2 \text{ perm.} - \frac{8}{(k_1 k_2 k_3)^2} \left(\frac{5}{k_1 k_2^2 k_3^3} + 5 \text{ perm.} \right) \right]. \tag{7.2.11}$$

7.3 Extracting Non-Gaussianity from CMB Anisotropies

In this section, we will briefly reflect on how (primordial) NG is extracted from CMB data. That is, we will describe how the non-linearity parameter f_{NL} is extracted from the CMB data. We will focus mainly on the local template of NG by following [55], as it mathematically the simplest and of significant importance for single-field models of inflation (see section 7.5). For details on the other two commonly used templates (i.e. equilateral and orthogonal), we refer to [54].

In essence, the relevant CMB data as obtained by e.g. the Planck collaboration [3] consists of temperature maps of the CMB.[4] Figure 7.2 shows such a temperature map with directional dependent temperature fluctuations $\delta T(\boldsymbol{n})$ around the mean temperature, where unit vector $\boldsymbol{n}$ denotes the direction along the sky.

[4]Polarization maps are also obtained, but will not be considered in this work.

For convenience, the map is decomposed into spherical harmonics as follows:

$$\Theta(\boldsymbol{n}) \equiv \frac{\delta T(\boldsymbol{n})}{\bar{T}} = \sum_{\ell,m} a_{\ell m} Y_{\ell m}, \tag{7.3.1}$$

where $\bar{T} = 2.72$ K denotes the mean temperature of the CMB. The standard two-sphere spherical harmonics are denoted by $Y_{\ell m}$, with $\ell = 0, 1, 2$ corresponding to the monopole, dipole and quadrupole, respectively. In addition, the multipole moments $a_{\ell m}$ can be written as:

$$a_{\ell m} = \int d\Omega \, Y^*_{\ell m}(\boldsymbol{n}) \Theta(\boldsymbol{n}) \tag{7.3.2}$$

Since the multipole moments are intrinsically statistical variables, we cannot make predictions about any individual $a_{\ell m}$, solely about the statistical distribution from which they are drawn. In particular, the variance of $a_{\ell m}$ is denoted is called the angular power spectrum C_ℓ^{TT} and defined as follows:

$$\langle a^*_{\ell m} a_{\ell' m'} \rangle = C_\ell^{TT} \delta_{\ell\ell'} \delta_{mm'}. \tag{7.3.3}$$

The multipole moments can be related directly to the Fourier modes of the comoving curvature perturbation via the relation:

$$a_{\ell m} = 4\pi(-i)^\ell \int \frac{d^3k}{(2\pi)^3} \, \Delta_{T\ell}(k) \, \mathcal{R}_k \, Y_{\ell m}. \tag{7.3.4}$$

Here, the transfer function relating the comoving perturbation to the temperature fluctuation is denoted as $\Delta_{T\ell}$. It is this function that can generate the second-order non-Gaussianity.

In terms of the angular power spectrum C_ℓ^{TT}, the connection to the comoving curvature perturbation is given by:

$$C_\ell^{TT} = \frac{2}{\pi} \int k^2 \, dk \, P_{\mathcal{R}}(k) \, \Delta^2_{T\ell}(k). \tag{7.3.5}$$

This result shows the direct relationship between the power spectrum (i.e. two-point correlation function) and the observed temperature anisotropies. By assuming a background cosmology (e.g. the flat FRW model), one can construct the transfer function and hence perform a so-called deconvolution of C_ℓ^{TT} to obtain the power spectrum $P_{\mathcal{R}}$. In Fig. 7.3, the angular power spectrum as measured by Planck is plotted.

7.3.1 (Non-)Gaussian CMB Anisotropies

If the temperature anisotropy $\delta T(\boldsymbol{x})$ is a Gaussian random field, its probability density function (PDF) is given by:

$$P_g(\delta T) = \frac{1}{(2\pi)^{N_{\rm pic}/2} |\xi|^{1/2}} \exp\left[-\frac{1}{2} \sum_{ij} \delta T_i (\xi)^{-1}_{ij} \delta T_j \right], \tag{7.3.6}$$

where g denotes the fact that the PDF is of the Gaussian form. In addition, $\xi_{ij} \equiv \langle \delta T_i \delta T_j \rangle$ is the covariance matrix or two-point correlation function, $|\xi|$ is the determinant of ξ_{ij} and $N_{\rm pix}$ is the number of pixels on the sky.

In terms of the multipole moments $a_{\ell m}$, obtained via the harmonic expansion of the field $\delta T(\boldsymbol{n})$, the PDF becomes:

$$P(a) = \frac{1}{(2\pi)^{N_{\rm harm}/2} |C|^{1/2}} \exp\left[-\frac{1}{2} \sum_{\ell m} \sum_{\ell' m'} a^*_{\ell m} (C^{-1})_{\ell m, \ell' m'} a_{\ell' m'} \right], \tag{7.3.7}$$

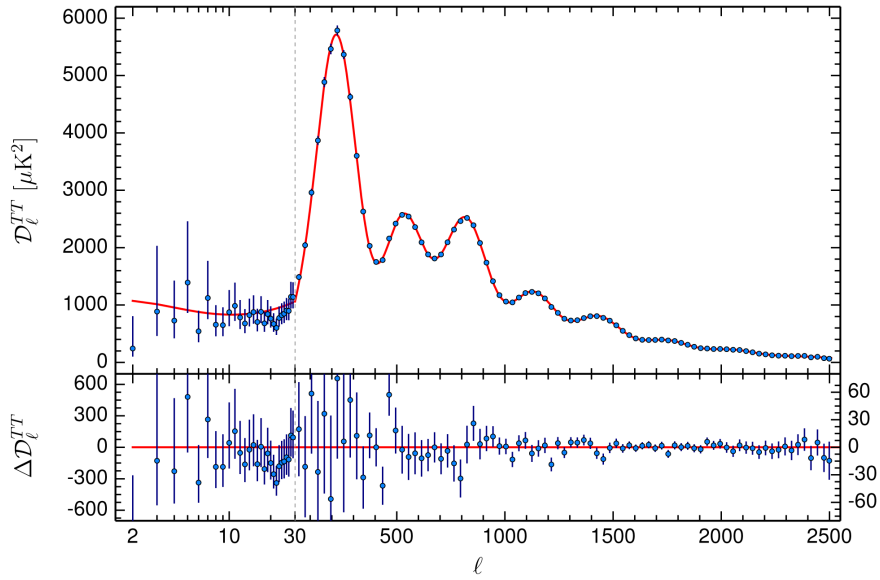


Figure 7.3: Temperature (TT) angular power spectrum of the CMB as measured by Planck. On the vertical axis, the combination $\mathcal{D}_\ell^{TT} \equiv \ell(\ell+1)/2\pi \times C_\ell^{TT}$ is plotted. Figure taken from [3].

where $C_{\ell m,\ell' m} \equiv \langle a^*_{\ell m}, a_{\ell' m'}\rangle$ and $N_{\rm harm}$ is the total number of ℓ and m [54]. Assuming $a_{\ell m}$ is statistically homogeneous and isotropic, we find that $C_{\ell m,\ell' m} = C_\ell \delta_{\ell\ell'}\delta_{mm'}$ and $P(a)$ reduces to the simple form:

$$P(a) = \prod_{\ell m} \frac{e^{-|a_{\ell m}|^2/2C_\ell}}{\sqrt{2\pi C_\ell}}. \tag{7.3.8}$$

As mentioned before, for Guassian anistropies, all statistical information is contained in the two-point correlation function C_ℓ and all higher order correlators vanish.

In case non-gaussianity is contained in the temperature anisotropy field $\delta T(\boldsymbol{n})$, the corresponding PDF is not given by $P_g(a)$. In fact, for NG fluctuations, it is generally not possible to write down the PDF. However, when the NG contribution is small, on may expand the PDF around the Gaussian form as follows [54]:

$$P(a) = \left[1 - \frac{1}{6}\sum_{\text{all } \ell_i m_j} \langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3}\rangle \frac{\partial}{\partial a_{\ell_1 m_1}} \frac{\partial}{\partial a_{\ell_2 m_2}} \frac{\partial}{\partial a_{\ell_3 m_3}}\right] P_g(a). \tag{7.3.9}$$

Note that we have truncated the expansion at the bispectrum, since we assume that higher order contributions are negligible with respect to the two- and three-point functions. Explicitly, the above expression gives:

$$P(a) = P_g(a)\Bigg[1 + \sum_{\text{all } \ell_i m_j} \langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3}\rangle \Big((C^{-1}a)_{\ell_1 m_1}(C^{-1}a)_{\ell_2 m_2}(C^{-1}a)_{\ell_3 m_3} - 3(C^{-1}a)_{\ell_1 m_1,\ell_2 m_2}(C^{-1}a)_{\ell_3 m_3}\Big)\Bigg]. \tag{7.3.10}$$

By means of this result, we can optimally estimate the bispectrum by maximizing the PDF based on the considered CMB data.

7.3.2 Extracting Non-Gaussianity from the CMB Data

In the previous section, we introduced the non-linearity parameter $f_{\rm NL}$ as a measure for NG. In general, we may define $f_{\rm NL}$ as the template-dependent amplitude of the angular bispectrum:

$$\langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle = \mathcal{G}^{m_1 m_2 m_3}_{\ell_1 \ell_2 \ell_3} \sum_{(i)} f^{(i)}_{\rm NL} b^{(i)}_{\ell_1 \ell_2 \ell_3}. \tag{7.3.11}$$

Here, $f^{(i)}_{\rm NL}$ labels the non-linearity parameter for template (i) (e.g. the local, equilateral or orthogonal template), the function $b^{(i)}_{\ell_1 \ell_2 \ell_3}$ is called the reduced bispectrum and encompasses the shape of the bispectrum for the considered template. Finally, $\mathcal{G}$ is the Gaunt integral, defined in following way:

$$\mathcal{G}^{m_1 m_2 m_3}_{\ell_1 \ell_2 \ell_3} \equiv \int d^2\boldsymbol{n}\, Y_{\ell_1 m_1}(\boldsymbol{n}) Y_{\ell_2 m_2}(\boldsymbol{n}) Y_{\ell_3 m_3}(\boldsymbol{n}), \tag{7.3.12}$$

enforcing that $\boldsymbol{\ell}_{1,2,3}$ form a triangle (in analogy with $\boldsymbol{k}_{1,2,3}$ in momentum space).[5]

Given the introduced form for the angular bispectrum, we can maximize $P(a)$ with respect to $f^{(i)}_{\rm NL}$ to obtain the optimal estimator. The equation for the optimal estimator is given by [54]:

$$f^{(i)}_{\rm NL} = \sum_j (F^{-1})_{ij} S_j, \tag{7.3.14}$$

and can be used to determine $f_{\rm NL}$ from the data for a specific template of interest. The quantity S_j is given by:

$$S_i \equiv \frac{1}{6} \sum_{\text{all } \ell_i m_j} \mathcal{G}^{m_1 m_2 m_3}_{\ell_1 \ell_2 \ell_3} b^{(i)}_{\ell_1 \ell_2 \ell_3} \times \Big[(C^{-1}a)_{\ell_1 m_1} (C^{-1}a)_{\ell_2 m_2} (C^{-1}a)_{\ell_3 m_3} - 3 (C^{-1}a)_{\ell_1 m_1, \ell_2 m_2} (C^{-1}a)_{\ell_3 m_3} \Big]. \tag{7.3.15}$$

The factor 1/6 is included so that the RHS of 7.3.14 coincides with the Fisher matrix of $f^{(i)}_{\rm NL}$. Hence, the two-point correlation function of $f^{(i)}_{\rm NL}$ is given by the inverse of F_{ij}, that is:

$$(F^{-1})_{ij} = \langle f^{(i)}_{\rm NL} f^{(j)}_{\rm NL} \rangle - \langle f^{(j)}_{\rm NL} f^{(i)}_{\rm NL} \rangle. \tag{7.3.16}$$

The 1σ uncertainty in $f^{(i)}_{\rm NL} f^{(j)}_{\rm NL}$ is now given by the trace $\Delta f^{(i)}_{\rm NL} \equiv (F^{-1})_{ii}$. Explicitly, the Fisher matrix can be written as [54]:

$$F_{ij} = \frac{f_{\rm sky}}{6} \sum_{\text{all } \ell_i m_j} \sum_{\text{all } \ell'_i m'_j} \mathcal{G}^{m_1 m_2 m_3}_{\ell_1 \ell_2 \ell_3} b^{(i)}_{\ell_1 \ell_2 \ell_3} \times (C^{-1})_{\ell_1 m_1, \ell'_1 m'_1} (C^{-1})_{\ell_2 m_2, \ell'_2 m'_2} (C^{-1})_{\ell_3 m_3, \ell'_3 m'_3} \mathcal{G}^{m'_1 m'_2 m'_3}_{\ell'_1 \ell'_2 \ell'_3} b^{(i)}_{\ell'_1 \ell'_2 \ell'_3}, \tag{7.3.17}$$

where $f_{\rm sky}$ corresponds to the fraction of the sky which is unmasked by point sources (e.g. nearby galaxies). Now we introduced all ingredients needed to compute $f^{(i)}_{\rm NL}$, except for the

[5] In the small angle or flat sky limit [54, 55], the Gaunt integral reduces to the 2D delta function:

$$\mathcal{G}^{m_1 m_2 m_3}_{\ell_1 \ell_2 \ell_3} \to (2\pi)^2\, \delta^{(2)}(\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2 + \boldsymbol{\ell}_3). \tag{7.3.13}$$

reduced bispectrum, which depends on the considered template. Below, we will discuss the reduced bispectrum for the local template, details on the other templates can be found in [54].

7.3.3 Reduced Bispectrum for the Local Template

Komatsu and Spergel [55] have derived the reduced bispectrum for the local template. Recall that the local parametrization of NG is given by:

$$\Phi(\boldsymbol{x}) = \Phi_g(\boldsymbol{x}) + f_{\rm NL}^{\rm local} \times \left(\Phi_g(\boldsymbol{x})^2 - \langle \Phi_g(\boldsymbol{x})^2 \rangle\right), \tag{7.3.18}$$

for the graviational potential. To obtain results in this subsection for the comoving curvature perturbation, one can use that the latter is related to the former via $\mathcal{R} = (3/5)\Phi$. Explicitly, the expression for the reduced bispectrum is given by [55]:

$$b^{\rm local}_{\ell_1\ell_2\ell_3} = 2\int r^2\, dr \Big[\beta_{\ell_1}\beta_{\ell_2}\alpha_{\ell_3} + 2 \text{ perms.}\Big], \tag{7.3.19}$$

where α_ℓ and β_ℓ are functions of the power spectrum and transfer function:

$$\begin{aligned} \alpha_\ell(r) &= \frac{2}{\pi}\int k^2\, dk\, \Delta_{T\ell}(k) j_\ell(kr), \\ \beta_\ell(r) &= \frac{2}{\pi}\int k^2\, dk\, P_\Phi(k) j_\ell(kr). \end{aligned} \tag{7.3.20}$$

This concludes our brief discussion on how NG signals are extracted from the CMB temperature data.

7.4 Komatsu-Spergel Local Bispectrum

In this section, we will explicitly derive the local bispectrum for the comoving curvature perturbation $\mathcal{R}$ as given in Eq. 7.2.8. Recall that the local parametrization of NG in the primordial potential Φ is given by:

$$\Phi(\boldsymbol{x}) = \Phi_g(\boldsymbol{x}) + f_{\rm NL}^{\rm local} \times \left(\Phi_g(\boldsymbol{x})^2 - \langle \Phi_g(\boldsymbol{x})^2 \rangle\right), \tag{7.4.1}$$

as first defined by Komatsu and Spergel [55]. Taking the Fourier transform, the following form in momentum space is found:

$$\begin{aligned} \Phi(\boldsymbol{k}) &= \Phi_g(\boldsymbol{k}) + \Phi_{\rm NL}(\boldsymbol{k}) \\ &= \Phi_g(\boldsymbol{k}) + f_{\rm NL}\left[\int \frac{d^3\boldsymbol{p}}{(2\pi)^3}\Phi_g(\boldsymbol{k}+\boldsymbol{p})\Phi_g^*(\boldsymbol{p}) - (2\pi)^3\delta^{(3)}(\boldsymbol{k})\langle\Phi_g(\boldsymbol{x})^2\rangle\right]. \end{aligned} \tag{7.4.2}$$

Here we used $1 = \int d^3\boldsymbol{k}\ \delta(\boldsymbol{k})$ in front of the $\langle\Phi_g(\boldsymbol{x})^2\rangle$ term. From now on, we will drop the g subscript on the Gaussian part of Φ. That is, Φ is understood to be the Gaussian part and only the non-gaussian contribution will be denoted explicitly as $\Phi_{\rm NL}$. Furthermore, for notational convenience, Fourier components are written as $\Phi(\boldsymbol{k})$ rather than the notation $\Phi_{\boldsymbol{k}}$ used before.

To first order in $\Phi_{\rm NG}$, the 3-point correlation function can now be written as follows:

$$\langle\Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\Phi(\boldsymbol{k}_3)\rangle = \langle\Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\Phi_{\rm NL}(\boldsymbol{k}_3)\rangle + 2 \text{ perm.} \tag{7.4.3}$$

The first term can be evaluated to be (see Appendix F.1):

$$\langle \Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\Phi_{\rm NG}(\boldsymbol{k}_3)\rangle = 2f_{\rm NL}\,(2\pi)^3\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)\,P_\Phi(k_1)P_\Phi(k_2). \tag{7.4.4}$$

Using this result, the Komatsu-Spergel 3-point correlation function can be written as:

$$\langle \Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\Phi(\boldsymbol{k}_3)\rangle = 2f_{\rm NL}\,(2\pi)^3\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)\Big[P_\Phi(k_1)P_\Phi(k_2)+2\text{ perm.}\Big]. \tag{7.4.5}$$

From the above expression, the local bispectrum $B_\Phi^{\rm local}$ can be extracted easily:

$$B_\Phi^{\rm local}(k_1,k_2,k_3) = 2f_{\rm NL}^{\rm local}\times\Big[P_\Phi(k_1)P_\Phi(k_2)+2\text{ perm.}\Big]. \tag{7.4.6}$$

To relate the above results to the comoving curvature perturbation $\mathcal{R}$, we use the fact that the gravitational potential Φ is related to $\mathcal{R}$ via a factor 3/5 and hence the 3-point correlator and local bispectrum for $\mathcal{R}$ are given by:

$$\langle \mathcal{R}(\boldsymbol{k}_1)\mathcal{R}(\boldsymbol{k}_2)\mathcal{R}(\boldsymbol{k}_2)\rangle = \frac{6}{5}f_{\rm NL}^{\rm local}\,(2\pi)^3\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)\Big[P_\mathcal{R}(k_1)P_\mathcal{R}(k_2)+2\text{ perm.}\Big], \tag{7.4.7}$$

$$B_\mathcal{R}^{\rm local}(k_1,k_2,k_3) = \frac{6}{5}f_{\rm NL}^{\rm local}\times\Big[P_\mathcal{R}(k_1)P_\mathcal{R}(k_2)+2\text{ perm.}\Big]. \tag{7.4.8}$$

These results will be used in the next section to derive a powerful consistency relation for single field inflation.

7.5 Single Field Consistency Relation

Here, we will derive a powerful consistency relation assuming single-field inflation, but making *no* other assumptions about the specific inflationary dynamics (e.g. the shape of the potential). The consistency relation implies that the 3-point correlation function for the comoving curvature perturbation $\mathcal{R}$:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle, \tag{7.5.1}$$

is very small in the squeezed triangle limit ($k_1 \ll k_2 \simeq k_3$) as it is suppressed by the factor $1-n_{\rm s}$. In particular, we will derive that in the squeezed limit the NG parameter is given by:

$$\boxed{\lim_{k_1\to 0} f_{\rm NL}^{\rm local} = \frac{5}{12}(1-n_{\rm s}),} \tag{7.5.2}$$

solely assuming that inflation is driven by a single degree of freedom. Assuming the spectral index to be $n_{\rm s}\simeq 0.96$, the above relation gives:

$$\lim_{k_1\to 0} f_{\rm NL}^{\rm local} = \mathcal{O}(10^{-3}) < \mathcal{O}(1), \tag{7.5.3}$$

for all single field inflationary models, independent of their details. Hence, a possible (future) measurement of $f_{\rm NL}^{\rm local} \gtrsim 1$ in the squeezed limit with sufficient accuracy would rule out all single-field inflationary models in a model-independent way. Indeed, the error margin in the measured value for $f_{\rm NL}^{\rm local}$ should be sufficiently small in order to exclude predicted value for single field scenarios (Eq. 7.5.3). In other words, the power of the consistency relation lies in the fact that it can easily be proven wrong by future observations, thereby ruling out all single-field inflationary models. The derivation of the consistency relation presented in this

section is based on the work done in [37, 40]. It should be noted, however, that the direct relationship between the NG parameter and the spectral index (Eq. 7.5.2) was not derived in [37, 40], but is a small extension performed in this work.

The approach to derive the consistency relation will be as follows. First, we will parametrize the background space-time according to the ADM formalism [37], taking the effect of the comoving curvature perturbation into account. Subsequently, the squeezed limit of the momenta-triangle will be introduced geometrically and the main physical mechanism behind the proof will be discussed. Finally, the 3-point correlation function will be derived explicitly in the squeezed limit.

Step 1: Parametrization of the Background

Following [37], the space-time will be parametrized using the ADM formalism, in order to conveniently incorporate the effect of the comoving curvature perturbation on the background.[6] According to the ADM formalism, the perturbed space-time line element can be written in terms of ADM variables N and N^i as:

$$ds^2 = -N^2 dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt), \tag{7.5.4}$$

where the effect of $\mathcal{R}$ is incorporated in a rescaled definition of the spatial metric:

$$h_{ij} \equiv a^2(t) e^{2\mathcal{R}(\boldsymbol{x},t)} \delta_{ij}. \tag{7.5.5}$$

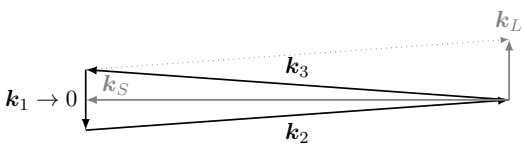
It can be proven that for Super-Horizon modes the ADM variables tend to their unperturbed values, i.e. $N \to 1$ and $N^i \to 0$ (see [37] for details). Hence, in the SH limit the space is described by:

$$ds^2 = -dt^2 + a^2(t) e^{2\mathcal{R}(\boldsymbol{x})} d\boldsymbol{x}^2, \tag{7.5.6}$$

where we omitted the dependence of $\mathcal{R}$ on time, as it is proven to be constant on SH scales.[7] Hence, on SH scales, the comoving curvature perturbation can be absorbed in the line element by a (local) rescaling of the coordinates: $\boldsymbol{x}' = e^{\mathcal{R}(\boldsymbol{x})}\boldsymbol{x}$.

Step 2: The Squeezed Limit and Physical Mechanism

Geometrically, the squeezed limit of the momenta-triangle, taking $\boldsymbol{k}_1 \to 0$, can be visualized using the following schematic:



[6] A detailed introduction to the ADM formalism will be given in chapter 9, here we will solely state the relevant results of formalism and use them. Derivations can be found in the stated chapter on the ADM formalism.

[7] The given proof for the constancy of $\mathcal{R}$ only holds for adiabatic perturbations. However, for the single field scenario, the generated perturbations will be adiabatic (by definition). Therefore, using the result $\dot{\mathcal{R}} \to 0$ on SH does not add any new assumptions to the current proof.

In addition to three momenta $\boldsymbol{k}_i$, two additional momenta are defined in the above schematic. Expressed in terms of $\boldsymbol{k}_i$, they can be written as:

$$\boldsymbol{k}_L \equiv \boldsymbol{k}_2 + \boldsymbol{k}_3, \qquad \boldsymbol{k}_S \equiv (\boldsymbol{k}_3 - \boldsymbol{k}_2)/2. \tag{7.5.7}$$

The labels L and S stand for long and short wavelengths, respectively, since they are related to the magnitude of the momenta as $\lambda = 2\pi/k$. In the squeezed limit, the mode $\mathcal{R}_{\boldsymbol{k}_1}$ leaves the horizon much earlier compared to the other two $\mathcal{R}_{\boldsymbol{k}_2,\boldsymbol{k}_3}$ (recall Fig. 3.2). Therefore, it may be regarded as background field, denoted as $\overline{\mathcal{R}}(\boldsymbol{x})$ in real space, which is simply induces a recaling of the spatial coordinates for the other two modes: $\boldsymbol{x}' = e^{\overline{\mathcal{R}}(\boldsymbol{x})}\boldsymbol{x}$.

Step 3: Computation

Based on the above discussion, the 3-point correlation function can be computed in a two-step process. First, we will compute the 2-point correlation function for $\mathcal{R}_{\boldsymbol{k}_2,\boldsymbol{k}_3}$ in the presence of the background field $\overline{\mathcal{R}}$:

$$\langle \mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle_{\overline{\mathcal{R}}}. \tag{7.5.8}$$

Subsequently, we will correlate the obtained 2-point function with the background field $\overline{\mathcal{R}}$. In mathematical terms, the 3-point function is thus approximated as:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle \simeq \langle \overline{\mathcal{R}}_{\boldsymbol{k}_1}\langle \mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle_{\overline{\mathcal{R}}}\rangle. \tag{7.5.9}$$

It proves convenient to compute the 2-point correlation function in the presence of the background field in real space, i.e. we compute:

$$\langle \mathcal{R}(\boldsymbol{x}_2)\mathcal{R}(\boldsymbol{x}_3)\rangle\Big|_{\overline{\mathcal{R}}}, \tag{7.5.10}$$

rather than its Fourier counterpart. The variation scale of the background field $\overline{\mathcal{R}}$, as set by the wavelength $\overline{\lambda} = 2\pi/k_1$ of the corresponding momentum mode k_1, is much larger than the spatial difference $|\boldsymbol{x}_2 - \boldsymbol{x}_3|$ in the squeezed limit.[8] Hence, we can assume that the background field varies only very slowly and the background field at positions $\boldsymbol{x}_{2,3}$ is well approximated by the value at the midpoint $\boldsymbol{x}_+ \equiv (\boldsymbol{x}_2 + \boldsymbol{x}_3)/2$, denoted as $\overline{\mathcal{R}}_+ \equiv \overline{\mathcal{R}}(\boldsymbol{x}_+)$ (see Fig. 7.4). Under the influence of the background field, the difference $\boldsymbol{x}_2 - \boldsymbol{x}_3$ then gets rescaled as:

$$\boldsymbol{x}_3' - \boldsymbol{x}_2' = e^{\overline{\mathcal{R}}_+}(\boldsymbol{x}_3 - \boldsymbol{x}_2) \simeq \boldsymbol{x}_3 - \boldsymbol{x}_2 + \overline{\mathcal{R}}(\boldsymbol{x}_+)(\boldsymbol{x}_3 - \boldsymbol{x}_2), \tag{7.5.11}$$

where in the last step the exponential is expanded to first order. In terms of the rescaled coordinates $\boldsymbol{x}_{2,3}'$, the two-point correlation function in the presence of the background field can be written conveniently as:

$$\langle \mathcal{R}(\boldsymbol{x}_2)\mathcal{R}(\boldsymbol{x}_3)\rangle\Big|_{\overline{\mathcal{R}}} = \langle \mathcal{R}(\boldsymbol{x}_2')\mathcal{R}(\boldsymbol{x}_3')\rangle. \tag{7.5.12}$$

Using Eq. 7.5.11, the above expression can be expanded to first order in the background field $\mathcal{R}_+$ as [40]:

$$\langle \mathcal{R}(\boldsymbol{x}_2)\mathcal{R}(\boldsymbol{x}_3)\rangle\Big|_{\overline{\mathcal{R}}} = \langle \mathcal{R}(\boldsymbol{x}_2)\mathcal{R}(\boldsymbol{x}_3)\rangle\Big|_0 + \overline{\mathcal{R}}(\boldsymbol{x}_+)\Big[(\boldsymbol{x}_3 - \boldsymbol{x}_2)\cdot\boldsymbol{\nabla}\langle \mathcal{R}(\boldsymbol{x}_2)\mathcal{R}(\boldsymbol{x}_3)\rangle\Big]\Big|_0. \tag{7.5.13}$$

[8] To be more precise, in the limit $k_1 \to 0$, the variation length tends to infinity $\overline{\lambda} \to \infty$.

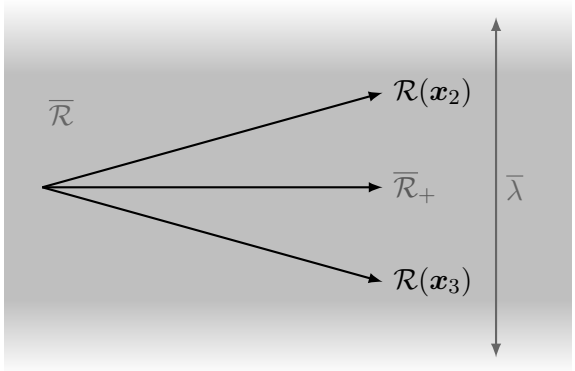


Figure 7.4: Schematic describing the evaluation of the background field $\mathcal{R}_+$ midway between $\mathcal{R}(\boldsymbol{x}_2)$ and $\mathcal{R}(\boldsymbol{x}_3)$. To good approximation, the background field is the same at $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ as the difference $|\boldsymbol{x}_3 - \boldsymbol{x}_2|$ is much smaller than the variation scale $\overline{\lambda}$ of the background field.

To get the 3-point correlation function (in real space), we correlate the above result with the background field $\overline{\mathcal{R}}(\boldsymbol{x}_1)$ and average over it, yielding:

$$\begin{aligned}\langle\overline{\mathcal{R}}(\boldsymbol{x}_1)\mathcal{R}(\boldsymbol{x}_2)\mathcal{R}(\boldsymbol{x}_3)\rangle &= \langle\overline{\mathcal{R}}(\boldsymbol{x}_1)\rangle\langle\mathcal{R}(\boldsymbol{x}_2)\mathcal{R}(\boldsymbol{x}_3)\rangle\Big|_0 \\ &\quad + \langle\overline{\mathcal{R}}(\boldsymbol{x}_1)\overline{\mathcal{R}}(\boldsymbol{x}_+)\rangle\Big[(\boldsymbol{x}_3-\boldsymbol{x}_2)\cdot\boldsymbol{\nabla}\langle\mathcal{R}(\boldsymbol{x}_2)\mathcal{R}(\boldsymbol{x}_3)\rangle\Big]\Big|_0. \end{aligned} \tag{7.5.14}$$

The first term is proportional to the mean or 1-point correlation function $\langle\overline{\mathcal{R}}(\boldsymbol{x}_1)\rangle$ and therefore is assumed to vanish. To proceed, we will perform a Fourier transform to find the corresponding expression for the 3-point correlation spectrum in momentum space. The momentum-space analog is given by:

$$\boxed{\langle\mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle \simeq -(2\pi)^3\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)P(k_1)P(k_2)\frac{d\ln k_2^3P(k_2)}{d\ln k_2}.} \tag{7.5.15}$$

This result will be derived explicitly in the Appendix F.2.

Finally, since $\mathcal{P}_{\mathcal{R}}(k_2) \propto k_2^3P(k_2)$, the derivative term is equal to the spectral index (by definition, see Eq. 5.4.33) and the 3-point correlation function in the squeezed limit becomes:

$$\boxed{\langle\mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle \simeq (2\pi)^3\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)(1-n_{\rm s})P_{\mathcal{R}}(k_1)P_{\mathcal{R}}(k_2).} \tag{7.5.16}$$

Comparing the above result to Eq. 7.4.7 with $k_3 \to 0$ we identify:

$$f_{\rm NL}^{\rm local} = \frac{5}{12}(1-n_{\rm s}). \tag{7.5.17}$$

Therefore, we derived the result that (in the squeezed local template) $f_{\rm NL}$ is suppressed by spectral index and consequently any substantial detection of $f_{\rm NL}$ in this configuration would rule out all single-field inflationary models, independent of their details.

Part IV

Non-Gaussianity in the Single-Field Scenario

Chapter 8

In-In Formalism of Quantum Field Theory

"Quantum field theory arose out of our need to describe the ephemeral nature of life."

— Anthony Zee

In order to compute the non-Gaussianity generated in specific models of (single-field) inflation, we have to evolve the comoving curvature perturbation (sourced by inflaton fluctuations) to super-horizon scales and evaluate quantum expectation values for higher order correlation functions. In particular, we are interested in the expectation value for the three-point correlation function or bispectrum:

$$\langle \mathcal{R}_{\boldsymbol{k}_1} \mathcal{R}_{\boldsymbol{k}_2} \mathcal{R}_{\boldsymbol{k}_3} \rangle,$$

since this is the lowest-order statistic that can discriminate between Gaussian and non-Gaussian statistics.

In this chapter, we will develop the theoretical formalism that we will use to evaluate the expectation value of such a correlation function. In standard quantum field theory (e.g. particle physics), computing correlation functions is done in the so-called In-Out formalism. The central object in this formalism is the scattering matrix S, which quantifies the probability that a state $|\text{in}\rangle$ at the far past becomes some state $|\text{out}\rangle$ in the far future:

$$\langle \text{out}(t \to +\infty) | \text{in}(t \to -\infty) \rangle \equiv \langle \text{out} | S | \text{in} \rangle,$$

during a scattering process. Notice that in this formalism, asymptotic conditions are imposed very early and very late times, since Minkowski space states are assumed to be non-interacting in far past and future, where the particles are far away from the interaction region. Often, the asymptotic states are taken to be the vacuum states of the free Hamiltonian of the system. The free Hamiltonian treats the considered fields as free fields. Notice that the expectation values do not possess time-dependence in the In-Out formalism.

However, in cosmology, we are interested in the expectation values of operators at a specific instant in time.[1] Furthermore, contrary to Minkowski space-time, in cosmology the background space-time (FRW metric) is time-dependent via the scale factor. Hence, defining time-dependence is less straightforward (we already saw this in Chapter 5 in constructing the vacuum state). Finally, boundary conditions are only imposed at very early times, when modes of cosmological interest are far inside the horizon. In this limit, it is convenient to impose boundary conditions since the space-time approaches the Minkowski limit.

[1] For single-field inflation, a suitable instant to evaluate the three-point function is after horizon exit, where the comoving curvature perturbation has frozen.

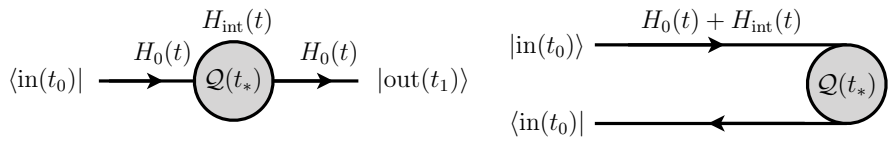


Figure 8.1: Left: Visualization of the In-Out formalism to compute quantum expectation values (correlators). The $|{\rm in}\rangle$ state defined at initial time t_0 is evolved with the free-field Hamiltonian to the moment of interest t_* where interactions are included via the interaction Hamiltonian $H_{\rm int}$ and the state is evolved to the far future. Boundary conditions are thus imposed both in the early and late time limits and one computes transition probabilities. Right: Pictorial representation of the In-In formalism, where the initial state is evolved towards the moment of interest and then evolved back to the initial time. Asymptotic boundary conditions are thus imposed only at early times.

Based on the above observations, we conclude that we require an alternative to the In-Out formalism, which is suitable for cosmological applications. In this chapter, we will construct this alternative formalism, known as the In-In formalism. In this formalism, boundary conditions are imposed only at very early times, i.e. the $|{\rm in}\rangle$ state.[2] Contrary to the In-Out formalism, the In-In formalism allows to evaluate expectation values at a specific instant in time (Fig. 8.1).

8.1 Preview of the In-In Formalism

For readers solely interested in the concrete results of the In-In formalism, we will first give the relevant results here and support them with detailed derivations in the remaining part of this chapter. The aim is to compute the n-point correlation function for a string of operators, compactly denoted as $\mathcal{Q} \equiv \psi_{\boldsymbol{k}_1}\psi_{\boldsymbol{k}_2}\cdots\psi_{\boldsymbol{k}_n}$. The considered field ψ will be left unspecified here but can be taken as for instance the comoving curvature perturbation $\mathcal{R}$ or gravitational wave polarization modes $E^{+,\times}$. Schematically, the correlation function evaluated at time t_* can then be written as:

$$\langle \mathcal{Q}(t_*)\rangle = \langle \Omega|\mathcal{Q}(t_*)|\Omega\rangle, \tag{8.1.1}$$

where $|\Omega\rangle$ is the vacuum state of the interacting theory at the far past (typically denoted as $t_) \to -\infty$). In general, constructing the interaction vacuum state is very difficult. Fortunately, we will see that in case interactions are weak, we can split the total Hamiltonian into a free part and interaction part and we replace the interaction vacuum by the free vacuum $|0\rangle$.

In order to determine $\langle \mathcal{Q}(t)\rangle$, we evolve the fields from $t_0 \to -\infty$ towards the moment of interest $t = t_*$ and then back to t using the total Hamiltonian (see Fig. 8.1). Computing the time evolution of the fields is not straightforward, because the total Hamiltonian includes interactions:

$$H = H_0 + H_{\rm int}, \tag{8.1.2}$$

resulting in non-linear equations of motion. To conveniently handle these complications, we introduce the interaction picture, where the leading evolution of the fields is determined by the free field Hamiltonian H_0. The free Hamiltonian is quadratic in the fields and therefore the equations of motion will be linear. Corrections due to interactions are treated perturbatively.

[2] In fact, this is the reason why the formalism goes under the name In-In.

Taking the above approach, we will show that the time-dependent expectation value $\langle \mathcal{Q}(t_*)\rangle$ is governed by the following equation:

$$\langle \mathcal{Q}(t_*)\rangle = \langle \left[\bar{T}\exp\left(+i\int_{-\infty^-}^{t_*} dt'\, H^I_{\rm int}(t')\right)\right]\mathcal{Q}^I(t_*)\left[T\exp\left(-i\int_{-\infty^+}^{t_*} dt'\, H^I_{\rm int}(t')\right)\right]\rangle, \quad (8.1.3)$$

where the angular brackets denote the vacuum state of the free field theory $|0\rangle$ and all fields are in the interaction picture, indicated using the superscript I. In order to regularize the time integral in the early time limit, we introduced the $i\epsilon$ prescription by defining:

$$-\infty^{\mp} \equiv -\infty(1 \pm i\epsilon). \quad (8.1.4)$$

As we will discuss in detail, this prescription assures that the integrand vanishes at early times and hence the integrals remain finite. The above equation for $\langle \mathcal{Q}(t_*)\rangle$ constitutes the main result of this chapter and will be derived in the subsequent sections.

8.2 Qauntum-Classical Split of the Hamiltonian

Let us consider the action for a generic field ψ, defined in terms of the Lagrangian (density) $\mathcal{L}(\psi, \dot{\psi})$ as follows:

$$S = \int d^4x\, \mathcal{L}(\psi, \dot{\psi}), \quad (8.2.1)$$

where the field $\psi = \psi(\boldsymbol{x}, t)$ is allowed to possess temporal as well as spatial dependence. The conjugate momentum of ψ is defined as:

$$\pi \equiv \frac{\partial \mathcal{L}}{\partial \dot{\psi}}. \quad (8.2.2)$$

For the sake of simplicity, we consider only one field, but the analysis below can easily be extended to multiple fields.

The total Hamiltonian of the system is obtained via the Lagrangian as follows:

$$H[\psi(t), \pi(t)] \equiv \int d^3x\, \mathcal{H}(\psi(\boldsymbol{x}, t), \dot{\psi}(\boldsymbol{x}, t)) = \int d^3x(\dot{\psi}\pi - \mathcal{L}). \quad (8.2.3)$$

Note that in the Hamiltonian the spatial-dependence of the fields is integrated out, while the Hamiltonian density still depends on spatial coordinates. According to the Hamilton equations, the time-evolution of the field and its conjugate momentum is generated by the full Hamiltonian via the commutator forms:

$$\dot{\psi}(\boldsymbol{x}, t) = i\left[H[\psi(t), \pi(t)], \psi\right], \quad (8.2.4)$$

$$\dot{\pi}(\boldsymbol{x}, t) = i\left[H[\psi(t), \pi(t)], \pi\right]. \quad (8.2.5)$$

To proceed, we split the field and its conjugate momentum into a classical background field and a quantum perturbation:

$$\psi(\boldsymbol{x}, t) = \bar{\psi}(\boldsymbol{x}, t) + \delta\psi(\boldsymbol{x}, t), \qquad \pi(\boldsymbol{x}, t) = \bar{\pi}(\boldsymbol{x}, t) + \delta\pi(\boldsymbol{x}, t). \quad (8.2.6)$$

We should note that this split is completely arbitrary.[3] The theory thus describes quantized perturbations on a classical time-dependent background field. Throughout, we will assume

[3] Although it was not made explicit at that point, in Chapter 5 we also performed this split, since we quantized the fluctuation in the inflaton field but took the background field as classical.

that the background fields are in their vacuum state, which may well be time-dependent. The classical background fields obey the Hamilton equations:

$$\dot{\bar{\psi}}(\boldsymbol{x},t) = \frac{\partial \mathcal{H}}{\partial \bar{\pi}}, \qquad \dot{\bar{\pi}}(\boldsymbol{x},t) = -\frac{\partial \mathcal{H}}{\partial \bar{\psi}}. \tag{8.2.7}$$

The perturbations to the field and its conjugate momentum are treated as quantum entities and are therefore quantized by imposing the canonical commutation relations:

$$[\delta\psi(\boldsymbol{x},t), \delta\pi(\boldsymbol{y},t)] = i\delta^{(3)}(\boldsymbol{x}-\boldsymbol{y}), \tag{8.2.8}$$

all other commutators vanish identically.

The Hamiltonian can be decomposed in accordance with the quantum-classical split of the field. In particular, we Taylor expand around the background Hamiltonian H_{bkg}, constituted solely by the background fields:

$$\begin{aligned} H[\psi(t),\pi(t);t] &= H_{\text{bkg}}[\bar{\psi}(t),\bar{\pi}(t)] \\ &\quad + \int d^3x\, \frac{\partial \mathcal{H}}{\partial \bar{\psi}} \delta\psi(\boldsymbol{x},t) + \int d^3x\, \frac{\partial \mathcal{H}}{\partial \bar{\pi}} \delta\pi(\boldsymbol{x},t) \\ &\quad + \tilde{H}[\delta\psi(t),\delta\pi(t);t], \end{aligned} \tag{8.2.9}$$

where the total Hamiltonian possesses explicit time-dependence, in addition to the implicit temporal dependence of the fields, due to the background fields. The first line corresponds to the background Hamiltonian and the second line contains the contributions linear in perturbations. Finally, $\tilde{H}$ encompasses the contribution of terms of second and higher order in the perturbations to the total Hamiltonian, that is:

$$\tilde{H} = \mathcal{O}(\delta\psi^2, \delta\pi^2). \tag{8.2.10}$$

We will now show that the terms in expanded Hamiltonian linear in perturbations have no effect on the evolutions of neither the background fields nor the perturbations. Substituting the quantum-classical split of the field (Eq. 8.2.6) and the expanded Hamiltonian (Eq. 8.2.9) into the commutator form of the equation of motion is given by:

$$\dot{\bar{\psi}}(\boldsymbol{x},t) + \delta\dot{\psi}(\boldsymbol{x},t) = i\left[H[\bar{\psi},\bar{\pi}] + \int d^3y \left(\frac{\partial \mathcal{H}}{\partial \bar{\psi}}\delta\psi + \frac{\partial \mathcal{H}}{\partial \bar{\pi}}\delta\pi\right) + \tilde{H}, \bar{\psi} + \delta\psi\right]. \tag{8.2.11}$$

On account of the fact that the background field $\bar{\psi}$ is classical, it commutes with the total Hamiltonian. In addition, the field perturbation $\delta\psi$ commutes with the contribution to the Hamiltonian that is linear in $\delta\psi$ (first term on second line of Eq. 8.2.6). The only non-vanishing commutator is the one between the Hamiltonian term proportional to $\delta\pi$ and $\delta\psi$, which can be manipulated as follows:

$$\begin{aligned} i\left[\int d^3y\, \frac{\partial \mathcal{H}}{\partial \bar{\pi}}\delta\pi, \delta\psi\right] &= i\int d^3y\, \dot{\bar{\psi}}(\boldsymbol{y},t)\,[\delta\pi(\boldsymbol{y},t), \delta\psi(\boldsymbol{x},t)] \\ &= \int d^3y\, \dot{\bar{\psi}}(\boldsymbol{y},t) \times \delta^{(3)}(\boldsymbol{x}-\boldsymbol{y}) = \dot{\bar{\psi}}(\boldsymbol{x},t). \end{aligned} \tag{8.2.12}$$

Notice that this term cancels the term $\dot{\bar{\psi}}(\boldsymbol{x},t)$ on the right hand side of 8.2.11 and therefore we conclude that the perturbative corrections to the background evolution of the field start

only at second order in perturbations. Now we can write down the Hamilton equations for the quantum perturbations:

$$\begin{aligned}\delta\dot{\psi}(\boldsymbol{x},t) &= i\left[\tilde{H}[\delta\psi(t),\delta\pi(t);t],\delta\psi\right],\\ \delta\dot{\pi}(\boldsymbol{x},t) &= i\left[\tilde{H}[\delta\psi(t),\delta\pi(t);t],\delta\pi\right].\end{aligned} \tag{8.2.13}$$

Hence, we derived that the perturbations are evolved using the second order perturbed Hamiltonian, which has explicit time dependence due to the background fields.

8.3 Evolution Operators and Interaction Picture

The next task will be to find the solutions to the equations of motion of the perturbations (Eqs. 8.2.13). We will follow the standard procedure in quantum field theory and write the solutions in terms of a unitary operator U as:

$$\delta\psi(\boldsymbol{x},t) = U^{-1}(t,t_0)\delta\psi(\boldsymbol{x},t_0)U(t,t_0). \tag{8.3.1}$$

Unitarity of U implies that the operator satisfies the following properties:

$$U(t,t) = 1, \qquad U^{-1}(t,t_0)U(t,t_0) = 1, \qquad U(t_3,t_2)U(t_2,t_1) = U(t_3,t_1), \tag{8.3.2}$$

where $t_3 > t_2 > t_1$. The operator U acts as an evolution operator, which evolves the initial value $\delta\psi(t_0)$ defined at some time t_0 to arbitrary later time t. Substituting the above ansatz into the equations of motion, we find that U is governed by the following differential equation:

$$\frac{d}{dt}U(t,t_0) = -i\tilde{H}[\delta\psi(t_0),\delta\pi(t_0);t]\;U(t,t_0). \tag{8.3.3}$$

The general solution to this equation is given by:[4]

$$U(t,t_0) = T\exp\left[-i\int_{t_0}^{t} dt'\;\tilde{H}(t')\right], \tag{8.3.5}$$

which in principle accounts for the evolution of the system.[5] However, the above solution not very convenient, since it is not straightforward to compute $\delta\psi(\boldsymbol{x},t)$ in the presence of interactions (as the equations of motion would be non-linear).

In order to describe the evolution of perturbations, taking interactions into account, we split the perturbed Hamiltonian $\tilde{H}$ into a free-field Hamiltonian H_0 (consisting only of second order perturbations), and the interaction Hamiltonian H_{int} as follows:

$$\tilde{H} = H_0 + H_{\text{int}}. \tag{8.3.6}$$

[4]In case it is convenient to make the Hamiltonian dependence manifest, the unitary operator is often represented by the following shorthand notation:

$$U(t,t_0) \equiv e^{-i\tilde{H}(t-t_0)}. \tag{8.3.4}$$

[5]The time-ordening operator T is included since we are typically dealing with time-dependent Hamiltonians.

The fields generated by the free-field Hamiltonian H_0 are called interaction picture fields, labeled by I. Interaction picture fields are defined to coincide with the fields of the complete theory[6] at some initial time t_0, that is:

$$\delta\psi^I(\boldsymbol{x}, t_0) \equiv \delta\psi(\boldsymbol{x}, t_0), \qquad \delta\pi^I(\boldsymbol{x}, t_0) \equiv \delta\pi(\boldsymbol{x}, t_0). \tag{8.3.7}$$

We choose the interaction picture fields such that we can construct the interaction Hamiltonian in terms of them. To summarize, the prescription is as follows: the leading dynamical evolution of the perturbations is encompassed in the interaction picture fields, governed by the free Hamiltonian H_0. Sub-dominant interaction effects are included by means of a series expansion of the interaction Hamiltonian $H_{\rm int}$.

By construction, the time-evolution of the interaction picture fields is generated via the free-field Hamiltonian:

$$\delta\dot{\psi}^I = i\left[H_0[\delta\psi^I(t), \delta\pi^I(t); t], \delta\psi^I(t)\right], \qquad \delta\dot{\pi}^I = i\left[H_0[\delta\psi^I(t), \delta\pi^I(t); t], \delta\pi^I(t)\right]. \tag{8.3.8}$$

From now on, we will often omit the field-dependence of the Hamiltonian and use a superscript I to indicate the Hamiltonian is composed of interaction picture fields. Since H_0 obviously commutes with itself, we can evolve H_0 to any moment in time inside the commutator. We take advantage of this fact by evolving H_0 to time t_0, so the interaction picture fields can be replaced by the fields in the complete theory:

$$H_0[\delta\psi^I(t), \delta\pi^I(t); t] = H_0[\delta\psi(t_0), \delta\pi(t_0); t], \tag{8.3.9}$$

note that explicit time-dependence of the Hamiltonian due to background fields is not affected by this procedure.

Now we can find the unitary evolution operators U_0 and F, associated with the free-field and interaction Hamiltonian, respectively. For the free-field part, the solution is obtained from:

$$\frac{d}{dt}U_0(t, t_0) = -iH_0[\delta\psi(t_0), \delta\pi(t_0); t]\; U_0(t, t_0), \tag{8.3.10}$$

and reads:

$$U_0(t, t_0) = T\exp\left[-i\int_{t_0}^{t} dt'\; H_0^I(t')\right]. \tag{8.3.11}$$

Hence, we can use the evolution operator U_0 to evolve the interaction picture fields, which are indeed calculated via the free-field Hamiltonian.

Next, we aim to solve for the evolution operator F associated with the interaction Hamiltonian. We want to define F such that it satisfies the following differential equation:

$$\frac{d}{dt}F(t, t_0) = -iH_{\rm int}^I(t)F(t, t_0), \tag{8.3.12}$$

note that the interaction Hamiltonian is in the interaction picture. However, an arbitrary operator (such as the Hamiltonian):

$$\mathcal{Q} = \mathcal{Q}[\delta\psi(\boldsymbol{x}, t), \delta\pi(\boldsymbol{x}, t)], \tag{8.3.13}$$

[6]For inflation, the full theory at early times corresponds to the Bunch-Davies vacuum and the associated initial conditions. The interaction picture fields would correspond to the linear solutions to Mukhanov-Sasaki equation (Eq. 5.2.14).

is not a priori constructed in the interaction picture.

In general terms, the evolution of such an operator $\mathcal{Q}$ is governed by the total Hamiltonian $\tilde{H}$ and the corresponding evolution operator U, yielding:

$$\mathcal{Q}(t) = U^{-1}(t, t_0)\mathcal{Q}(t_0)U(t, t_0). \tag{8.3.14}$$

In order to recast $\mathcal{Q}$ in the interaction picture, we should define U such that the operator is squeezed between the free-field evolution operator U_0, which takes the operator to the interaction picture since we know that:

$$\mathcal{Q}^I(t) \equiv U_0^{-1}(t, t_0)\mathcal{Q}(t_0)U_0(t, t_0). \tag{8.3.15}$$

This is achieved by choosing the operator F as follows:

$$F(t, t_0) \equiv U_0^{-1}(t, t_0)U(t, t_0), \tag{8.3.16}$$

since then we can write $\mathcal{Q}(t)$ as follows:

$$\begin{aligned}\mathcal{Q}(t) &= U^{-1}(t, t_0)\mathcal{Q}(t_0)U(t, t_0),\\ &= F^{-1}(t, t_0)\ U_0^{-1}(t, t_0)\mathcal{Q}(t_0)U_0(t, t_0)\ F(t, t_0)\\ &= F^{-1}(t, t_0)\mathcal{Q}^I(t_0)F(t, t_0).\end{aligned} \tag{8.3.17}$$

Therefore, the expectation value of operator $\mathcal{Q}$ evaluated at the moment of interest t_* can be recasted in the interaction picture via the relation:

$$\langle \mathcal{Q}(t_*)\rangle = \langle \Omega | F^{-1}(t_*, t_0)\mathcal{Q}^I(t_0)F(t_*, t_0)|\Omega\rangle. \tag{8.3.18}$$

The last task is to show explicitly that the parametrization of F indeed satisfies Eq. 8.3.12. Taking the time-derivative of F yields:

$$\begin{aligned}\frac{d}{dt}F(t, t_0) &= -U_0^{-2}\dot{U}_0 U + U_0^{-1}\dot{U}\\ &= -U_0^{-1}(-iH_0)U - U_0^{-1}(i\tilde{H})U\\ &= -iU_0^{-1}(\tilde{H} - H_0)U\\ &= -iU_0^{-1}H_{\text{int}}[\delta\psi(t_0), \delta\pi(t_0); t](U_0U_0^{-1})U\\ &= -iH^I_{\text{int}}[\delta\psi^I(t_0), \delta\pi^I(t_0); t]F(t, t_0),\end{aligned} \tag{8.3.19}$$

which is indeed equivalent to Eq. 8.3.12. In the second line we used the evolution equations U_0 and U. Then, in the third line we used the fact that $H_{\text{int}} = \tilde{H} - H_0$. Finally, we insert a conveniently chosen unity $U_0U_0^{-1} = 1$ to rewrite the interaction Hamiltonian in the interaction picture and we obtain the desired result. The solution to Eq. 8.3.12, in terms of H^I_{int}, reads:

$$F(t, t_0) = T\exp\left[-i\int_{t_0}^{t} dt'\ H^I_{\text{int}}(t')\right]. \tag{8.3.20}$$

On account of unitarity, the inverse operator is given by:

$$F^{-1}(t, t_0) = F^\dagger = \bar{T}\exp\left[+i\int_{t_0}^{t} dt'\ H^I_{\text{int}}(t')\right]. \tag{8.3.21}$$

The expectation value $\langle \mathcal{Q}(t_*)\rangle$ can now be written as follows:

$$\langle \mathcal{Q}(t_*)\rangle = \langle\Omega|\left[\bar{T}\exp\left(i\int_{t_0}^{t_*} dt'\; H^I_{\text{int}}(t')\right)\right]\mathcal{Q}^I(t_*)\left[T\exp\left(-i\int_{t_0}^{t_*} dt'\; H^I_{\text{int}}(t')\right)\right]|\Omega\rangle. \quad (8.3.22)$$

The above expression is almost equivalent to the desired form (Eq. 8.1.3), the only operation that is still to be justified is the replacement of the interaction vacuum $|\Omega\rangle$ by the free-field vacuum $|0\rangle$.

8.4 Relating the Interaction and Free-Field Vacua

The fact that we consider interactions via the interaction Hamiltonian H_{int} around the free-field evolution determined by H_0 introduces a second vacuum, the interaction vacuum $|\Omega\rangle$, in addition to the free-field vacuum associated with H_0 (the latter is discussed in section 5.4 as well). Constructing the free-field vacuum is trivial by means of the creation and annihilation operators arising after quantization, in contrast to the interaction vacuum, which is not straightforward to construct.

Fortunately, it is possible to find a direct relationship between the interaction vacuum and the free-field vacuum in the early time limit.[7] This relation is particularly convenient when considering the evaluation of expectation values $\langle \mathcal{Q}(t_*)\rangle$, as the interaction picture fields will be written in terms of creation and annihilation operators, which act in a very well-defined way on the free-field vacuum state $|0\rangle$. Here, we will explicitly construct the interaction vacuum in terms of the free-field vacuum.

To start, we expand the free-field Hamiltonian in terms of a complete set of energy eigenstates $|n\rangle$ of the complete theory (governed by $\tilde{H}$) as follows:

$$|0\rangle = \sum_n |n\rangle\langle n|0\rangle. \quad (8.4.1)$$

We can evolve the free-field vacuum state from some initial time t_0 to some later time t by applying the evolution operator of the total Hamiltonian:

$$U(t,t_0)|0\rangle = T\exp\left[-i\int_{t_0}^{t} dt'\; \tilde{H}(t')\right]|0\rangle \equiv e^{-i\tilde{H}(t-t_0)}|0\rangle, \quad (8.4.2)$$

where in the last expression we introduced the shorthand notation for the evolution operator. In terms of the set of energy eigenstates of the full theory we obtain:

$$\begin{aligned} U(t,t_0)|0\rangle = e^{-i\tilde{H}(t-t_0)}|0\rangle &= e^{-i\tilde{H}(t-t_0)}\sum_n |n\rangle\langle n|0\rangle \\ &= \sum_n e^{-iE_n(t-t_0)}|n\rangle\langle n|0\rangle \\ &= e^{-iE_\Omega(t-t_0)}|\Omega\rangle\langle\Omega|0\rangle + \sum_{n\neq\Omega} e^{-iE_n(t-t_0)}|n\rangle\langle n|0\rangle. \end{aligned} \quad (8.4.3)$$

By construction, we will set the vacuum energy of the free-field theory to zero, i.e. $H_0|0\rangle = 0|0\rangle$. Since we are comparing the vacuum energy of the free and interaction theory (and we

[7] As mentioned before, this early time limit corresponds the beginning of inflation, when all modes of interest are deep inside the horizon.

set the vacuum energy of the free theory to zero), the vacuum energy of the interacting theory will presumably be small but non-vanishing. Therefore, the vacuum energy of the total Hamiltonian can be computed as:

$$\tilde{H}|\Omega\rangle = H_{\text{int}}|\Omega\rangle = E_\Omega|\Omega\rangle \neq 0|\Omega\rangle. \tag{8.4.4}$$

However, since $|\Omega\rangle$ is the vacuum state, its energy will still be lower compared to the energy of all other energy-eigenstates:

$$E_n > E_\Omega \quad \forall\, |n\rangle \neq |\Omega\rangle. \tag{8.4.5}$$

Now we consider the early time limit, i.e. we take $t_0 \to -\infty$. In taking this limit, the system will still be highly oscillatory due to the complex exponential and the contribution to the time-integral over the Hamiltonian will average out. In mathematical terms, however, the integral will diverge and we have to regularize the integral to enforce the converging behavior in the early time limit.[8] Therefore, we introduce a small imaginary part to time t_0 as follows:

$$\tilde{t}_0 \equiv t_0(1 - i\epsilon), \tag{8.4.6}$$

where ϵ is small. This formal procedure is called the $i\epsilon$-prescription and effectively turn of the (interaction) Hamiltonian at early times.

Taking the limit $t_0 \to -\infty$, we find that the all exponentials become suppressed. Note that non-vacuum states will be stronger damped than the vacuum state, since $E_n > E_\Omega$. To a good approximation, we therefore get:

$$\lim_{\tilde{t}_0 \to -\infty^+} e^{-i\tilde{H}(t-\tilde{t}_0)}|0\rangle = \lim_{\tilde{t}_0 \to -\infty^+} e^{-iE_\Omega(t-\tilde{t}_0)}\langle\Omega|0\rangle|\Omega\rangle, \tag{8.4.7}$$

where we reintroduced the notation of Eq. 8.1.4 and we find that only the interaction vacuum survives the $i\epsilon$-prescription. Solving this equation for $|\Omega\rangle$ provides the direct relation between the interaction and free-field vacuum in the early time limit:

$$|\Omega\rangle = \lim_{\tilde{t}_0 \to -\infty^+} \frac{e^{-i\tilde{H}(t-\tilde{t}_0)}|0\rangle}{e^{-iE_\Omega(t-\tilde{t}_0)}\langle\Omega|0\rangle} \tag{8.4.8}$$

We can recast the above expression in explicit evolution operator form by manipulating the numerator:

$$\begin{aligned} e^{-i\tilde{H}(t-\tilde{t}_0)}|0\rangle = U(t,t_0)|0\rangle &= U^{-1}(t_0,t)|0\rangle \\ &= U^{-1}(t_0,t)U_0(t_0,t)U_0^{-1}(t_0,t)|0\rangle \\ &= F^{-1}(t_0,t)U_0^{-1}(t_0,t)|0\rangle \\ &= F(t,t_0)U_0(t,t_0)|0\rangle = F(t,t_0)|0\rangle, \end{aligned} \tag{8.4.9}$$

where we used the fact that $H_0|0\rangle = 0$ in the last line. Omitting the early time limit and $i\epsilon$-prescription, we obtain for the intarction vacuum $|\Omega\rangle$ and its complex conjugate:

$$|\Omega\rangle = \frac{F(t,t_0)|0\rangle}{e^{-iE_\Omega(t-t_0)}\langle\Omega|0\rangle}, \qquad \langle\Omega| = \frac{\langle 0|F(-t_0,t)}{e^{iE_\Omega(t+t_0)}\langle 0|\Omega\rangle}. \tag{8.4.10}$$

[8] A similar argument will be applied to the early time limit of Eq. 8.1.3. In order to evaluate for instance $\langle\mathcal{R}\mathcal{R}\mathcal{R}\rangle$, we will have to compute the time-integral and therefore we will insert the mode-function for $\mathcal{R}$, which oscillates like $e^{-ik\tau}$ near $\tau_0 \to -\infty$.

Invoking the normalization condition $\langle\Omega|\Omega\rangle \equiv 1$, we obtain the equation:

$$\langle 0|F(-t_0,t_0)|0\rangle = e^{2iE_\Omega t_0}|\langle 0|\Omega\rangle|^2. \tag{8.4.11}$$

Using these results, we can rewrite the expectation value $\langle\mathcal{Q}(t_*)\rangle$ to the desired form (Eq. 8.1.3). Starting from Eq. 8.3.18, we insert the obtained results for $|\Omega\rangle$ and $\langle\Omega|$ to obtain:

$$\begin{aligned}\langle\mathcal{Q}(t_*)\rangle &= \langle\Omega|F^{-1}(t_*,t_0)\mathcal{Q}^I(t_0)F(t_*,t_0)|\Omega\rangle \\ &= \frac{\langle 0|F(-t_0,t)}{e^{iE_\Omega(t+t_0)}\langle 0|\Omega\rangle}\, F^{-1}(t_*,t_0)\mathcal{Q}^I(t_0)F(t_*,t_0)\, \frac{F(t,t_0)|0\rangle}{e^{-iE_\Omega(t-t_0)}\langle\Omega|0\rangle} \\ &= \frac{\langle 0|F(-t_0,t_0)|0\rangle}{e^{2iE_\Omega t_0}|\langle 0|\Omega\rangle|^2}\, \langle 0|F^{-1}(t_*,t_0)\mathcal{Q}^I(t_0)F(t_*,t_0)|0\rangle. \end{aligned} \tag{8.4.12}$$

Finally, on account of Eq. 8.4.11, we find that we can effectively replace the interacting vacuum by the free-field vacuum:

$$\langle\mathcal{Q}(t_*)\rangle = \langle 0|F^{-1}(t_*,t_0)\mathcal{Q}^I(t_0)F(t_*,t_0)|0\rangle. \tag{8.4.13}$$

Explicitly writing out the expressions for F and introducing the $i\epsilon$-prescription in the early time integration limit, we arrive at the desired result (Eq. 8.1.3):

$$\langle\mathcal{Q}(t_*)\rangle = \langle\left[\bar{T}\exp\left(+i\int_{-\infty^-}^{t_*} dt'\, H^I_{\rm int}(t')\right)\right]\mathcal{Q}^I(t_*)\left[T\exp\left(-i\int_{-\infty^+}^{t_*} dt'\, H^I_{\rm int}(t')\right)\right]\rangle, \tag{8.4.14}$$

where the angular brackets are understood to denote the free-field vacuum $|0\rangle$ expectation value. The $i\epsilon$-prescription is introduced to effectively turn off the interaction Hamiltonian at early times t_0 (see Fig. 8.1). In the context of single-field inflation, the $\tau \to -\infty$ mode functions go as $e^{\pm ik\tau}$ (Eq. 5.4.23) and thus rapidly oscillate. The contribution of these oscillations to the expectation value $\langle\mathcal{Q}(t_*)\rangle$ averages out and we regularize the integral by invoking the $i\epsilon$-prescription.

8.5 Dyson Series and Contractions

In the previous section, we derived the central result of this chapter, i.e. Eq. 8.1.3. Notice that is valid up to any order in the interaction Hamiltonian. In practical terms, however, the representation of Eq. 8.1.3 is inconvenient since we typically work to a specific order in the interaction Hamiltonian. Therefore, it is natural to expand the interaction Hamiltonian using a Dyson series. To second order in $H_{\rm int}$ the Dyson series gives:

$$\begin{aligned} T\exp\left[+i\int_{-\infty^-}^{t_*} dt'\, H^I_{\rm int}(t')\right] &= 1 - i\int_{-\infty^-}^{t_*} dt'\, H^I_{\rm int}(t') \\ &\quad + \frac{i^2}{2}\int_{-\infty^-}^{t_*} dt' \int_{-\infty^-}^{t_*} dt''\, H^I_{\rm int}(t')\, H^I_{\rm int}(t'') + \mathcal{O}(H^3_{\rm int}). \end{aligned} \tag{8.5.1}$$

In terms of Feynman diagrams, each factor of $H_{\rm int}$ corresponds to an interaction vertex and the coupling strength is thus encoded in the interaction Hamiltonian.[9] This is illustrated for a three-point and four-point correlator of $\delta\psi$ in Fig. 8.2.

[9] For inflation, as expected, we will find that the coupling constant is given in terms of slow roll parameters, therefore the coupling is weak and the perturbative analysis by means of the Dyson series is justified.

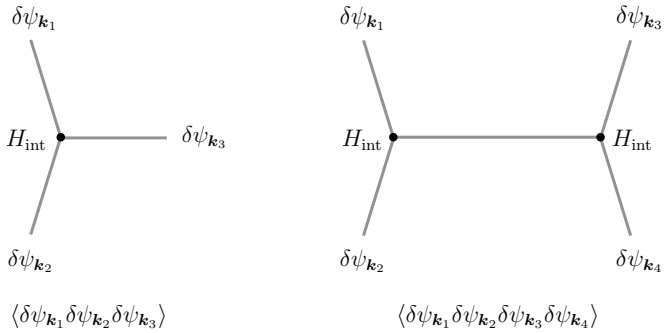


Figure 8.2: The left diagram shows the first order contribution in $H_{\rm int}$ to the three-point correlator (notice that the diagram contains only on vertex). Similarly, the right diagram shows the $\mathcal{O}(H_{\rm int}^2)$ contribution to the four-point correlator.

In the remaining part of this work, we will work only to leading order in $H_{\rm int}$ and hence we can expand Eq. 8.1.3 to first order as:

$$\begin{aligned}\langle\mathcal{Q}(t_*)\rangle &= \langle\left[\mathcal{Q}^I - i\mathcal{Q}^I i\int_{-\infty^-}^{t_*} dt'\; H_{\rm int}^I(t') + i\mathcal{Q}^I i\int_{-\infty^-}^{t_*} dt'\; H_{\rm int}^I(t')\right]\rangle + \mathcal{O}(H_{\rm int}^2)\\ &= \langle\mathcal{O}^I\rangle - i\int_{-\infty^-}^{t_*} dt'\langle 0|\left[\mathcal{Q}^I(t), H_{\rm int}^I(t')\right]|0\rangle + \mathcal{O}(H_{\rm int}^2).\end{aligned} \tag{8.5.2}$$

We conclude that the leading (first order) correction to the three-point correlator, denoted as $\langle\mathcal{O}^I\rangle_1$, is given by:

$$\langle\mathcal{Q}^I\rangle_1 = -i\int_{-\infty^-}^{t_*} dt'\langle 0|\left[\mathcal{Q}^I(t), H_{\rm int}^I(t')\right]|0\rangle. \tag{8.5.3}$$

By Hermicity, we may also write the leading correction in terms of solely purely real or imaginary parts as follows:

$$\langle\mathcal{Q}^I\rangle_1 = -2{\rm Re}\left[i\int_{-\infty^-}^{t_*} dt'\langle 0|\mathcal{Q}^I(t)H_{\rm int}^I(t')|0\rangle\right] = 2{\rm Im}\left[\int_{-\infty^-}^{t_*} dt'\langle 0|\mathcal{Q}^I(t)H_{\rm int}^I(t')|0\rangle\right]. \tag{8.5.4}$$

In order to evaluate the leading order contribution $\langle\mathcal{O}^I\rangle_1$, we need to compute (anti-)time-ordered integrals, for which the integrands are products of fields and their conjugate momenta sandwiched between two free-field vacuum states. Contrary to the static Minkowski background of standard QFT, in the inflationary background, there is no propagator analogous to the Feynman diagram that takes care of the time-ordering [82]. Therefore, we will evaluate the integrands and leave the time-ordering procedure for the final integration (over time).[10]

To enlighten the difference between the In-Out and In-In formalism we will first briefly go over the approach in the standard In-Out formalism. To compute the integrands in

[10]There are approaches possible in which on can construct propagators to easily compute higher order correlation functions [77]. However, those propagators can only be constructed in the de Sitter limit and therefore only form a reasonable approximation to true inflationary background, which is indeed quasi de Sitter. We will take this approach when we compute the bispectrum for the multi-field scenario [76] in the next part of this work.

the In-Out formalism, we use Wick's theorem. The prescription is to expand $\mathcal{Q}$ and H_{int} explicitly into interaction picture fields and use the commutation relations for the quantized interaction picture fields to contract into a product of Feynman propagators. However, in the inflationary background, there is no propagator analogous to the Feynman propagator so we take a different approach.

With respect to the procedure of normal ordering, the In-In approach is equivalent to In-Out: all annihilation operators are placed to the right so they annihilate the vacuum. Hence, we define normal ordering as usual. Consider a quantized scalar field ψ in the interaction picture, written in terms of creation and annihilation operators:

$$\psi_I(\boldsymbol{x},t) = \psi_I^+(\boldsymbol{x},t) + \psi_I^-(\boldsymbol{x},t) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\left[f_k(t)\hat{a}_{\boldsymbol{k}}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} + f_k^*(t)\hat{a}_{\boldsymbol{k}}^\dagger\, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\right], \tag{8.5.5}$$

and the mode function is denoted as $f_k(t)$. The so-called positive frequency mode ψ_I^+ annihilates the free-field vacuum, that is:

$$\psi_I^+|0\rangle = 0. \tag{8.5.6}$$

In terms of $\psi_I^\pm$, the normal ordered string is defined as:

$$:\psi_1\psi_2\cdots\psi_n: \equiv \psi_1^-\psi_2^-\cdots\psi_n^+. \tag{8.5.7}$$

Notice that the vacuum expectation value of a normal ordered string vanishes by construction, that is:

$$\langle 0|\ :\psi_1\psi_2\cdots\psi_n:\ |0\rangle = 0. \tag{8.5.8}$$

For a product of operators, the definition is equivalent to the one above with ψ replaced by the considered operators, since operators can also be expanded into the positive and negative frequency parts of the field.

A contraction between two field operators can now be defined as:

$$\overbracket{\psi_1\psi_2} \equiv [\psi_1^+, \psi_2^-] = (2\pi)^3\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2)\ f_{k_1}(t_1)f_{k_2}^*(t_2). \tag{8.5.9}$$

Although we used equal time commutation relations, we leave the time-ordering to the final integration and therefore the two mode functions can be evaluated at two different times $t_{1,2}$.[11] Furthermore, notice that we contract into so-called Wightman propagators, or more precisely the absolute value squared of the mode function, which are manifestly real (contrary to the Feynman propagator in Minkowski space, which is imaginary). The fact that we contract in the real functions can be traced back to the chosen contour of integration, which starts $-\infty(1-i\epsilon)$ to t_* (moment at which $\langle\mathcal{Q}^I\rangle_1$ is evaluated) and goes back to $-\infty(1+i\varepsilon)$, hence the contour does not close. This causes the contraction into real functions.[12]

Combining the concepts of normal ordering and contractions, the vacuum expectation value of a string of fields can now be evaluated using Wick's theorem:

$$\langle 0|\psi_1\psi_2\cdots\psi_n|0\rangle = \langle 0|\ :\psi_1\psi_2\cdots\psi_n:\ |0\rangle + \ :\ \text{all possible contractions}\ :. \tag{8.5.10}$$

Here, *all possible contractions*, which are also normal ordered, means there will be one contraction term for each possible way of contracting the n operators into pairs. Notice that in the above formulation of Wick's theorem we excluded the time-ordering operator which is present in the usual In-Out version of the theorem, since we leave the time-ordering to the final integration in Eq. 8.1.3.

[11] Another way of justifying this procedure is the fact that the objects relevant for the contraction are the creation and annihilation operators, which are time-independent in contrast to the mode functions.

[12] In the In-Out formalism, the contour can be closed and the resulting propagators are manifestly imaginary.

8.6 Proof of Wick's Theorem

In this final section, we will explicitly derive Wick's theorem, i.e. we will prove that we can write the correlator $\langle 0|\psi_1\psi_2\cdots\psi_n|0\rangle$ as in Eq. 8.5.10. Formally, we can state Wick's theorem as follows. Consider the field operators $\psi(\boldsymbol{x}_{(i)}) \equiv \psi_i$, where $i = 1, \ldots, n$ for an n-point correlation function, and consider the string of field operators:

$$\{\psi_1\psi_2\cdots\psi_n\} = :\ \{\psi_1\psi_2\cdots\psi_n + \text{ all possible contractions}\}\ :\ . \tag{8.6.1}$$

Notice that we consider the fields to be in real space, however one could easily perform the proof in momentum space.

Let us first consider the simple case of $n = 2$, for which Wick's theorem states:

$$\langle 0|\psi_1\psi_2|0\rangle = :\psi_1\psi_2: + \overbracket{\psi_1\psi_2}. \tag{8.6.2}$$

Explicitly expanding the two field modes into positive and negative frequency parts gives:

$$\begin{aligned}\psi_1\psi_2 &= (\psi_1^+ + \psi_1^-)(\psi_2^+ + \psi_2^-)\\ &= \psi_1^+\psi_2^+ + \psi_1^+\psi_2^- + \psi_1^-\psi_2^+ + \psi_1^-\psi_2^-\\ &= \psi_1^+\psi_2^+ + \psi_2^-\psi_1^+ + [\psi_1^+, \psi_2^-] + \psi_1^-\psi_2^+ + \psi_1^-\psi_2^-,\end{aligned} \tag{8.6.3}$$

notice that the non-commutator terms are indeed normal ordered and can be written collectively as $:\psi_1\psi_2:$. Hence, for $n = 2$ Wick's theorem is indeed satisfied.

We will now proceed by induction. Assuming Wick's theorem holds for $n-1$ operators, we will show that it also holds for n operators. The string of n field operators can be written as:

$$\{\psi_1\cdots\psi_n\} = \psi_1\ :\{\psi_2\psi_3\cdots\psi_n + \text{ all possible contractions}\}:\ . \tag{8.6.4}$$

Upon expanding ψ_1 into positive and negative frequency parts, we aim to move them into to normal ordening operator (denoted using $:\cdots:$). For the negative frequency contribution ψ_1^- this is trivial, being on the left, it is already in normal order. The term ψ_1^+ is put in normal order by commuting it to the right past the string $\psi_2\cdots\psi_n$. Consider for instance the first term:

$$\begin{aligned}\psi_1^+ :\{\psi_2\psi_3\cdots\psi_n\}: &= :\{\psi_2\psi_3\cdots\psi_n\}:\ \psi_1^+ + [\psi_1^+, :\{\psi_2\psi_3\cdots\psi_n\}:]\\ &= :\{\psi_1^+\psi_2\cdots\psi_n\}:\\ &+ \ \{[\psi_1^+, \psi_2^-]\psi_3\cdots\psi_n + \psi_2[\psi_1^+, \psi_3^-]\psi_4\cdots\psi_n + \cdots\}:\end{aligned} \tag{8.6.5}$$

The first term in the last line combines with the negative frequency part into $:\{\psi_1\psi_2\cdots\psi_n\}:$ and hence we derived the first term on the right hand side of Wick's theorem (Eq. 8.6.1), as well as all possible terms involving a single contraction of ψ_1 with another field. Similarly, all terms in Eq. 8.6.5 containing a single contraction will generate all possible terms involving that contraction and one contraction of ψ_1 with one of the other field operators in the string. Repeating this argument eventually generates all possible contraction of all n field operators. Therefore, we conclude that the induction step is complete: Wick's theorem is proven.

Chapter 9

ADM Formalism in Inflationary Cosmology

"Not all gauges are born equal and some are cleverer than others."

— Eugene A. Lim

To compute non-Gaussianities in inflation models using the In-In prescription (Eq. 8.1.3), we have to derive the interaction Hamiltonian H_{int}. This can be done by expanding the single-field inflaton-gravity action (Eq. 2.4.17) to third order in perturbations[1] and then reading off the third order Lagrangian to find the interaction Hamiltonian. This perturbative expansion of the action will be performed explicitly in the next chapter.

In order to expand the action to third order, we parametrize the perturbed metric by using the Arnowitt-Deser-Misner (ADM) formalism. For an introduction to the ADM formalism, see e.g. [46]. A review of the formalism by the original author's is given in [8]. Using this formalism to decompose the metric has two main advantages. First, the dynamical degrees of freedom will be manifest in the action (i.e. we can easily distinguish between physical and gauge modes). Second, the ADM formalism uses a convenient foliation of space-time, such that there is a well-defined notion of time. Before going into details, we will discuss the reasoning and motivation behind the ADM formalism in the first section.

9.1 Philosophy of the ADM Formalism

In the previous chapter on the In-In formalism, we considered a Hamiltonian and (implicitly) assumed it is constructed from physical degrees of freedom of the theory.[2] In the case of general relativity, physical degrees of freedom are encoded in the metric $g_{\mu\nu}$, whose evolution is governed by the Einstein-Hilbert action. However, this action is subject to gauge-invariances, caused by diffeomorphism invariance of the space-time coordinates.[3] Due to the gauge invariance, when casting the theory in canonical form, unphysical (gauge) degrees of freedom will arise in the action, caused by the freedom of choosing the coordinate system. By definition, the canonical form of the theory involves the minimal number of variables specifying the state of the considered system.

In more practical terms, when varying the action with respect to the metric, some of the obtained equations will be redundant and serve as constraint equations. These constraints must be satisfied by the solutions to the physical modes but they not dynamically evolve the physical modes. Consequently, the gauge modes are encoded in the constraint

[1] Recall from the previous chapter that the interaction Hamiltonian starts at third order in perturbations.
[2] Physical degrees of freedom are, in principle, measurable quantities.
[3] Diffeomorphism invariance refers to the freedom in choosing a coordinate system.

equations. Constraint equations can be included conveniently in the action via the use of Lagrange multipliers. Lagrange multipliers are defined to be invariant under diffeomorphism transformations (i.e. reparametrizations of the space-time coordinates). That is, a Lagrange multiplier N transforms as:

$$\int N(x)\, dx^\mu \to \int N(\tilde{x})\, d\tilde{x}^\mu, \tag{9.1.1}$$

under the coordinate reparametrization $x^\mu \to \tilde{x}^\mu$.

By decomposing the metric such that constraint equations appear explicitly in the action in combination with a Lagrange multiplier, we can easily distinguish between physical degrees of freedom and gauge modes. The ADM decomposition of metric exactly satisfies these requirements, i.e. it factorizes the action into a form where the constraint equations appear explicitly accompanied by Lagrange multipliers. Moreover, it turns out that we can solve the Lagrange multipliers in terms of physical degrees of freedom and hence we can formulate the action completely in terms of physical degrees of freedom.

9.2 Foliation of Space-Time

The ADM formalism relies on a foliation of space-time by constant-time hypersurfaces, denoted as Σ_t. In mathematical terms, these spatial hypersurfaces are thus the level sets of some function of time $t(x^\mu)$, so that:

$$\Sigma_{t_0} = \{x^\mu : t(x^\mu) \equiv t_0\}, \tag{9.2.1}$$

for some time $t = t_0$. The associated time-like normal vector n^μ scales as:

$$n_\mu \propto \partial_\mu t, \tag{9.2.2}$$

and obeys the constraint $n_\mu n^\mu = -1$. See also the left schematic in Fig. 9.1. Now, we induce coordinates $(t, \boldsymbol{y})$ on the space-time by the coordinate transformation:

$$x^\mu = x^\mu(t, \boldsymbol{y}), \tag{9.2.3}$$

in the following way:

- ▷ For any fixed value of time $t = t_0$, the condition:

$$x^\mu_{t_0}(\boldsymbol{y}) \equiv x^\mu(t_0, \boldsymbol{y}), \tag{9.2.4}$$

  provides the identification of hypersurface Σ_t as the hypersurface Σ_{t_0}, that is:

$$x_{t_0} \ : \ \Sigma \to \Sigma_{t_0}. \tag{9.2.5}$$

  More informally, we have chosen $x^\mu_{t_0}$ such that Σ_t coincides with Σ_{t_0}.

- ▷ The curves:

$$x^\mu_{y_0}(t) \equiv x^\mu(t, \boldsymbol{y}_0) \tag{9.2.6}$$

  connects points on different hypersurfaces corresponding to the same spatial coordinates $\boldsymbol{y}_0$. The curves therefore encode a notion of time evolution from one hypersurface to another.

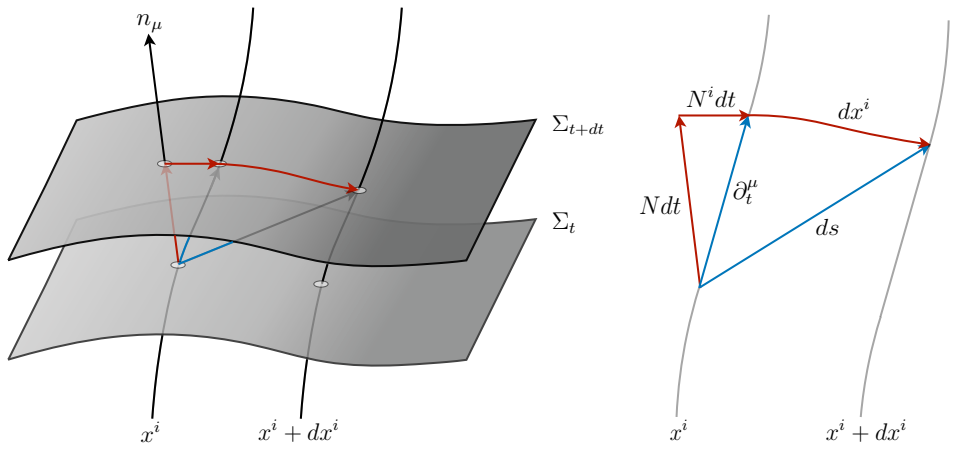


Figure 9.1: Left: Visualization of the space-time foliation in terms of two spatial hypersurfaces Σ_t and Σ_{t+dt}. The normal vector n^μ and two infinitesimally separated curves of constant spatial coordinates x^i and $x^i + dx^i$ are also shown. Right: Definitions of the the lapse N, shift N^i, time evolution vector field ∂_t^μ and the infinitesimal line element ds.

Given the chosen coordinate system $x^\mu = x^\mu(t, \boldsymbol{y})$, the tangent vector to the hypersurface Σ_t can be defined as follows:

$$E_i^\mu = \left.\frac{\partial x^\mu}{\partial y^i}\right|_t, \tag{9.2.7}$$

evaluated at constant time t. In addition, the time evolution vector field ∂_t^μ is defined as:

$$\partial_t^\mu = \left.\frac{\partial x^\mu}{\partial t}\right|_{\boldsymbol{y}}, \tag{9.2.8}$$

and is evaluated while keeping the spatial coordinates $\boldsymbol{y}$ fixed. In general, the curves $x^\mu_{y_0}$ are not required to be normal to hypersurfaces Σ. Hence, we can decompose the time vector field ∂_t^μ into normal and tangential contributions to the hypersurfaces:

$$\partial_t^\mu = N N^\mu + E_i^\mu N^i. \tag{9.2.9}$$

Here N and N^i, called lapse and shift, respectively parametrize the freedom in chosing ∂_t^μ.

In terms of the above quantities, the infinitesimal space-time interval dx^μ reads:

$$dx^\mu = \partial_t^\mu dt + E_i^\mu dy^i = (N N^\mu + E_i^\mu N^i) dt + E_i^\mu dy^i. \tag{9.2.10}$$

The differential line element $ds^2 \equiv g_{\mu\nu} dx^\mu dx^\nu$ becomes:

$$ds^2 = -N^2 dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt), \tag{9.2.11}$$

where we used $N^\mu N_\mu = -1$ the normalization contraint for $N^\mu \equiv (N, N^i)$ and h_{ij}, called the induced metric on spatial hypersurfaces, is given by:

$$h_{ij} \equiv g_{\mu\nu} E_i^\mu E_j^\nu. \tag{9.2.12}$$

The factorization of the differential line element ds^2 as performed in Eq. 9.2.11 is also referred to as the $3+1$ split of the metric and forms the starting point for the Hamiltonian formalism

of general relativity and field theories on a gravitational background (such as inflation). The geometric meaning of the various quantities introduced above (e.g. N and N^i), is visualized in the right schematic of Fig. 9.1.

The components of the metric and its inverse corresponding to the $3+1$ split of the space-time read:

$$g_{\mu\nu} = \begin{bmatrix} -N^2 + N_i N^i & N_i \\ N_j & h_{ij} \end{bmatrix}, \qquad g^{\mu\nu} = \begin{bmatrix} -1/N^2 & N^i/N^2 \\ N^j/N^2 & h^{ij} - N^i N^j/N^2 \end{bmatrix}. \tag{9.2.13}$$

The determinant of the metric can be written simply as: $\sqrt{-g} = N\sqrt{h}$. Finally, the future directed time-like normal vector n_μ to Σ_t can be written in terms of the lapse as:

$$n_\mu = -N\partial_\mu t, \tag{9.2.14}$$

and its explicit components are:

$$n_\mu = -N\delta^0_\mu, \qquad n^\mu = (1/N, -N^i/N). \tag{9.2.15}$$

Using the normal time vector n_μ, the time vector field ∂_t^μ can be decomposed as:

$$\partial_t^\mu = N n^\mu + N^\mu, \tag{9.2.16}$$

and the metric becomes:

$$g_{\mu\nu} = h_{\mu\nu} + n_\mu n_\nu, \qquad g^{\mu\nu} = h^{\mu\nu} + n^\mu n^\nu. \tag{9.2.17}$$

On account of the constraint $n_\mu n^\mu = -1$, the mixed form of the induced metric reads:

$$h^\mu_\nu = \delta^\mu_\nu + n^\mu n_\nu. \tag{9.2.18}$$

The mixed form h^μ_ν is the natural projection operator of any tensorial quantity onto the spatial hypersurface.

9.3 Intrinsic and Extrinsic Curvature

Let us now consider the curvature of the space-time. Since the space-time is foliated into a time-ordered sequence of spatial hypersurfaces, we distinguish two classes of curvature: intrinsic curvature (defined on the spatial hypersurfaces) and extrinsic curvature, describing how the spatial hypersurfaces are bent in the four-dimensional space-time.

The covariant derivative, denoted as $\mathcal{D}_\sigma$ associated with the induced metric $h_{\mu\nu}$ is defined as:

$$\mathcal{D}_\sigma h_{\mu\nu} = \partial_\rho h_{\mu\nu} - \Gamma^\lambda_{\sigma\mu} h_{\lambda\nu} - \Gamma^\lambda_{\sigma\nu} h_{\mu\lambda} = 0. \tag{9.3.1}$$

The Christoffel symbol is understood to be defined with respect to the induced metric $h_{\mu\nu}$, that is:

$$\Gamma^\mu_{\rho\sigma} = \frac{1}{2} h^{\mu\lambda} \left(\partial_\rho h_{\sigma\lambda} + \partial_\sigma h_{\rho\lambda} - \partial_\lambda h_{\rho\sigma}\right). \tag{9.3.2}$$

Using the projection operator h^μ_ν, we can write $\mathcal{D}_\sigma$ associated with $h_{\mu\nu}$ in terms of the covariant derivative ∇_σ constructed from $g_{\mu\nu}$ as follows:

$$\mathcal{D}_\sigma T^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_l} = (h^{\mu_1}_{\rho_1}\cdots h^{\mu_k}_{\rho_k})(h^{\kappa_1}_{\nu_1}\cdots h^{\kappa_l}_{\nu_l}) h^\lambda_\sigma \nabla_\lambda T^{\rho_1\cdots\rho_k}_{\kappa_1\cdots\kappa_k}. \tag{9.3.3}$$

Next, we define the three-dimensional Riemann tensor ${}^{(3)}R^{\sigma}_{\mu\nu\rho}$. Consider a vector V_σ defined on the spatial hypersurface, which is realized by imposing the condition $V_\sigma n^\sigma = 0$ on the vector (such a vector is called a spatial form). The commutator of two covariant derivatives acting on V_σ then defines the three-dimensional Riemann tensor:

$$[\mathcal{D}_\mu, \mathcal{D}_\nu]V_\rho \equiv {}^{(3)}R^{\sigma}_{\mu\nu\rho}V_\sigma. \tag{9.3.4}$$

The Riemann tensor measures the change of a vector when it is parallel transported around a closed loop on the spatial hypersurface. Accordingly, the Ricci tensor and scalar are defined as:

$${}^{(3)}R_{\mu\nu} = {}^{(3)}R^{\rho}_{\mu\rho\nu}, \qquad {}^{(3)}R = h^{\mu\nu}\,{}^{(3)}R_{\mu\nu}. \tag{9.3.5}$$

Note that all quantities introduced above are defined on the spatial slice Σ. In order to describe the full space-time geometry, we need to quantify how the spatial hypersurfaces are embedded in the full four-dimensional space-time. Stated otherwise, we have to describe how the surfaces Σ are bent in the full space-time. Hence, we are naturally led to define the extrinsic curvature tensor $K_{\mu\nu}$ as the covariant derivative of the vector n_μ, which is normal to the spatial hypersurfaces:

$$K_{\mu\nu} \equiv \mathcal{D}_\mu n_\nu. \tag{9.3.6}$$

Essentially, the extrinsic curvature tensor quantifies how the normal vector n_ν changes when slided along the hypersurface Σ. In terms of the covariant derivative associated with $g_{\mu\nu}$, the extrinsic curvature tensor can be written as:

$$K_{\mu\nu} = h^\rho_\mu h^\sigma_\nu \nabla_\rho n_\sigma = h^\rho_\mu \nabla_\rho n_\nu. \tag{9.3.7}$$

In the second equality, we used the projection operator h^σ_ν and invoked the condition $n_\sigma n^\sigma = -1$, giving the constraint:

$$n^\sigma \nabla_\rho n_\sigma = 0. \tag{9.3.8}$$

Below, we will list some relevant properties of the extrinsic curvature tensor. Non-trivial properties will be derived explicitly.

- ▷ *Symmetry.*—The extrinsic curvature tensor is symmetric: $K_{\mu\nu} = K_{\nu\mu}$. This property also shows that the covariant derivative of the normal vector is insensitive to interchanging the indices, i.e. $\nabla_\mu n_\nu = \nabla_\nu n_\mu$.

- ▷ *Spatially Non-Vanishing.*—Only purely spatial components of $K_{\mu\nu}$ are non-vanishing. This property can be derived as follows. First, we write $K_{\mu\nu}$ in terms of the normal vector:

$$K_{\mu\nu} = \nabla_\mu n_\nu + n^\rho n_\mu \nabla_\rho n_\nu. \tag{9.3.9}$$

  Using Eq. 9.3.8, we conclude that $n^\mu K_{\mu\nu} = n^\nu K_{\mu\nu} = 0$. By the fact that $n_i = 0$, we obtain $K^{\mu 0}$, constraining the curvature tensor to be purely spatial:

$$K^{\mu\nu} = K^{ij}. \tag{9.3.10}$$

  This does not necessarily hold for $K_{\mu\nu}$. Nevertheless, the contraction $K^{\mu\nu}K_{\mu\nu}$ will be purely spatial on account of the above equation:

$$K^{\mu\nu}K_{\mu\nu} = K^{ij}K_{ij}. \tag{9.3.11}$$

▷ *Connection to Lie-Derivative of Metric.*—The extrinsic curvature tensor is closely related to the Lie derivative of the metric:

$$K_{\mu\nu} = \frac{1}{2}\mathcal{L}_n h_{\mu\nu}. \tag{9.3.12}$$

The Lie derivative $\mathcal{L}_n$ is defined as:[4]

$$\mathcal{L}_n h_{\mu\nu} \equiv n^\rho \nabla^\rho h_{\mu\nu} + h_{\rho\nu}\nabla_\mu n^\rho + h_{\mu\rho}\nabla_\nu n^\rho. \tag{9.3.13}$$

To derive this relation, we use metric compatibility of the metric with the covariant derivative $\nabla_\rho g_{\mu\nu} = 0$ and the symmetric property of $K_{\mu\nu}$ to derive:

$$\begin{aligned} 2K_{\mu\nu} &= (K_{\mu\nu} + K_{\nu\mu}) \\ &= h^\rho_\mu \nabla_\rho n_\nu + h^\rho_\nu \nabla_\rho h_\mu \\ &= n^\rho n_\mu \nabla_\rho n_\nu + n^\rho n_\nu \nabla_\rho n_\mu + \nabla_\mu n_\nu + \nabla_\nu n_\mu \\ &= n^\rho \nabla_\rho (n_\mu n_\nu) + g_{\rho\nu}\nabla_\mu n^\rho + g_{\mu\rho}\nabla_\nu n^\rho \\ &= n^\rho \nabla_\rho h_{\mu\nu} + h_{\rho\nu}\nabla_\mu n^\rho + h_{\mu\rho}\nabla_\nu n^\rho \equiv \mathcal{L}_n h_{\mu\nu}. \end{aligned} \tag{9.3.14}$$

The scalar lapse is also compatible with the covariant derivative, so that we can rewrite $K_{\mu\nu}$ as follows:

$$K_{\mu\nu} = \frac{1}{2N}\left(N n^\rho \nabla_\rho h_{\mu\nu} + h_{\rho\nu}\nabla_\mu (N n^\rho) + h_{\mu\rho}\nabla_\nu (N n^\rho)\right). \tag{9.3.15}$$

Now using the fact that $N n^\mu = \partial^\mu_t - N^\mu$, we obtain:

$$K_{\mu\nu} = \frac{1}{2N} h^\rho_\mu h^\sigma_\nu \mathcal{L}_{\partial_t - N} h_{\rho\sigma} = \frac{1}{2N} h^\rho_\mu h^\sigma_\nu (\mathcal{L}_t h_{\rho\sigma} - \mathcal{L}_N h_{\rho\sigma}), \tag{9.3.16}$$

where we included inoffensive projections. On account of the relations:

$$h^\rho_\mu h^\sigma_\nu \mathcal{L}_{\partial_t} h_{\rho\sigma} \equiv \dot{h}_{\mu\nu}, \qquad \mathcal{L}_N h_{\rho\sigma} = \mathcal{D}_\rho N_\sigma + \mathcal{D}_\sigma N_\rho, \tag{9.3.17}$$

we arrive at the following final expression for the extrinsic curvature tensor:

$$K_{\mu\nu} = \frac{1}{2N}(\dot{h}_{\mu\nu} - \mathcal{D}_\mu N_\nu - \mathcal{D}_\nu N_\mu). \tag{9.3.18}$$

9.4 Codazzi Equation and the Ricci Scalar

Recall that in order to recast the inflaton-gravity action (Eq. 2.4.17) in terms of the ADM formalism, we will have to rewrite the Ricci scalar – appearing in the Einstein-Hilbert part of the action – in ADM variables. That is, we require a direct relation of the form:

$$R = R(K, K_{\mu\nu}), \tag{9.4.1}$$

where K represents the trace of the extrinsic curvature tensor ($K \equiv K^\sigma_\sigma$). In order to derive this relation, we will use the so-called *Codazzi* equation, which we will first derive now.

[4]This definition of the Lie derivative differs by a sign form the one defined in section 6.4. That is, for arbitrary $A_{\mu\nu}$ and b^ρ we have: $\mathcal{L}_b A_{\mu\nu} \equiv -\Delta_b A_{\mu\nu}$.

Recall that the three-dimensional Riemann tensor is defined in terms of the commutator of the covariant derivative with respect to the induced metric $h_{\mu\nu}$ as follows:

$$[\mathcal{D}_\mu, \mathcal{D}_\nu]V_\rho \equiv {}^{(3)}R^\sigma_{\mu\nu\rho}V_\sigma. \tag{9.4.2}$$

In addition, the double covariant derivative of V_ρ can be projected on the covariant derivative associated with $g_{\mu\nu}$ as follows:

$$\mathcal{D}_\mu\mathcal{D}_\nu V_\rho = \mathcal{D}_\mu(h^\alpha_\nu h^\beta_\rho \nabla_\alpha V_\beta) = h^\varepsilon_\mu h^\kappa_\nu h^\eta_\rho \nabla_\varepsilon(h^\alpha_\kappa h^\beta_\eta \nabla_\alpha V_\beta). \tag{9.4.3}$$

Explicitly expanding the last expression for $\mathcal{D}_\mu\mathcal{D}_\nu V_\rho$ yields:

$$\mathcal{D}_\mu\mathcal{D}_\nu V_\rho = h^\varepsilon_\mu h^{\alpha\beta}_{h\rho}\nabla_\varepsilon\nabla_\alpha V_\beta + h^\beta_\rho K_{\mu\nu}n^\alpha\nabla_\alpha V_\beta + h^\alpha_\nu K_{\mu\rho}n^\beta\nabla_\alpha V_\beta, \tag{9.4.4}$$

where we used the relation:

$$h^\alpha_\mu h^\beta_\nu \nabla_\alpha h^\rho_\beta = K_{\mu\nu}n^\rho. \tag{9.4.5}$$

The middle term in Eq. 9.4.4 vanishes when anti-symmetrized over μ and ν, which is exactly the case in the commutator of Eq. 9.4.2. Furthermore, we can rewrite the last term by invoking:

$$h^\alpha_\nu n^\beta \nabla_\alpha V_\beta = h^\alpha_\nu \nabla_\alpha(n^\beta V_\beta) - h^\alpha_\nu V_\beta \nabla_\alpha n^\beta = -K^\beta_\nu V_\beta. \tag{9.4.6}$$

Combining these results, we arrive at the *Codazzi* equation:

$${}^{(3)}R^\sigma_{\mu\nu\rho} = h^\alpha_\mu h^\beta_\nu h^\delta_\rho\, R^\sigma_{\alpha\beta\delta} + K_{\mu\rho}K^\sigma_\nu - K_{\nu\rho}K^\sigma_\mu, \tag{9.4.7}$$

which relates the three-dimensional curvature, the full space-time curvature and the extrinsic curvature.

Contracting the *Codazzi* equation by setting $\nu = \sigma$ gives the three-dimensional Ricci tensor:

$${}^{(3)}R_{\mu\rho} = h^\alpha_\mu h^\beta_\sigma h^\delta_\rho R^\sigma_{\alpha\beta\delta} + K_{\mu\rho}K - K_{\sigma\rho}K^\sigma_\mu. \tag{9.4.8}$$

Finally, contracting the Ricci tensor with the induced metric $h^{\mu\rho}$ gives the direct relation between the four-dimensional Ricci scalar, the three-dimensional one and the extrinsic curvature.

$$\begin{aligned} {}^{(3)}R = h^{\mu\rho}\,{}^{(3)}R_{\mu\rho} &= h^{\mu\rho}h^\alpha_\mu h^\beta_\sigma h^\delta_\rho R^\sigma_{\alpha\beta\delta} + h^{\mu\rho}K_{\mu\rho}K - h^{\mu\rho}K_{\sigma\rho}K^\sigma_\mu \\ &= h^{\alpha\delta}R_{\alpha\delta} + K^2 - K^\mu_\sigma K^\sigma_\mu \\ R &= R^{(3)} - K^2 + K_{ij}K^{ij}. \end{aligned} \tag{9.4.9}$$

In the line, we used the fact solely the purely spatial part of the extrinsic curvature tensor is non-vanishing.

9.5 Inflaton-Gravity Action in the ADM Formalism

In the previous sections we derived expressions for the metric (Eq. 9.2.13) and Ricci scalar (Eq. 9.4.9) in terms of ADM variables. Now, those results can be used to recast the inflaton-gravity action (Eq. 2.4.17):

$$S = S_{\rm EH} + S_\phi = \int d^4x\sqrt{-g}\left[\frac{1}{2}M^2_{\rm pl}R - \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right]. \tag{9.5.1}$$

in the ADM formalism. Below, we will show how to transform the different parts of the action to the ADM variables.

▷ *Metric Determinant.*—The metric determinant can be written in terms of the scalar lapse function N and the induced metric determinant $\sqrt{h}$ as follows:

$$\sqrt{-g} = N\sqrt{h}. \tag{9.5.2}$$

Hence, the space-time integration measure becomes:

$$\int d^4x\, \sqrt{-g} = \int d^4x\, N\sqrt{h}. \tag{9.5.3}$$

▷ *Gravitational Sector.*—Using Eq. 9.4.9, we can rewrite the Einstein-Hilbert contribution to the action:

$$\frac{1}{2}M_{\rm pl}^2 R = \frac{1}{2}M_{\rm pl}^2\left(R^{(3)} - K^2 + K_{ij}K^{ij}\right). \tag{9.5.4}$$

▷ *Kinetic Sector Inflaton.*—On account of the ADM metric (Eq. 9.2.13), the kinetic part of the inflaton Lagrangian reads:

$$g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi = -\frac{1}{N^2}\left(\dot{\phi}^2 - N^i\partial_i\phi\right)^2 + \partial^i\phi\partial_i\phi. \tag{9.5.5}$$

To recast this in the canonical momentum form, we define π_ϕ in the ADM formalism as:

$$\pi_\phi^2 \equiv \frac{1}{N^2}\left(\dot{\phi}^2 - N^i\partial_i\phi\right)^2, \tag{9.5.6}$$

such that the kinetic sector becomes:

$$g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi = -\pi_\phi^2 + \partial^i\phi\partial_i\phi. \tag{9.5.7}$$

Finally, we obtain the following expression for the inflaton-gravity action in the terms of the ADM variables:

$$S = \int d^4x\, N\sqrt{h}\left[\frac{M_{\rm pl}^2}{2}\left(R^{(3)} + K_{ij}K^{ij} - K^2\right) + \frac{1}{2}\left(\pi_\phi^2 - \partial^i\phi\partial_i\phi\right) - V(\phi)\right]. \tag{9.5.8}$$

For convenience, often the following field redefinition is made to make the curvature tensor manifestly dimenionless:

$$E_{ij} \equiv NM_{\rm pl}K_{ij} = \frac{M_{\rm pl}}{2}(\dot{h}_{ij} - \mathcal{D}_iN_j - \mathcal{D}_jN_i). \tag{9.5.9}$$

In terms of the dimensionless extrinsic curvature perturbation, the ADM inflaton-gravity action becomes:

$$S = \frac{1}{2}\int d^4x\, N\sqrt{h}\left[M_{\rm pl}^2R^{(3)} + N^{-2}(E_{ij}E^{ij} - E^2) + \pi_\phi^2 - \partial^i\phi\partial_i\phi - 2V(\phi)\right]. \tag{9.5.10}$$

Notice that this action only contains first order time derivatives and that the lapse N and shift N^i contain no derivatives. Hence, they will act as non-dynamical Lagrange multipliers, yielding constraint equations.

The constraint equations can be obtained explicitly from the above action by varying with respect to the shift and lapse variables. For the lapse N we have:

$$\frac{\delta S}{\delta N} = M_{\rm pl}^2R^{(3)} - N^{-2}(E_{ij}E^{ij} - E^2) - \partial^i\phi\partial_i\phi - \pi_\phi^2 - 2V(\phi) = 0, \tag{9.5.11}$$

notice that the conjugate momentum π_ϕ contains N-dependence as well. This constraint equation is also called the *Hamilton* constraint equation. Now we do the same for the shift function N^i. The only terms depending on the shift in the ADM action are the second and third term. Varying the π_ϕ^2 contribution with respect to N^i yields:

$$\frac{\delta \pi_\phi^2}{\delta N^i} = 2\pi_\phi \frac{\delta \pi_\phi}{\delta N^i} = -\frac{2\pi_\phi \partial_i \phi}{N}. \tag{9.5.12}$$

For the term $E_{ij}E^{ij}$ we obtain:

$$\frac{\delta}{\delta N^k}(E_{ij}E^{ij}) = 2E^{ij}\frac{\delta E_{ij}}{\delta N^k} = 2E^{ij}h_{kl}\frac{\delta E_{ij}}{\delta N_l} = -4\mathcal{D}_m E^m_k, \tag{9.5.13}$$

and for E^2 we get:

$$\frac{\delta E^2}{\delta N^k} = 2Eh_{kl}\frac{\delta}{\delta N_l}(h^{ij}E_{ij}) = -4D_k E. \tag{9.5.14}$$

The constraint equation regarding the shift can be written as:

$$\frac{\delta S}{\delta N^i} = \frac{\delta \pi_\phi^2}{\delta N^i} + \frac{\delta}{\delta N^i}(E^{ij}E_{ij} - E^2), \tag{9.5.15}$$

yielding the so-called momentum constraint equation:

$$2\pi_\phi \partial_i \phi + \frac{4}{N}\mathcal{D}_j(E^j_i - E\delta^j_i) = 0. \tag{9.5.16}$$

In the literature, it is often anticipated already that the spatial derivative of the inflaton vanishes ($\partial_i \phi = 0$), since at background level the inflaton is spatially homogeneous. However, we will perturb the above constraint equation(s) later on to include fluctuations, for which the spatial gradient does not vanish starting at first order. Nevertheless, in the next chapter we will work in the comoving gauge, for which $\delta\phi = 0$, so we are left with:

$$\mathcal{D}_i\left[N^{-1}(E^i_j - E\delta^i_j)\right] = 0. \tag{9.5.17}$$

Solution to the Lapse and Shift for a Flat FRW Metric

Now we will constrain to the flat FRW background[5] and examine the background solutions to of the constraint equations in terms of the lapse and shift. In particular, we will focus on the Hamilton constraint equation, since the momentum constraint equation is satisfied at background order by $N = 1$ (see Eq. 9.5.17). Specifying to the flat FRW background implies the induced spatial metric can be written as:

$$h_{ij} = a^2\gamma_{ij} = a^2\delta_{ij}, \tag{9.5.18}$$

i.e. h_{ij} becomes the spatial three-metric determined by scale factor a (Eq. 1.2.4).

At zeroth order, the inflaton is spatially homogeneous as imposed by the Cosmological Principle, so that we can set the shift to zero at background level: $N^i = 0$. The Hamilton constraint equation becomes:

$$E^2 - E_{ij}E^{ij} = \dot{\phi}^2 + 2V. \tag{9.5.19}$$

[5]Since we consider a flat geometry, we will set $R^{(3)} = 0$ throughout from now on.

Evaluation of the left hand side yields:

$$E_{ij}E^{ij} = h^{im}h^{jn}E_{ij}E_{mn} = 3M_{\rm pl}^2 H^2, \qquad E^2 = (h^{ij}E_{ij})^2 = 9M_{\rm pl}^2. \tag{9.5.20}$$

where we used that $E_{ij} = M_{\rm pl} a\dot{a}\delta_{ij}$. On account of those results, we then find that the Hamiltonian constraint equation becomes:

$$H^2 = \frac{\rho_\phi}{3M_{\rm pl}^2}, \tag{9.5.21}$$

where the scalar field energy density ρ_ϕ is given by Eq. 2.6.4. We recognize the above equation as the Friedmann equation for a universe dominated by a scalar field. Therefore, the Friedmann equation essentially is a Hamilton constraint equation, rather than an equation of motion for the scale factor.

Chapter 10

Bispectrum for Single-Field Inflation

"Cosmology is a science with only a few observational facts to work with."

— Robert W. Wilson

Now that we have developed the In-In formalism to compute higher-order correlation functions in an expanding background as well as the ADM form of the inflaton-gravity action, we can compute the three-point correlator or bispectrum of the comoving curvature perturbation $\mathcal{R}$ explicitly. This was first done for single-field slow-roll inflation in the pioneering work by Maldacena [64] and later generalized to so-called $P(X, \phi)$ theories of inflation (also called k-inflation), allowing for non-standard kinetic terms by Chen and collaborators [36]. The standard kinetic term is understood to be the one in the inflaton-gravity action (Eq. 2.4.17). See also similar work by Seery and Lidsey [77]. A pedagogical review of the computation of non-Gaussianity in single-field inflation was given in [32].

Notice that, in the same spirit as the preceding part of this thesis, the above work assumes the inflaton scalar field to be the only non-gravitational degree of freedom sourcing the comoving curvature perturbation during inflation. However, the microscopic theory of inflation is still unknown. Therefore, during inflation other degrees of freedom may be present as well, and interactions between those fields and the inflaton may well contribute to the bispectrum. In particular, we know that new degrees of freedom must arise around the Planck scale in order to obtain a UV-complete theory of gravity [17]. There are even scenarios possible in which multiple scalar fields are responsible for the accelerated expansion and, as such, where the notion of a single inflaton field is not applicable.

In terms of the mass or particle spectra of the additional fields during inflation (Fig. 10.2), we distinguish between three different scenarios by which additional fields may leave imprints in the bispectrum and higher order correlation functions. Within the context of effective field theory, those scenarios will be discussed in the first section. In particular, we will motivate why it often is justified – within the effective description of single-field inflation – to neglect the additional degrees of freedom. In the subsequent sections, will follow Maldacena and derive the leading order contribution to bispectrum of single-field inflation.

10.1 Effective Field Theory of Inflation and Particle Spectra

In the absence of a microscopic theory of inflation, the phenomenological physics of inflation is described in the context of effective field theory (EFT) [17, 38, 93]. In this section, we will discuss the essentials of EFT and then apply at a mostly qualitative level to inflation. Our findings will be used to justify, under certain well-defined assumptions, that the bispectrum

can be computed at low energies by neglecting the influence of non-gravitational degrees of freedom other than the inflaton field.

10.1.1 EFT Fundamentals

Effective field theories describe the physics of light degrees of freedom below the cutoff scale Λ, which sets the range of validity of the effective theory. Light particles with masses $m < \Lambda$ are included in theory, while heavy particles with masses $M > \Lambda$ are integrated out in a sense that we will make precise below.

Let ϕ and χ denote the light and heavy field(s) with respect to the cutoff scale, respectively. Their combined Lagrangian, allowing for interactions between the two classes of fields, can be written as:

$$\mathcal{L}[\phi, \chi] = \mathcal{L}_l[\phi] + \mathcal{L}_h[\chi] + \mathcal{L}_{lh}[\phi, \chi], \tag{10.1.1}$$

where the interactions are governed by $\mathcal{L}_{lh}$. To obtain the effective action for the light modes, we integrate out the heavy degrees of freedom by means of the path integral:

$$e^{iS_{\text{eff}}[\phi]} = \int [\mathcal{D}\chi]\, e^{iS[\phi,\chi]}. \tag{10.1.2}$$

In practice, the path-integral approach is rarely taken to obtain the effective action, in fact it only works in case the high energy or UV theory in known (this is not the case for inflation). Instead, a matching calculation order-by-order in perturbation theory is often performed [17].

Via the last approach, the obtained effective Lagrangian will consist of a part containing relevant operators (with mass dimension δ_i lower than four) and an infinite sum of higher order operators:

$$\mathcal{L}_{\text{eff}}[\phi] = \mathcal{L}_{\delta_i<4} + \sum_i c_i \frac{\mathcal{O}_i(\phi)}{\Lambda^{\delta_i-4}}, \tag{10.1.3}$$

where the so-called *Wilsonian* coefficients c_i in the higher order sum are typically of order unity and $\mathcal{O}_i$ are the operators of dimension $\delta_i > 4$ made out of the light fields. By this procedure, typically all operators consistent with the symmetries of the UV complete theory are generated.

In case only the low-energy effective theory is accessible to experiment, i.e. the experiment probes energy scales well below the cutoff scale $E \ll \Lambda$, only a finite number of terms need to be considered in the sum. The reason is that the effects of operators of mass dimension δ_i are proportional to powers of the ratio E/Λ as follows:

$$\frac{\mathcal{O}_i(\phi)}{\Lambda^{\delta_i-4}} \propto \left(\frac{E}{\Lambda}\right)^{\delta_i-4}. \tag{10.1.4}$$

Hence, the higher the dimension of an operator, the smaller its contribution to low-energy observables accessible to experiment. In particular, the effect of operators with dimension $\delta_i > 4$ on the low-energy regime is strongly suppressed in case $E/\Lambda \ll 1$. Notice that for inflation, not even the low-energy effective description is not directly accessible to experiment, since at the moment we can only probe imprints of the physics during inflation in the CMB.[1]

Based on the energy scaling according to Eq. 10.1.4, we discriminate between three different types of operators related to their respective mass dimension δ_i.

[1] An exception to this statement would be the detection of primordial gravitational waves, which would provide a direct measure of the energy scale during inflation.

▷ *Relevant Operators* ($\delta_i < 4$).—These operators dominate the low-energy limit or IR regime the full theory, therefore form the leading contributions to the effective theory and become vanishingly small at high energies ($E \to \Lambda$). This can be seen as well from the energy-scaling relation of the operators (Eq. 10.1.4). For relevant operators $\delta_i < 4$ so that the power $\delta_i - 4$ of the dimensionless energy scale E/Λ becomes negative. Meaning that the contribution of these operators becomes large when the studied energy scale E is much smaller than the cutoff scale Λ. Assuming a four-dimensional space-time, the number of possible relevant operators are $\delta_i = 0$ for the unit operator, $\delta_i = 2$ for a bosonic mass term and $\delta_i = 3$ for cubic scalar field interactions.[2]

▷ *Marginal Operators* ($\delta_i = 4$).—Notice that for $\delta_i = 4$, i.e. the mass dimension equals the number of space-time dimensions, the operator scales with $(E/\Lambda)^0$, so they lie between relevant and irrelevant operators. Since quantum effects could affect their scaling behavior to the IR as well as the UV regime, they are a priori equally important at all energy scales.[3]

▷ *Irrelevant Operators* ($\delta_i > 4$).—At low energies, these operators are suppressed by powers of E/Λ and higher, thus they contribute only weakly to the low-energy observables. Therefore, those operators can typically be neglected in a low-energy analysis. However, this does not mean that they are not important. In fact, typically the irrelevant operators contain information on the physics on high scales. They are only irrelevant in the sense that these operators are weak at low energies.

10.1.2 EFT and Inflation

Now that we have discussed the essentials of EFT, we can apply it to inflation, this was done first in [38, 93]. In the context of EFT, the inflaton-gravity action becomes:

$$S = \int d^4x\sqrt{-g}\left[\frac{1}{2}M_{\rm pl}^2 R + \mathcal{L}_l[\phi] + \sum_i c_i \frac{\mathcal{O}_i[\phi]}{\Lambda^{\delta_i - 4}}\right], \tag{10.1.5}$$

where the low-energy Lagrangian $\mathcal{L}_l[\phi]$ contains the kinetic term and operators with dimension $\delta_i \leq 4$. As mentioned above, the operators with dimension greater than four contained in the sum parametrize the effect of the heavy fields on the light inflaton field.

To quantitatively discriminate between light and heavy fields, we have to determine a suitable cutoff scale Λ. As mentioned in the introduction to this chapter, we expect new degrees of freedom to become relevant around the Planck scale, making $\Lambda \sim M_{\rm pl}$ a natural choice for the cutoff scale. Here, we will make this statement more precise. Recall that the low-energy degree of freedom for gravity is the metric $g_{\mu\nu}$, whose leading dynamics is determined by Einstein-Hilbert action:

$$S_{\rm EH} = \frac{M_{\rm pl}^2}{2}\int d^4x\sqrt{-g}R. \tag{10.1.6}$$

[2] We know that in natural units, the action is dimensionless $[S] = 0$ and the space-time volume measure has dimension $[d^4x] = -4$, so that $[\mathcal{L}] = 4$. The rule is that the dimensions of various parts in a contribution to the Lagrangian have to add up to four.

[3] Additional symmetries or conditions on the field theory could possibly weaken this statement.

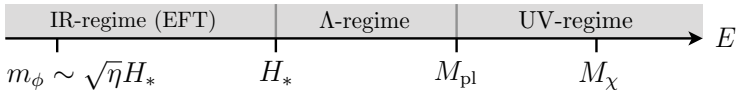


Figure 10.1: Relevant energy scales in the EFT of (single-field) inflation. The three different regimes indicate the low-energy regime (IR), governed by the EFT, the range for the cutoff scale Λ and the UV-regime, containing heavry degrees of freedom M_χ that are excluded by the EFT.

To motive the choice $\Lambda \sim M_{\rm pl}$, we expand the Einstein-Hilbert action around a Minkowski background [17] by writing the metric as:

$$g_{\mu\nu} = \eta_{\mu\nu} + \frac{1}{M_{\rm pl}} h_{\mu\nu}. \tag{10.1.7}$$

Schematically, the expanded form of the Einstein-Hilbert action up to order $\mathcal{O}(h^3)$ is given by:

$$S_{\rm EH} = \int d^4x \partial_\alpha h_{\mu\nu} \partial^\alpha h^{\mu\nu} \left[1 + \frac{h}{M_{\rm pl}} + \frac{h^2}{M_{\rm pl}^2} + \mathcal{O}(h^3/M_{\rm pl}^3) \right], \tag{10.1.8}$$

where we defined the trace as $h \equiv h^\beta_\beta$. Notice that higher order operators $\delta_i > 4$ are suppressed by factors of $M_{\rm pl}$. Hence, in the context of GR, gravity is to be regarded as an effective field theory with cutoff scale $\Lambda = M_{\rm pl}$.

Based on the above considerations concerned with the gravitational sector of the theory, we have an upper limit $\Lambda < M_{\rm pl}$ on the cutoff scale. Furthermore, the cosmological context provides upper bound to Λ. Notice that we are mostly in the quantum fluctuations around the background evolution as described by the inflaton-gravity action. In order to make predictions about these fluctuations, e.g. the bispectrum of the curvature perturbation, they are first evolved to super-horizon scales. Hence, the EFT of inflation should include the fluctuations outside the horizon, corresponding to the energy scale $E \sim H_*$, where H_* is the Hubble scale during inflation. Therefore, we find the following approximate bounds on the cutoff scale:

$$H_* \lesssim \Lambda \lesssim M_{\rm pl}, \tag{10.1.9}$$

and consequently all fields with masses smaller than the Hubble scale during inflation should be included in the EFT:

$$m \lesssim H_*. \tag{10.1.10}$$

In fact, for single-field slow-roll inflation, the inflaton mass should be much smaller than the Hubble scale. This condition is imposed by the second (potential) slow roll parameter:

$$\eta_{\rm v} \equiv M_{\rm pl}^2 \frac{V_{\phi\phi}}{V} \ll 1. \tag{10.1.11}$$

Identifying the second field derivative of the potential with the mass of the inflaton,[4] i.e. $V_{\phi\phi} \equiv m_\phi^2$, we find using Eq. 2.7.2 that:

$$\eta_{\rm v} = \frac{m_\phi^2}{3H^2}, \tag{10.1.13}$$

[4]This identification is a generalization of the quadratic mass term:

$$V = \frac{1}{2} m_\phi^2 \phi^2, \tag{10.1.12}$$

for which $V_{\phi\phi}$ is indeed equal to the inflaton mass squared.

so we find $m_\phi \ll H_*$ in case of slow-roll. Unfortunately, in the context of EFT, it is challenging to maintain this hierarchy between the inflaton mass and the Hubble scale, as we will explain below.

10.1.3 The Mass-Hierarchy and Eta Problem of Inflation

In the absence of additional symmetries, quantum corrections to the low-energy operators tend to drive a scalar mass, and in particular the inflaton mass, to the cutoff scale. To make this statement more quantitative, we review the results of [26], which were discussed in the context of inflation in [17]. Following [17, 26], we consider the following Lagrangian which couples the dynamics of the light scalar field ϕ (which may be regarded as the inflaton) to the heavy degree of freedom χ:

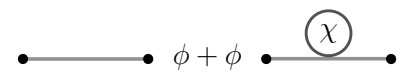
$$\mathcal{L} = -\frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m_\phi^2\phi^2 - \frac{1}{2}M_\chi^2\chi^2 - \frac{1}{4}g\phi^2\chi^2, \tag{10.1.14}$$

where g is the coupling constant between the light and heavy fields and the bare masses of the fields are assumed to obey the hierarchy $m_\phi \ll m_\chi$. The EFT for the light degree of freedom takes the form:

$$\begin{aligned}\mathcal{L}_{\text{eff}}[\phi] = &-\frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m_{\text{E}}^2\phi^2 \\ &- \sum_i \left(\frac{c_i(g)}{M^{2i}}\phi^{4+2i} + \frac{d_i(g)}{M^{2i}}(\partial\phi)^2\phi^{2i} + \dots\right),\end{aligned} \tag{10.1.15}$$

where the Wilsonian coefficients c_i and d_i depend on the coupling constant g and are associated with non-derivative and derivative operators, respectively. The effective mass of the light scalar is denoted as m_{E}.

The effective mass m_{E}, which would be the actual mass of the inflaton in the EFT picture, can be expressed in terms of the bare mass of the light scalar field m_ϕ via a matching procedure [17], which introduces a quantum correction to the effective mass. In particular, the bare mass of the light scalar receives so-called loop corrections, corresponding to Feynman diagrams in which the heavy field χ runs in a loop. We will not go into details here, but merely present the result of the (one-loop) calculation performed in [17, 26]. At one-loop, i.e. considering only quantum corrections corresponding to Feynman diagrams with one loop, the relation between the bare mass m_ϕ and the effective mass m_{E} is given by:

$m_{\text{E}}^2 =$ 

$= m_\phi^2 + \frac{g}{32\pi^2}(\Lambda^2 - M_\chi^2 L),$ (10.1.16)

where $L \equiv \ln(\Lambda^2/{}_\phi^2)$ and we have set the arbitrary renormalization scale to be the mass of the inflaton ${}_\phi$.

The first term, proportional to Λ^2, in the correction depends on the renormalization scheme and is therefore unphysical [17]. However, the second contribution, $M_\chi^2 L$, is physical and hence the effective mass gets a large contribution due to the mass of the heavy scalar. The existence of a light scalar field is therefore unnatural, in the respect that large quantum corrections of order $\mathcal{O}(M_\chi)$ must be canceled by a large bare mass m_ϕ with opposite sign, in order to render the effective mass of the (inflaton) scalar field small. In other words, significant fine-tuning is required to obtain a small effective inflaton mass. This is a challenge

for inflation, since for typical (i.e. non fine-tuned) values of m_{E} the scalar field ϕ will not even be part of the low-energy EFT.

It should be mentioned in the multi-field scenario, the existence of light scalar fields responsible is less difficult to realize. In particular, in assuming string theory to be UV complete theory, from which inflation should be derived as one of the low-energy consequences, the existence of a large number of light scalar fields is natural. Therefore, from this viewpoint, inflationary cosmologies with a large number of light scalars seems to be plausible [4].

In addition to the mass hierarchy problem, there is another (related) problem, called the η-problem [17]. If the inflaton scalar field has large couplings to the heavy fields, integrating out these heavy degrees of freedom leads to an effective Lagrangian of the form of Eq. 10.1.15 with Wilsonian couplings c_i and d_i of order unity. Those large couplings typically lead to corrections to the potential of the form:

$$\delta V = c_{\mathrm{W}} V(\phi) \frac{\phi^2}{\Lambda^2}, \tag{10.1.17}$$

where c_{W} denotes a Wilsonian coefficient of order unity appearing in the effective Lagrangian. Such a correction to the potentials yields the following contribution to the η paramater:

$$\delta\eta = M_{\mathrm{pl}}^2 \frac{\delta V_{\phi\phi}}{V} = 2c_{\mathrm{W}} \left(\frac{M_{\mathrm{pl}}}{\Lambda} \right)^2 > 1, \tag{10.1.18}$$

where we assumed $M_{\mathrm{pl}} \geq \Lambda$ as motivated in the previous section. Notice that such a correction the η parameter seems to make slow-roll inflation unnatural [17], as $\eta \ll 1$ is required to sustain slow-roll inflation. In conclusion, a plausible microscopic theory of inflation must address the mass hierarchy issue and η-problem.

10.1.4 Particle Spectra during Inflation

Now that the EFT of inflation, its challenges and the relevant energy scales during inflation are discussed, we will now introduce three generic inflationary scenarios, called single-field, quasi single-field and multi-field inflation, based on their particle spectrum. The particle spectrum is illustrated for the three scenarios in Fig. 10.2.

Single-Field Inflation

In the single-field scenario, the inflaton field is assumed to be much lighter than the Hubble scale H_*, although this is unnatural from the EFT viewpoint in the sense that the existence of light scalars requires fine-tuning of the bare mass. This scenario is illustrated in Fig. 10.2 (a). We assume that heavy fields χ satisfy the hierarchy $M_\chi \gg \Lambda$. In addition, we assume the Hubble scale, i.e. the scale of inflaton fluctuations, to be well contained in the EFT regime. In that case, the curvature perturbation $\mathcal{R}$ will also be in the low-energy range of the theory. To be more precise, since $\mathcal{R}$ is typically first evolved to super-horizon scales before observables such as the bispectrum will be evaluated, we expect its energy scale $E_{\mathcal{R}}$ to be comparable to the Hubble scale:

$$E_{\mathcal{R}} \sim H_* \ll \Lambda, \tag{10.1.19}$$

and therefore to be well within the EFT regime. Hence, in computing the bispectrum, we can, to a good approximation, neglect the effects of (unknown) heavy degrees of freedom. This is illustrated in the space-time diagram of Fig. 10.3. There is another way of justifying this

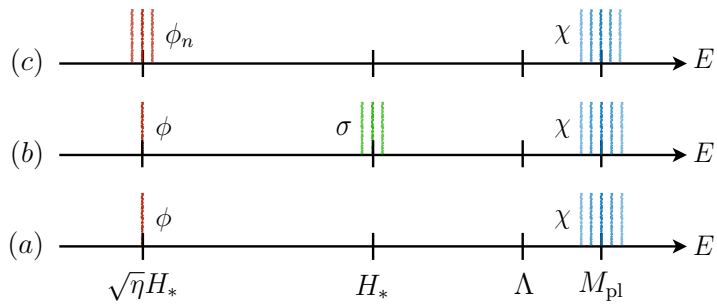


Figure 10.2: Particle Mass Spectra in (a) single-field, (b) quasi single-field and (c) multi-field inflation. The light scalar field(s) responsible for inflation are denoted as $\phi_{(n)}$. Fields with mass comparable to the Hubble scale are collectively denoted as σ. Heavy fields, i.e. fields above the EFT cutoff scale Λ and comparable to the Planck mass, are represented by χ.

simplification. Assuming the heavy fields to be at the minimum of their potential, quantum effects could in principle replace the fields slightly from the minimum, thereby generating quantum fluctuations. However, since fields are so heavy compared to the Hubble scale H_* at which quantum fluctuations are probed, the quantum fluctuations in those fields will be strongly suppressed. In other words, due to the heavy masses, it is difficult for quantum effects to de-locate the heavy fields from their respective equilibrium positions. In the remaining part of this chapter, we will assume that this scenario is valid in order to compute the Bispectrum in accordance with the treatment of Maldacena [64].

Quasi Single- and Multi-Field Inflation

In case of multi-field inflation, the hierarchy is the same as for single-field inflation, as visualized in Fig. 10.2 (c). The only difference is the fact that there are multiple light scalar field responsible for the accelerated expansion and hence the notion of a single inflaton field ceases to be valid. Non-Gaussianities generated by multi-field models will be discussed in the next part of this thesis and, similarly to the approach in the single-field case, we assume the effects of heavy degrees of freedom to be negligible. Although we neglect the imprints of heavy exotic degrees of freedom, it is indeed interesting to examine the effects of such exotic modes on low-energy observables such as the bispectrum of $\mathcal{R}$. Under certain circumstances, it turns out to be possible to extract information about heavy modes, such as the mass and spin of the corresponding quanta, from the low-energy observables. This approach to probe high energy physics via cosmology goes under the name of *Cosmological Collider Physics* [7].

The last scenario (see Fig. 10.2 (b)), known as quasi single-field inflation, forms the bridge between the single- and multi-field cases. Here, we extend the single-field scenario by allowing for fields σ with masses comparable to the Hubble scale during inflation:

$$m_\sigma \sim H_*. \tag{10.1.20}$$

Classically, these fields are too heavy to contribute to the dynamics of inflation and the fields remain at the minimum of the potential. However, the fields σ are light enough for quantum effects to displace the field slightly from its classical equilibrium position and therefore quantum fluctuations $\delta\sigma$ will be generated. Those fluctuations could couple to the inflaton and therefore be observable. In [35], observable imprints in the bispectrum of $\mathcal{R}$ where computed

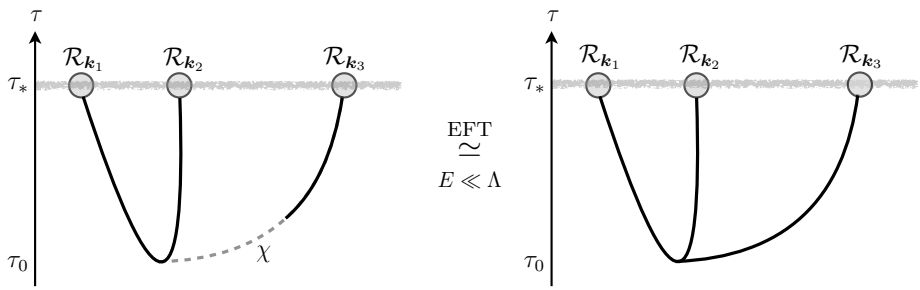


Figure 10.3: In the EFT of single-field inflation, we can neglect the imprints of heavy degrees of freedom on low energy observables, in this case the bispectrum of the comoving curvature perturbation. The space-time diagram visualizes the three-point correlation function of the three $\mathcal{R}$ modes, propagating from the initial time τ_0 to the time of evaluation τ_*. In the full theory, there may be interactions with heavy fields χ as illustrated in the left panel. However, in the low-energy EFT, i.e. considering only processes at energies $E \ll \Lambda$, we can neglect the effects of heavy modes.

for this scenario. This scenario will not be considered in this work but is mentioned for completeness.

10.2 Perturbative Solutions to the Constraint Equations

As discussed in detail in the previous chapter, we solve the Hamilton and momentum constraint equations for the lapse N and shift N^i and plug the results back into the ADM action, so that the action is written solely in terms of physical degrees of freedom. Notice that we want to write the action up to and including third order perturbations, so that we can derive the interaction Hamiltonian. Therefore, the first task is to determine to what order we should solve the constraint equations in order to get the action correction correct up to third order.

10.2.1 Order of Perturbative Constraint Equations

Let us the generalize the question by asking up to which order the action will be correct if we solve the constraint equations up to n-th order. To answer this question, we derive the theorem given in [70]. The theorem states, when solving the constraint equations up to n-th order, the perturbed action will correct to $(2n+1)$-th order in perturbations, up to a total derivative.

Consider a Lagrangian, depending on a generic constraint N, which could be the shift or lapse and possibly a perturbation δX (which would be $\mathcal{R}$ in our case):

$$\mathcal{L} = \mathcal{L}(N, \delta X). \tag{10.2.1}$$

Assuming N is a Lagrange multiplier, there will be no time derivatives acting on it, and the Euler-Lagrange equation will be:

$$\frac{\delta \mathcal{L}}{\delta N} \equiv \frac{\partial \mathcal{L}}{\partial N} - \partial_i \left(\frac{\delta \mathcal{L}}{\partial \partial_i N} \right) = 0. \tag{10.2.2}$$

Solving the constraint N up to n-th order, the EL equations will be satisfied up to and including n-th order:

$$\frac{\delta\mathcal{L}}{\delta N} = 0 + \mathcal{O}(N_{m>n}, \partial_i N_{m>n}, \delta X). \tag{10.2.3}$$

Let us denote the formal exact solution to the Lagrange multiplier by $N_{\rm sol}$, which may or may not exist analytically.[5] The Lagrangian $\mathcal{L}(N_{\rm sol})$ can be expanded in terms of the n-th and higher order solutions by means of the following Taylor expansion:

$$\begin{aligned}\mathcal{L}(N_{\rm sol}, \partial_i N_{\rm sol}, \delta X) &\equiv \mathcal{L}(N_{m\leq n} + \delta N_{m>n}, \partial N_{m\leq n} + \partial\delta N_{m>n}, \delta X)\\ &= \mathcal{L}(N_{m\leq n}, \partial N_{m\leq n}, \delta X) + \delta N_{m>n}\frac{\partial\mathcal{L}}{\partial N}(N_{m\leq n}, \partial N_{m\leq n}, \delta X)\\ &+ \partial\delta N_{m>n}\frac{\partial\mathcal{L}}{\partial\partial_i N}(N_{m\leq n}, \partial N_{m\leq n}, \delta X) + \mathcal{O}\big((n+1)^2\big).\end{aligned} \tag{10.2.4}$$

Using the product rule, we can generate a total spatial derivative term as follows:

$$\begin{aligned}\mathcal{L}(N_{\rm sol}, \partial_i N_{\rm sol}, \delta X) &= \mathcal{L}(N_{m\leq n}, \partial N_{m\leq n}, \delta X)\\ &+ \delta N_{m>n}\left[\frac{\partial\mathcal{L}}{\partial N} - \partial_i\left(\frac{\partial\mathcal{L}}{\partial\partial_i N}\right)\right](N_{m\leq n}, \partial_i N_{m\leq n}, \delta X)\\ &+ \partial_i\left(\delta N_{m>n}\frac{\partial\mathcal{L}}{\partial\partial_i N}\right)(N_{m\leq n}, \partial N_{m\leq n}, \delta X) + \mathcal{O}\big((n+1)^2\big).\end{aligned} \tag{10.2.5}$$

We recognize the combination between the square brackets in the second line as Eq. 10.2.2, which is solved to n-th order by $N_{m\leq n}$. Setting the second line to zero introduces an error of order $\mathcal{O}(2n+2)$, since the term $\delta N_{m>n} = N_{\rm sol} - N_{m\leq n}$ introduces an error of order $n+1$. Furthermore, setting $\delta\mathcal{L}/\delta N$ to zero while solving only up to n-th order also introduces an error at the order $n+1$. The combined error will thus be of the order $\mathcal{O}(2n+2)$. Hence, when solving the constraint equations up to n-th order, the action will correct up to order $\mathcal{O}(2n+1)$, which verifies the advocated result.

10.2.2 Constraint Equations to First Order in Perturbations

Now we will solve the constraint equation for the lapse and shift to first order and those solutions will be sufficient to obtain the third order action S_3. We will work in the comoving gauge, which, for scalar perturbations, can be defined as follows:

$$\delta\phi = 0, \qquad h_{ij} = a^2 e^{2\mathcal{R}}\delta_{ij} = a^2(1+2\mathcal{R})\delta_{ij} + \mathcal{O}(\mathcal{R}^2), \tag{10.2.6}$$

since the gravitational potential Ψ equals $\mathcal{R}$ in this gauge. To connect to the notation of Maldacena [64], we parametrize the scale factor as $a \equiv e^{\rho}$, such that $\dot\rho = H$.

Before getting into mathematical details, we will set up notations here. As advocated above, we will obtain expressions for N and N_i to first order in perturbations. We know that the zeroth order (background) solutions are given by $N^{(0)} = 1$ and $N^{(0)} = 0$, and hence we can write the lapse and shift to first order in perturbations as:

$$N = 1 + N^{(1)} + \mathcal{O}(N^{(2)}), \qquad N_i = N_i^{(1)} + \mathcal{O}(N_i^{(2)}). \tag{10.2.7}$$

[5] In case the solution N can only be obtained perturbatively, $N_{\rm sol}$ corresponds to infinite order, i.e. $N_{\rm sol} = \lim_{n\to\infty}\sum_{i=0}^{n} N_i$.

Since the shift is a spatial vector, we decompose it as usual according to the Helmholtz decomposition prodecure into the gradient of scalar ψ and a divergenceless vector $\hat{N}_i$. That is:

$$N_i^{(1)} = \partial_i \psi + \hat{N}_i, \qquad \partial^i \hat{N}_i = 0. \tag{10.2.8}$$

where we constrained to first order but this decomposition would also apply to higher order terms. Furthermore, we will have to be careful with contracting indices. Indices can only be contracted by using the induced space-time metric h_{ij} and *not* by using the Kronecker delta δ_{ij}. To make this difference manifest, for a generic vector A_i, we define:

$$A_i A^i \equiv h^{ij} A_i A_j, \qquad A_i A_i \equiv \delta^{ij} A_i A_j. \tag{10.2.9}$$

That is, the indices in $A_i A_i$ are summed over but *not* contracted. In particular the gradient operator is given by:

$$\partial_i \partial^i \equiv \nabla^2 = h^{ij} \partial_i \partial_j = a^{-2} \delta^{ij} \partial_i \partial_j = a^{-2} \partial^2, \tag{10.2.10}$$

where we defined $\partial^2 \equiv \partial_i \partial_i$. Finally, we will set the Planck mass $M_{\rm pl}$ to unity in the rest of this chapter and reinstate it in the final result: the bispectrum for single-field inflation.

In order to perturb the Hamilton and momentum constraint equations to first order, the three-dimensional Ricci tensor $R^{(3)}$ and curvature terms $E_{ij}E^{ij} - E^2$ should be expressed in terms of first-order perturbations. We start by deriving the expression for $R^{(3)}$. The Christoffel symbol can be written as:

$$\Gamma^k_{ij} = a^2 h^{kl} (\partial_j \mathcal{R} \delta_{il} + \partial_i \mathcal{R} \delta_{jl} - \partial_l \mathcal{R} \delta_{ij}), \tag{10.2.11}$$

which is first order in perturbations. Hence, terms in the definition of the spatial Ricci tensor R_{ij} which are quadratic in Christoffel symbols can be neglected in a first order analysis:

$$R_{ij} = \partial_k \Gamma^k_{ij} - \partial_j \Gamma^k_{ik} + \mathcal{O}(\mathcal{R}^2). \tag{10.2.12}$$

The three-dimensional Ricci scalar can be derived by contracting with the induced spatial metric:

$$R^{(3)} = h^{ij} R_{ij} = -4 \partial_k \partial^k \mathcal{R}. \tag{10.2.13}$$

Similarly, we will perturb the rescaled curvature tensor E_{ij}, which is defined as (recall $M_{\rm pl} \equiv 1$):

$$E_{ij} = \frac{1}{2} (\dot{h}_{ij} - \mathcal{D}_i N_j - \mathcal{D}_j N_i). \tag{10.2.14}$$

We can simplify the covariant derivatives $\mathcal{D}_i N_j$ by noting that the Christoffel symbols and shift will start at first order in perturbations (at background level N_i vanishes). Hence, we can make the replacement:

$$\mathcal{D}_i N_j = \partial_i N_j + \mathcal{O}(\mathcal{R}). \tag{10.2.15}$$

Taking the time derivative of the spatial metric, we obtain the following expression:

$$E_{ij} = a^2 e^{2\mathcal{R}} (H + \dot{\mathcal{R}}) \delta_{ij} - \partial_{(i} \partial_{j)} \psi, \tag{10.2.16}$$

where we made the substitution $N_i^{(1)} = \partial_i \psi$ since we only consider scalar perturbations. Now we can write the term $E_{ij}E^{ij} - E^2$ to first order in perturbations as:

$$\begin{aligned} E_{ij}E^{ij} - E^2 &= (h^{ik} h^{jl} - h^{ij} h^{kl}) E_{ij} E_{kl} \\ &= -6(H + \dot{\mathcal{R}})^2 + 4(H + \dot{\mathcal{R}}) \partial_i \partial^i \psi. \end{aligned} \tag{10.2.17}$$

Momentum Constraint

First we consider the momentum constraint, which is given by (Eq. 9.5.17):

$$\mathcal{D}_i \left[N^{-1}(E^i_j - E\delta^i_j)\right] = 0. \tag{10.2.18}$$

Contracting the first-order perturbed expression for E_{ij} with the induced metric (at zeroth order in perturbations) yields:

$$E = 3(H + \dot{\mathcal{R}}) - \partial_i\partial^i\psi. \tag{10.2.19}$$

The mixed form of the perturbed curvature tensor E^i_j can be written as:

$$E^i_j = 3(H + \dot{\mathcal{R}})\delta^i_j - \partial^{(i}\partial_{j)}\psi. \tag{10.2.20}$$

Using these results, the momentum constraint equation becomes:

$$\partial_i \left[(1 - N^{(1)})(H + \dot{\mathcal{R}})\delta^i_j\right] = 0, \tag{10.2.21}$$

where we replaced the covariant derivative by the partial derivative and the contributions proportional to ψ from E and E^j cancel. Extracting the purely first order contribution from this equation gives:

$$\partial_j(\dot{\mathcal{R}} - N^{(1)}H) = 0, \qquad N^{(1)} = \frac{\dot{\mathcal{R}}}{H}, \tag{10.2.22}$$

from which we obtained the first order solution to the lapse function in terms of the comoving curvature perturbation.

Hamilton Constraint

Similar to the perturbed momentum constraint, we can perturb the Hamilton constraint as follows:

$$-\partial_k\partial^k\big(\mathcal{R} + H\psi\big) + \frac{1}{2}\dot{\phi}^2 N^{(1)} - 3H\big(HN^{(1)} - \dot{\mathcal{R}}\big) = 0, \tag{10.2.23}$$

where we used the background Friedmann equation to simplify the result (i.e. to get rid of all zeroth order terms).On account of the obtained expression for the first order perturbed lapse, we find that the last term vanishes and ψ can be solved as:

$$\psi = -\frac{\mathcal{R}}{H} + \frac{\dot{\phi}^2 a^2}{2H^2}\partial^{-2}\dot{\mathcal{R}}, \tag{10.2.24}$$

where we used that the inverse Laplacian can be written as $\nabla^{-2} = a^2\partial^{-2}$. The inverse Laplacian is most easily defined in Fourier space, where it equals $-k^{-2}$. Using the definition of the slow roll parameter ε, we can define the field $\partial^2\chi \equiv a^2\varepsilon\dot{\mathcal{R}}$ and we obtain:

$$\psi = -\frac{\mathcal{R}}{H} + \chi. \tag{10.2.25}$$

10.3 Perturbed Inflaton-Gravity Action

Now that we have obtained explicit expressions for the lapse and shift to first order in perturbations, can substitute them into the ADM form or the inflaton-gravity action to obtain an action which is written solely in terms of physical degrees of freedom. Below, we will construct the action to first, second and finally third order in perturbations.

10.3.1 Linear Action

We expect the first order action in perturbations to vanish on account of the background solution of the system (i.e. the Friedmann equation for a universe dominated by a scalar field). The reason is that, as discussed in section 8.2, we expect first order contributions to cancel in the Hamiltonian and hence perturbative corrections to the background evolution start quadratic in perturbations. Here, we will aim to verify this prediction.

Perturbing the ADM inflaton-gravity action (Eq. 9.5.10) to first order in perturbations using the expressions for ψ, $N^{(1)}$ and Eq. 10.2.17, we find:

$$S_1 = \frac{1}{2}\int d^4x\, a^3\left[4M_{\rm pl}^2\partial_i\partial^i(H\psi - R) - 12M_{\rm pl}^2 H\dot{\mathcal{R}} + \frac{\dot{\mathcal{R}}}{H}(6M_{\rm pl}^2H^2 - 2V - \dot{\phi}^2)\right], \tag{10.3.1}$$

where we reinstated dimensional factors of the Planck mass. We recognize the last term (proportional to $\dot{\mathcal{R}}/H$) to be vanishing on account of the background Friedmann equation (conform Eq. 9.5.21). Now, we obtain:

$$S_1 = \frac{1}{2}\int d^4x\, a^3\big[4M_{\rm pl}^2\partial_i\partial^i(H\psi - R) - 12M_{\rm pl}^2 H\dot{\mathcal{R}}\big]. \tag{10.3.2}$$

Since there is no spatial boundary term in the inflationary universe, we can neglect the spatial boundary term in S_1. However, non-boundary terms involving temporal derivatives should be treated with more care. In particular the term proportional to $\dot{\mathcal{R}}$ will not neccesarily be zero the moment of evaluation $\dot{\mathcal{R}}(t_*)$.

The non-vanishing behavior of the linear action via such non-boundary terms induces so-called tadpole contributions to the bispectrum. However, it is shown in [30] that the tadpole contribution in the action can be set to zero via the procedure of renormalization and hence it will be neglected in the further analysis. Moreover, for single-field inflation there exists a dynamical argument as well, concerned with the constancy of the comoving curvature perturbation outside the horizon. Since we will evaluate the bispectrum when the modes are outside the horizon, the comoving curvature perturbation will be vanishing, rendering the last term zero so that we can typically neglect it.

In addition, note that we have a hidden non-derivative term in the linear action S_1, associated with the term:

$$\begin{aligned} S_1 &\supset \frac{1}{2}\int d^4x\, a^3\left(4M_{\rm pl}^2\partial_i\partial^i(H\psi)\right) \\ &= 2\int d^4x\, a^3 M_{\rm pl}^2\partial_i\partial^i(H\psi). \end{aligned} \tag{10.3.3}$$

Using the expression for ψ, we find that this contribution contains a spatial boundary term as well as a non-boundary term involving $\dot{\mathcal{R}}$ and the latter is given by:

$$S_1 \supset 2\int d^4x\, a^3 M_{\rm pl}^2 H\varepsilon\dot{\mathcal{R}}. \tag{10.3.4}$$

Similarly to the reasoning above, we will neglect this boundary term as well.

10.3.2 Quadratic Action

Now we proceed by considering the second order action. As discussed in section 8.3, the evolution of the interaction picture fields is governed by the quadratic action. Perturbing the

ADM action to second order in perturbations gives [64]:

$$
\begin{aligned}
S_2 = \frac{1}{2}\int d^4x \Bigg[& a e^{\mathcal{R}}\left(1+\frac{\dot{\mathcal{R}}}{H}\right)\left(-4\partial^2\mathcal{R} - (\partial\mathcal{R})^2 - 2Va^2e^{2\mathcal{R}}\right) \\
& + a^3 e^{3\mathcal{R}}\left(1-\frac{\dot{\mathcal{R}}}{H}+\frac{\dot{\mathcal{R}}^2}{H^2}\right)\Big(-6(H+\dot{\mathcal{R}})^2 + \dot{\phi}^2 \\
& + 4a^{-2}e^{-2\mathcal{R}}(H+\dot{\mathcal{R}})\left(\partial_i\psi\partial_i\mathcal{R} + \partial^2\psi\right)\Big)\Bigg],
\end{aligned}
\tag{10.3.5}
$$

where we introduced the notation $(\partial\mathcal{R})^2 = \delta^{ij}\partial_i\mathcal{R}\partial_j\mathcal{R}$. This result can be simplified by performing a number of integrations by parts to get [32, 36, 64, 88]:

$$
\boxed{S_2 = \frac{1}{2}\int d^4x \frac{\dot{\phi}^2}{M_{\rm pl}^2 H^2}\left[a^3\dot{\mathcal{R}}^2 - a(\partial\mathcal{R})^2\right] + \partial S_2,}
\tag{10.3.6}
$$

where we have contained the boundary terms generated by the integration by parts in the term ∂S_2. Explicitly, ∂S_2 reads [88]:

$$
\begin{aligned}
\partial S_2 = & -\int d^4x\, \partial_0\left[\frac{a}{H}\mathcal{R}\partial^2\mathcal{R} - 2a^3H(2+6\mathcal{R}+9\mathcal{R}^2)\right] \\
& - \int d^4x\, a^2\partial_k\left[\left(2+\mathcal{R}+\frac{\dot{H}}{H^2}\mathcal{R}+\frac{\dot{\mathcal{R}}}{H}\right)\partial_k\mathcal{R} - \frac{\mathcal{R}}{H}\partial_k\mathcal{R}\right] \\
& + \int d^4x\, a^3\partial_k\left[4H\mathcal{R}\partial_k\psi - \partial^2\psi\partial_k\psi + \partial_i\psi\partial_i\partial_k\psi\right].
\end{aligned}
\tag{10.3.7}
$$

The terms in the third line above are again no proper boundary terms, leading to tadpole contributions. Following the standard literature [32, 36, 64], we neglect these terms in our computations.

Discarding the boundary term and making the field redefinition $f \equiv z\mathcal{R}$ with $z \equiv a\dot{\phi}/H$, we obtain the Mukhanov-Sasaki action (Eq. 5.2.13):

$$
S_{\rm MS} = \frac{1}{2}\int d\tau\, d^3x\left[f'^2 - (\partial f)^2 - \frac{z''}{z}f^2\right].
\tag{10.3.8}
$$

In order to connect to the exact expression in chapter 5, we make the approximation $z''/z \simeq a''/a$, since a evolves much faster than both the inflaton $\dot{\phi}$ and the Hubble parameter H in the slow roll regime. Since we derived the Mukhanov-Sasaki action from the ADM approach, all results (e.g. the power spectrum) given in chapter 5 follow. Below, we will re-derive the mode function for the comoving curvature perturbation in a form, the so-called de Sitter limit of the mode function, that will be convenient for computing the bispectrum later on in this chapter.

10.3.3 De Sitter Limit of the Mode Functions

Here we will derive the mode function for $\mathcal{R}$ explicitly. Varying the Mukhanov-Sasaki Lagrangian with respect to f, we obtain:

$$
\left(\partial_\tau^2 + k^2 - \frac{\nu^2 - 1/4}{\tau^2}\right)f_k(\tau) = 0,
\tag{10.3.9}
$$

which is know as the Mukhanov-Sasaki equation [92]. The time dependent term z''/z can be written is terms of slow-roll parameters as follows (see Eq. 5.4.7):

$$\frac{z''}{z} = \frac{\nu^2 - 1/4}{\tau^2}, \tag{10.3.10}$$

where $\nu = 3/2 + 3\varepsilon + 3\eta/2$ to first order in slow-roll parameters.

Assuming the Bunch-Davies vacuum, we found in section 5.4 that the solution to the mode function in terms of the first Hankel function reads:

$$f_k(\tau) = \sqrt{-\tau}\left[\frac{\sqrt{\pi}}{2}\Delta^{(+)}\right] H_\nu^{(1)}(-k\tau), \tag{10.3.11}$$

conform Eq. 5.4.27 and the mode function for $\mathcal{R}$ is related simply by $\mathcal{R}_k = f_k/z$. To compute the bispectrum or three-point correlation function of $\mathcal{R}$, we will have to perform time-integrals over the mode function for $\mathcal{R}_k$. These integrals simplify significantly when the order ν of the Hankel function is integer or half-integer, which happens in the perfect de Sitter limit $(\varepsilon, \eta) \to 0$ so that $\nu = 3/2$. Driven by the computational advantage, we will from on consider the mode function in the limit $\nu \to 3/2$ and refer to those mode functions as the perfect de Sitter mode functions. However, this simplification comes with additional reduction in terms of accuracy and we discuss the induced error in section REF.

For now, we take the limit $\nu \to 3/2$ of the Hankel function $H_\nu^{(1)}(-k\tau)$, which reads:

$$\lim_{\nu \to 3/2} H_\nu^{(1)}(-k\tau) = \frac{1}{\sqrt{x^3}}\sqrt{\frac{2}{\pi}}(k\tau - i)e^{ix}, \tag{10.3.12}$$

where we defined $x \equiv -k\tau$. Consequently, the mode function $f_k(\tau)$ becomes:

$$f_k(\tau) = \frac{Ha}{\sqrt{2k^3}}(i - k\tau)e^{-ik\tau}. \tag{10.3.13}$$

Now using the relation $\tau = -(aH)^{-1}$ and the definition of the slow-roll parameter ε (with $M_{\rm pl} \equiv 1$) to recast the quantity z, we obtain the following expression for the de Sitter mode function of the comoving curvature perturbation [32, 88]:

$$\boxed{\mathcal{R}_k = \frac{iH}{\sqrt{4\varepsilon k^3}}(1 + ik\tau)e^{-ik\tau}.} \tag{10.3.14}$$

Notice that the de Sitter limit mode function still depends on the slow-roll parameter ε and hence in the perfect de Sitter limit the expression is not well-defined. It should be emphasized that this solution is referred to as the de Sitter mode function since it is obtained by setting $\nu \to 3/2$.

We will now show explicitly that the de Sitter limit of the mode function satisfies the quadratic equation of motion. Directly taking the variation of the second order Lagrangian with respect to $\mathcal{R}$, i.e. without introducing the Mukhanov-Sasaki variable f, we obtain:[6]

$$\frac{\delta\mathcal{L}_2}{\delta\mathcal{R}} \equiv \frac{d}{dt}\frac{\partial\mathcal{L}_2}{\partial\dot{\mathcal{R}}} - \frac{\partial\mathcal{L}_2}{\partial\mathcal{R}} = \frac{d}{dt}(\varepsilon a\partial^2\mathcal{R}) - \varepsilon a\partial^2\mathcal{R} = 0. \tag{10.3.16}$$

[6] Actually, this variation is with respect to the Lagrangian:

$$\mathcal{L}_2 = \frac{1}{2}\varepsilon(a^3\dot{\mathcal{R}}^2 - a(\partial\mathcal{R})^2), \tag{10.3.15}$$

which differs by a factor of 1/2 from the Lagrangian corresponding to the second order action S_2. This discrepancy is purely convential and arises since there exists no general agreement in the literature on the numeric prefactor of the second order action.

The non-trivial minus sign in the derivative $\partial\mathcal{L}_2/\partial\mathcal{R}$ is due to the variation of the derivative term:

$$\frac{\partial\mathcal{L}_2}{\partial\mathcal{R}} = -\varepsilon a\frac{\partial}{\partial\mathcal{R}}(\partial\mathcal{R})^2 = \varepsilon a\partial^2\mathcal{R}. \tag{10.3.17}$$

The solution to the above equation of motion for $\mathcal{R}$ can be written in Fourier space and conformal time, where ∂^2 is replaced by $-k^2$, as:

$$\mathcal{R}_k = \frac{iH}{\sqrt{4\varepsilon k^3}}(1+ik\tau)e^{-ik\tau}, \tag{10.3.18}$$

which is indeed the $\nu \to 3/2$ limit of the mode function.

10.3.4 Cubic Action

Finally, we will compute the third order action in perturbations, which we will use in order to find the interaction Hamiltonian. In analogy with Eq. 10.3.5, the third order action can be written as:

$$\begin{aligned} S_3 = \frac{1}{2}\int d^4x\Bigg[&ae^{\mathcal{R}}\left(1+\frac{\dot{\mathcal{R}}}{H}\right)\left(-4\partial^2\mathcal{R}-(\partial\mathcal{R})^2-2Va^2e^{2\mathcal{R}}\right)\\ &+a^3e^{3\mathcal{R}}\left(1-\frac{\dot{\mathcal{R}}}{H}+\frac{\dot{\mathcal{R}}^2}{H^2}\right)\Big(-6(H+\dot{\mathcal{R}})^2+\dot{\phi}^2\\ &+4a^{-2}e^{-2\mathcal{R}}(H+\dot{\mathcal{R}})\left(\partial_i\psi\partial_i\mathcal{R}+\partial^2\psi\right)\Big)\\ &+a^{-4}e^{-4\mathcal{R}}\Big((\partial_i\partial_j\psi)^2-(\partial^2\psi)^2-4\partial_i\partial_j\psi(\partial_{(i}\psi\partial_{j)}\mathcal{R})\Big)\Bigg], \end{aligned} \tag{10.3.19}$$

After performing numerous integrations by parts, which we will not discuss here, the final form for the third order action reads [32, 36, 88]:

$$\begin{aligned} S_3 = \int d^4x\Bigg[&a^3\varepsilon^2\mathcal{R}\dot{\mathcal{R}}^2-2a\varepsilon\dot{\mathcal{R}}(\partial_i\mathcal{R})(\partial_i\chi)+a\varepsilon^2\mathcal{R}(\partial\mathcal{R})^2\\ &+\frac{a^3}{2}\varepsilon\dot{\eta}\mathcal{R}^2\dot{\mathcal{R}}+\frac{\varepsilon}{2a}\partial_i\mathcal{R}\partial_i\chi\partial^2\chi+\frac{\varepsilon}{4a}(\partial\chi)^2\partial^2\mathcal{R}+2f(\mathcal{R})\frac{\delta\mathcal{L}_2}{\delta\mathcal{R}}\Bigg]. \end{aligned} \tag{10.3.20}$$

This form of the cubic action is in exact agreement with Eq. 169 given in [88], but differs from the expressions given in [32, 36] by a factor of 2 in from of the term proportional to $f(\mathcal{R})$. The function $f(\mathcal{R})$ is given by:

$$f(\mathcal{R}) = \frac{\eta}{4}\mathcal{R}^2+\frac{1}{H}\mathcal{R}\dot{\mathcal{R}}+\frac{1}{4a^2H^2}\Big[-(\partial\mathcal{R})^2+\partial^{-2}\left(\partial_i\partial_j(\partial_i\mathcal{R}\partial_j\mathcal{R})\right)\Big], \tag{10.3.21}$$

and encompasses all terms in the cubic action which are proportional to the quadratic equation of motion.

10.4 From Action to Hamiltonian

In order to find the interaction Hamiltonian, we will have to transform the cubic action or equivalently the corresponding $\mathcal{L}_3$ into a third order Hamiltonian, which will be the interaction Hamiltonian. In the literature [32, 36, 64], often the substitution:

$$\mathcal{H}_{\text{int}} = -\mathcal{L}_3, \tag{10.4.1}$$

where $\mathcal{L}_3$ contains only terms cubic or higher order in perturbations, is made without further comment. This prescription is not wrong, but it is not true in general. For instance, at fourth order interaction Hamiltonian is typically not simply the negative of the fourth order Lagrangian $\mathcal{L}_4$. The reason is that interaction terms may contain temporal derivative couplings, i.e. terms proportional to $\dot{\mathcal{R}}$.

As a result, the momentum conjugate π will not simply be proportional to $\dot{\mathcal{R}}$ but will contain terms quadratic and possibly higher order in $\dot{\mathcal{R}}$ as well. Therefore, it is generally not possible to find the invert the momentum conjugate:

$$\pi(\dot{\mathcal{R}}) = \frac{\partial \mathcal{L}}{\partial \dot{\mathcal{R}}}, \tag{10.4.2}$$

where $\mathcal{L}$ contains terms starting quadratic in perturbations. We assume here that terms first order in perturbations are absent: in section 8.2 it was shown that perturbative corrections to the background evolution start at quadratic order. To proceed, we assume that $\dot{\mathcal{R}}$ and π are of the same order and iteratively compute the Hamiltonian. Below, we will prove that, at least at cubic order, this approach is equivalent to making the substitution given by Eq. 10.4.1.

Consider the following cubic order perturbed Lagrangian for some perturbation φ (we drop the δ-notation for notational convenience):

$$\mathcal{L} = \frac{1}{2}\dot{\varphi}^2 - V(\varphi, \partial_i\varphi) + g_1\dot{\varphi}^3 + g_2\dot{\varphi}^2 + g_3\dot{\varphi}, \tag{10.4.3}$$

where we have not included terms linear in the perturbation since we assume that perturbative corrections to the background evolution start quadratic in perturbations. In addition, we included all terms that do not contain temporal derivative coupling in the potential $V(\varphi, \partial_i\varphi)$. The above Lagrangian can be split into a quadratic and cubic part in perturbations:

$$\begin{aligned} \mathcal{L}_2 &= \frac{1}{2}\dot{\varphi}^2 - V_0(\varphi, \partial_i\varphi), \\ \mathcal{L}_3 &= -V_{\text{int}}(\varphi, \partial_i\varphi) + g_1\dot{\varphi}^3 + g_2\dot{\varphi}^2 + g_3\dot{\varphi}. \end{aligned} \tag{10.4.4}$$

We have denoted the potential term second and third order in perturbations as V_0 and V_{int}, respectively. Notice that at quadractic order $\mathcal{L}_2$ we assume temporal derivative coupling to be absent, i.e. terms proportional to $\varphi\dot{\varphi}$ are absent:

$$\mathcal{L}_2 \not\supset f\varphi\dot{\varphi}, \tag{10.4.5}$$

which is in line with the quadratic action for the comoving curvature perturbation S_2 (Eq. 10.3.6). Then, $\dot{\varphi}$ only appears in the canonical kinetic term:

$$\mathcal{L}_2 \supset \frac{1}{2}\dot{\varphi}^2. \tag{10.4.6}$$

The above requirement constrains the prefactors $g_{1,2,3}$ of the derivative terms in the third order action. In particular, we require g_2 and g_3 to be at least first and second order in φ:

$$g_2 = \mathcal{O}(\varphi, \partial_i\varphi), \qquad g_3 = \mathcal{O}(\varphi^2, \partial_i\varphi\partial_j\varphi), \tag{10.4.7}$$

so that those terms do not contribute to the quadratic Lagrangian. The prefactor g_3 need not be constrained since this term is already third order in perturbations.

Computing the conjugate momentum π to the field φ, we find:

$$\pi(\dot{\varphi}) = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = \dot{\varphi} + 3g_1\dot{\varphi}^2 + 2g_2\dot{\varphi} + g_3 = \dot{\varphi} + \mathcal{O}(g\dot{\varphi}), \tag{10.4.8}$$

so that at first order the conjugate momentum still coincide with the time derivative $\dot{\varphi}$ and corrections start at higher order since the couplings $g_{2,3}$ start at first order in perturbations. To first order, we can invert the relation to obtain:

$$\dot{\varphi} = \pi - 3g_1\dot{\varphi}^2 - 2g_2\dot{\varphi} - g_3 + \mathcal{O}(g\dot{\varphi}). \tag{10.4.9}$$

This expression can be used in the Legendre transform of the Lagrangian, which is given by:

$$H = \int d^3x\, \mathcal{H} = \int d^3x\, (\pi\dot{\varphi} - \mathcal{L}). \tag{10.4.10}$$

In particular, the Hamiltonian density becomes:

$$\begin{aligned}\mathcal{H} &= \pi\dot{\varphi} - \mathcal{L} \\ &= \pi^2 - 3g_1\dot{\varphi}^3 - 2g_2\dot{\varphi}^2 - g_3\dot{\varphi} - \frac{1}{2}(\pi - 3g_1\dot{\varphi}^2 - 2g_2\dot{\varphi} - g_3)^2 \\ &+ V(\varphi, \partial_i\varphi) - g_1\dot{\varphi}^3 - g_2\dot{\varphi}^2 - g_3\dot{\varphi}.\end{aligned} \tag{10.4.11}$$

In the expansion of the term $\pi\dot{\varphi}$, we have set $\pi = \dot{\varphi}$ in all terms except for π^2. This is justified since corrections would only enter at higher order according to Eq. 10.4.8. Expanding the quadratic term and again replacing π by $\dot{\varphi}$ in all resulting terms except for $\pi^2/2$, we obtain:

$$\begin{aligned}\mathcal{H} = &\frac{\pi^2}{2} + V_0(\varphi, \partial_i\varphi) \\ &+ V_{\text{int}}(\varphi, \partial_i\varphi) - g_1\dot{\varphi}^3 - g_2\dot{\varphi}^2 - g_3\dot{\varphi} \\ &- \frac{g_3^2}{2} - 2g_2g_3\dot{\varphi} - 2g_2^2\dot{\varphi}^2 - 3g_1g_3\dot{\varphi}^2 - 6g_1g_2\dot{\varphi}^3 - \frac{9}{2}g_1^2\dot{\varphi}^4,\end{aligned} \tag{10.4.12}$$

We recognize the first line as the free field Hamiltonian (density) $\mathcal{H}_0$, composed out of terms second order in perturbations. The second and third line constitute third and fourth order terms, respectively, and is hence identified as the interaction Hamiltonian $\mathcal{H}_{\text{int}}$. Notice that at third order, i.e. the second line above, the interaction Hamiltonian is indeed the negative of the third order Lagrangian:

$$\mathcal{H}_{\text{int}} = V_{\text{int}}(\varphi, \partial_i\varphi) - g_1\dot{\varphi}^3 - g_2\dot{\varphi}^2 - g_3\dot{\varphi} = -\mathcal{L}_3 + \mathcal{O}(\varphi^4, \dot{\varphi}^4, \partial_i\varphi^4). \tag{10.4.13}$$

Therefore, we verified that at third order, the Hamiltonian density is the negative of the Lagrangian, as advocated.

Applying the prescription of Eq. 10.4.1, we find from Eq. 10.3.20 that the Hamiltonian density is given by:

$$\begin{aligned}\mathcal{H}_{\text{int}} = &- a^3\varepsilon^2\mathcal{R}\dot{\mathcal{R}}^2 + 2a\varepsilon\dot{\mathcal{R}}(\partial_i\mathcal{R})(\partial_i\chi) - a\varepsilon^2\mathcal{R}(\partial\mathcal{R})^2 \\ &- \frac{a^3}{2}\varepsilon\dot{\eta}\mathcal{R}^2\dot{\mathcal{R}} - \frac{\varepsilon}{2a}\partial_i\mathcal{R}\partial_i\chi\partial^2\chi - \frac{\varepsilon}{4a}(\partial\chi)^2\partial^2\mathcal{R} - 2f(\mathcal{R})\frac{\delta\mathcal{L}_2}{\delta\mathcal{R}}.\end{aligned} \tag{10.4.14}$$

Notice that the first line contains interaction terms up to quadratic order in slow roll parameters and therefore forms the leading contribution to be bispectrum. The terms in the

first line will be considered in calculating the leading bispectrum. The second line contains terms cubic or higher in slow-roll parameters, except for the first one, which is proportional to $\dot{\eta}$. However, generically one can express the time derivative $\dot{\eta}$ in terms of a function that is quadratic in slow-roll parameters [32]:

$$\dot{\eta} = \mathcal{O}(\varepsilon^2), \tag{10.4.15}$$

where ε^2 collectively denotes either ε or η. Hence, this term may be regarded as cubic in slow-roll parameters as well and therefore induces a sub-dominant contribution to the bispectrum.

Having said that, it should be noted that at the end of inflation, when we typically compute the bispectrum, both η and ε become of order unity and $\dot{\eta}$ can become even larger depending on the shape of the potential. Stated otherwise, one can construct potentials which are very flat ($\varepsilon, \eta \ll 1$) and therefore suitable for inflation, while $\dot{\eta}$ can be very large. Hence, the first term on the second line can in principle generate large NG contributions. This possibility is discussed in detail in [34].

10.5 De Sitter Limit and Maldacena's Field Redefinition

When evaluated on-shell, the variation of the second order Lagrangian with respect to the comoving curvature perturbation will zero:

$$\frac{\delta \mathcal{L}_2}{\delta \mathcal{R}} = 0, \tag{10.5.1}$$

so that all terms contained in $f(\mathcal{R})$ can be neglected. However, there is a subtle technicality concerned with the above argument [9]. The variation $\delta\mathcal{L}_2/\delta\mathcal{R}$ only vanishes when the mode function $\mathcal{R}_k$ is taken to be the de Sitter limit mode function as given by Eq. 10.3.14, i.e. when we take the limit $\nu \to 3/2$. Hence, using this form of the mode function effectively turns off the contributions of the terms contained in $f(\mathcal{R})$ to the bispectrum.

However, those terms cannot simply be omitted based on this observation, since the de Sitter limit of the mode function only solves the Mukhanov-Sasaki equation (Eq. 10.3.9) in the limit $\nu \to 3/2$ (i.e. for $(\varepsilon, \eta) \to 0$). In other words, the de Sitter mode function only *approximately* solves the Mukhanov-Sasaki equation of motion. In mathematical terms, evaluating equation of motion with the de Sitter mode functions yields zero up to, but not including, first order slow-roll parameters:

$$\left.\frac{\delta \mathcal{L}_2}{\delta \mathcal{R}}\right|_{\nu \to 3/2} = 0 + \mathcal{O}(\varepsilon, \eta), \tag{10.5.2}$$

where the limit $\nu \to 3/2$ signifies the fact that the equations of motions are evaluated using the de Sitter mode function corresponding to $(\varepsilon, \eta) \to 0$.

10.5.1 Field Redefinition

In order to resolve this subtlety, Maldacena proposed to perform a field shift or redefinition of the form [64]:

$$\mathcal{R}_n \equiv \mathcal{R} - f(\mathcal{R}), \tag{10.5.3}$$

which removes the term proportional to $f(\mathcal{R})$ from the cubic action S_3.[7] When enforcing this field redefinition in the obtained action $S_3[\mathcal{R}_n]$, we will be computing the bispectrum of $\mathcal{R}_n$ rather than $\mathcal{R}$. In contrast to $\mathcal{R}$, the shifted field does evolve outside the horizon. This is most easily seen by realizing that the function f contains terms involving the scale factor, which is emphatically not constant outside the horizon.

In addition, note that after the field redefinition, the cubic action will contain a contribution of the form:

$$S_3[\mathcal{R}_n] \supset \frac{1}{2}\int d^4x\, a^3 \varepsilon \dot{\eta} \mathcal{R}_n^2 \dot{\mathcal{R}}_n. \tag{10.5.4}$$

At the end of inflation, when we typically evaluate the bispectrum, this term would vanish when using $\mathcal{R}$ since by then all modes of interest will be outside the horizon and hence $\dot{\mathcal{R}} = 0$. However, in this case, we evaluate the action in terms of $\mathcal{R}_n$ and this argument ceases to be valid. One may argue that contribution will still be negligible since it is proportional to $\dot{\eta}$, which approaches zero in the slow-roll regime. However, at the end of inflation, the slow-roll conditions, i.e. smallness of ε and η, will not be satisfied anymore and hence $\dot{\eta}$ may well be of order one, rendering the term non-negligible.

Consequently, the three-point correlation function for $\mathcal{R}_n$ will still contain intrinsic temporal dependence. Therefore, we will have to account for this leftover time dependence at the end of the computation by shifting back to the original field $\mathcal{R}$. In particular, for a field redefinition of the form:

$$\mathcal{R} = \mathcal{R}_n + \lambda \mathcal{R}_n^2 \tag{10.5.5}$$

the leading correction in going from the bispectrum of $\mathcal{R}_n$ back to that of $\mathcal{R}$ is given by [9, 32]:

$$\begin{aligned}\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle &= \langle \mathcal{R}_n(\boldsymbol{k}_1)\mathcal{R}_n(\boldsymbol{k}_2)\mathcal{R}_n(\boldsymbol{k}_3)\rangle \\ &\quad + 2\lambda\left(\langle \mathcal{R}_n(\boldsymbol{k}_1)\mathcal{R}_n(\boldsymbol{k}_2)\rangle\langle \mathcal{R}_n(\boldsymbol{k}_1)\mathcal{R}_n(\boldsymbol{k}_3)\rangle + 2 \text{ perms.}\right) + \mathcal{O}(\eta^2 A_{\mathrm{S}}^3).\end{aligned} \tag{10.5.6}$$

From now on, we will drop the subscript n on $\mathcal{R}_n$ for notational convenience, keeping in mind that at the end we will have to apply the prescription given by the above equation.

By inspection of $f(\mathcal{R})$, we note the parameter λ will in principle be sourced by various different terms. However, to a good approximation, we can set λ to be:

$$\lambda \equiv \frac{\eta}{4}, \tag{10.5.7}$$

i.e. we consider only the first contribution to $f(\mathcal{R})$. This is justified since the other contributions to the function $f(\mathcal{R})$ contain at least one derivative with respect to space or time acting on the field. When evaluated outside the horizon, spatial gradients can be neglected and $\mathcal{R}$ approaches a constant value.

Nevertheless, even with the considerable simplification obtain by setting $\lambda = \eta/4$, the parameter η still has to be evaluated at the end of inflation, where it may attain a value of order unity. In such a scenario, the perturbative expansion in ε and η ceases to be appropriate. However, one may expect the simplification by means of Eq. 10.5.7 to remain valid, since contributions of order η^2 in Eq. 10.5.6 are accompanied by cubic factors of the power spectrum amplitude A_{S}^3 and we know from observations that:

$$A_{\mathrm{S}} = \mathcal{O}(10^{-10}), \tag{10.5.8}$$

[7] Note that at quadratic order, it does not matter whether we evaluate f using $\mathcal{R}$ or $\mathcal{R}_n$, we have: $f(\mathcal{R}) = f(\mathcal{R}_n)$ [64].

so that corrections proportional to η^2 can be neglected safely, even if η approaches unity at the end of inflation.

10.5.2 Bispectrum Contribution due to Field Redefinition

Based on the above arguments, we write Eq. 10.5.6 as follows from now on:

$$\begin{aligned}\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle &= \langle \mathcal{R}_n(\boldsymbol{k}_1)\mathcal{R}_n(\boldsymbol{k}_2)\mathcal{R}_n(\boldsymbol{k}_3)\rangle \\ &\quad + \frac{\eta}{2}\left(\langle \mathcal{R}_n(\boldsymbol{k}_1)\mathcal{R}_n(\boldsymbol{k}_2)\rangle\langle \mathcal{R}_n(\boldsymbol{k}_1)\mathcal{R}_n(\boldsymbol{k}_3)\rangle + 2 \text{ perms.}\right). \end{aligned} \tag{10.5.9}$$

We will now compute the term on the second line explicitly, i.e. we are interested in the contribution:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle_\eta \equiv \frac{\eta}{2}\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\rangle\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_3}\rangle + 2 \text{ perms.} \tag{10.5.10}$$

In other words, we compute the contribution to the bispectrum caused by the field redefinition.

We know that, in momentum space, the two-point correlation function is given by:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\rangle = (2\pi)^3\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2)P(k_1). \tag{10.5.11}$$

We know that quantity $P(k_1) \equiv |\mathcal{R}_{\boldsymbol{k}_1}|^2$ is related to the power spectrum of the comoving curvature perturbation $\mathcal{R}$ as follows (see Eq. 3.2.34):

$$P(k_1) = \frac{2\pi^2}{k_1^3}\mathcal{P}_\mathcal{R}. \tag{10.5.12}$$

In the limit of a scale-invariant power spectrum, i.e. when the spectral index n_s approaches unity, the power spectrum can be written as (Eq. 5.4.32):

$$\mathcal{P}_\mathcal{R} = \frac{H^2}{8\pi^2\varepsilon M_\mathrm{pl}^2}, \tag{10.5.13}$$

where we left the Planck mass explicitly in order to make the dimensions manifest. Using these results in Eq. 10.5.10, we obtain the following contribution:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle_\eta = (2\pi)^3\delta^{(3)}(\boldsymbol{K}) \times \frac{\eta}{2}\frac{H^4}{(4\varepsilon)^2 M_\mathrm{pl}^2}\frac{k_1^3+k_2^3+k_3^3}{k_{123}^3}, \tag{10.5.14}$$

where we traded the two momentum delta functions for one delta function over the total momentum vector $\boldsymbol{K} \equiv \boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3$ and we defined $k_{123} \equiv k_1k_2k_3$. Loosely speaking, this term connects the bispectra of the comoving curvature perturbation $\mathcal{R}$ and the redefined field $\mathcal{R}_n$ via Eq. 10.5.6. In the next section, we will show that this correction term is also closely related to the boundary terms in the cubic action, which have always been omitted in the analysis up till now.

10.6 Cubic Action and Boundary Terms

In the derived cubic action S_3, we have discarded all boundary terms, both spatial and temporal, thereby neglecting the possible contributions to non-proper boundary terms similar

to the ones appearing at quadratic order (see the third line of Eq. 10.3.7). In this section, we will examine the omitted boundary terms in more detail, following [9, 27]. As mentioned earlier, omitting spatial boundary terms is justified as in Fourier space the spatial derivatives are replaced by the momenta k and on super-horizon scales, where we evaluate the bispectrum, we can take the limit $k \to 0$. Even by explicit inclusion of the spatial boundary terms in the cubic action, they would not contribute to bispectrum as they are proportional to the total momentum $\boldsymbol{K}$ in Fourier space (which replaces the spatial derivatives in momentum space). After integrating over the overall momentum delta function, the derivative terms thus vanish. In conclusion, the vanishing behavior of boundary terms is a direct consequence of momentum conservation.

However, temporal boundary terms should be considered with more care for reasons we will explain now. Recall that we are interested in computing the three-point correlation function by means of the In-In formalism, which requires an explicit form for the interaction Hamiltonian. Evaluation of a string of operators $\mathcal{Q}(t)$, in our case the bispectrum, is then performed via Eq. 8.1.3, which is defined in terms of the interaction Hamiltonian H_{int}. Equivalently, one may compute the $\langle \mathcal{Q}(t_*) \rangle$ in the Lagrangian path integral formalism, for which the prescription reads [27]:

$$\langle \mathcal{Q}(t_*) \rangle = \int [\mathcal{D}\phi_+ \mathcal{D}\phi_-] \; \mathcal{Q}(t_*) \; e^{iS[\phi_+ - \phi_-]} \; \delta(\phi_+(t_*) - \phi_-(t_*)). \tag{10.6.1}$$

For current purposes, this form is more convenient as it allows for direct implementation of the action, instead of the interaction Hamiltonian. As enforced by the temporal delta function, the domain of integration in field space is restricted to fields $\phi_\pm$ that coincide at the time of interest t_*. Hence, boundary operators that contain no temporal derivatives solely produce a phase [27]. Those phases cancel between the $+$ contour (from ∞^+ to t_*) and the $-$ contour (from t_* to ∞^-).

Let us now consider the implications of the path integral prescription for temporal boundary terms, which we omitted up till now. For the cubic action, the temporal boundary terms are given by [9, 27]:

$$\begin{aligned} \partial S_3 = \int d^4x \, \frac{d}{dt} \Big[& -9a^3 H \mathcal{R}^3 + \frac{a}{H}(1-\varepsilon)\mathcal{R}(\partial\mathcal{R})^2 - \frac{1}{4aH^3}(\partial\mathcal{R})^2 \partial^2 \mathcal{R} \\ & - \frac{a^3}{H}\varepsilon \mathcal{R}\dot{\mathcal{R}}^2 + \frac{1}{2aH^2}\mathcal{R}\left(\partial_i\partial_j\mathcal{R}\partial_i\partial_j\chi - \partial^2\mathcal{R}\partial^2\chi\right) \\ & - \frac{\eta a}{2}\mathcal{R}^2\partial^2\chi - \frac{1}{2aH}\mathcal{R}\left((\partial_i\partial_j\chi)^2 - (\partial^2\chi)^2\right) \Big]. \end{aligned} \tag{10.6.2}$$

Since the operators in the first line involve no time derivatives, they evaluate to cancelling phases on the $+$ and $-$ contours. Therefore, we can safely omit the terms on the first line. In addition, all terms containing spatial derivatives can be neglected as well.

However, note that the temporal delta function in the path integral formulation of the In-In formalism (Eq. 10.6.1) in no way enforces the time derivatives of the fields $\phi_\pm$ to coincide at time t_*. Hence, we will have to more careful with boundary terms involving time derivatives of fields. This procedure leaves us with the following boundary terms that should be considered with care:

$$\partial S_3 \supset -\int d^4x \, \frac{d}{dt}\left[\frac{a^3}{H}\varepsilon\mathcal{R}\dot{\mathcal{R}}^2 + \frac{1}{2}a\eta\mathcal{R}^2\partial^2\chi\right]. \tag{10.6.3}$$

Now, we argue, based on a field redefinition, that the boundary term proportional to $\mathcal{R}\dot{\mathcal{R}}^2$ can be neglected as well. More generally, the analysis to be performed below shows that all boundary terms quadratic in $\dot{\mathcal{R}}$ (or higher) can be neglected. Consider a field redefinition of the form [27]:

$$\mathcal{R} = \pi + \pi\dot{\pi}. \tag{10.6.4}$$

Schematically, the three-point correlation function of the comoving curvature perturbation transforms as follows under the considered field redefinition:

$$\langle \mathcal{R}\mathcal{R}\mathcal{R}\rangle = \langle \pi\pi\pi\rangle + 3\langle \pi\pi\rangle\langle \pi\dot{\pi}\rangle + \mathcal{O}(\dot{\pi}^2). \tag{10.6.5}$$

Computing the two-point correlation function for π, we find that it is equivalent to the one for $\mathcal{R}$, that is:

$$\langle \pi_{\boldsymbol{k}_1}\pi_{\boldsymbol{k}_2}\rangle = (2\pi)^3\,\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2)\frac{2\pi^2}{k_1^3}\left(\frac{H_*^2}{8\pi^2\varepsilon M_{\rm pl}^2}\right), \tag{10.6.6}$$

which is manifestly time independent. This implies that the time derivative term can vanishes:

$$\langle \pi\dot{\pi}\rangle = 0. \tag{10.6.7}$$

Therefore, the middle term in Eq. 10.6.5 vanishes and the bispectra of $\mathcal{R}$ and π are equivalent up to the order $\mathcal{O}(\dot{\pi}^2)$. Notice that the invoked field redefinition will inevitably produce bulk (i.e. non-boundary) terms proportional to the linear equations of motion $\delta\mathcal{L}_2/\delta\pi$. However, those terms will not contribute to the bispectrum since equations of motion will be evaluated using the de Sitter mode functions.

Using these results, we can show explicitly that the boundary term quadratic in $\dot{\mathcal{R}}$ indeed only contributes at quadratic and higher order in $\dot{\pi}$, therefore can be neglected in the computation of the bispectrum. Let us substitute the field redefinition into the relevant boundary term:

$$\partial S_3 \supset -\int d^4x\,\frac{d}{dt}\left(\frac{a^3}{H}\varepsilon\mathcal{R}\dot{\mathcal{R}}^2\right) = -\int d^4x\,\frac{d}{dt}\left(\frac{a^3}{H}\varepsilon\pi\dot{\pi}^2 + \mathcal{O}(\dot{\pi}^3)\right). \tag{10.6.8}$$

This boundary term is indeed of order $\dot{\pi}^2$ and hence by virtue of Eq. 10.6.5 this term does not affect the three-point function of the comoving curvature perturbation $\mathcal{R}$ at leading order.

In conclusion, we are left with a single temporal boundary term that cannot simply be neglected:

$$\partial S_2 \supset -\int d^4x\,\frac{d}{dt}\left(\frac{a\eta}{2}\mathcal{R}^2\partial^2\chi\right) = -\int d^3x\frac{d}{dt}\left(\frac{a^3\varepsilon\eta}{2}\mathcal{R}^2\dot{\mathcal{R}}\right). \tag{10.6.9}$$

Below, we will explicitly compute the contribution of this non-negligible boundary term to the bispectrum by means of the In-In formalism.

10.6.1 Bispectrum Contribution due to Boundary Term

Since we have proven that the interaction Hamiltonian is simply the negative of the Lagrangian at third order, the interaction Hamiltonian for the non-vanishing boundary term is given by:

$$\mathcal{H}_{{\rm int},\eta} = \frac{d}{dt}\left[\frac{1}{2}a^3\eta\varepsilon\mathcal{R}^2\dot{\mathcal{R}}\right] = \frac{1}{a}\frac{d}{d\tau}\left[\frac{1}{2}a^2\eta\varepsilon\mathcal{R}^2\mathcal{R}'\right], \tag{10.6.10}$$

where we switched to conformal time instead of cosmic time. Now using Eq. 8.5.4 with the operator $\mathcal{Q} = \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}$, we find:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle_\eta = 2\mathrm{Im}\Bigg\langle \Bigg[\int a d\tau'\, d^3x\, \mathcal{R}_{\boldsymbol{k}_1}(\tau)\mathcal{R}_{\boldsymbol{k}_2}(\tau)\mathcal{R}_{\boldsymbol{k}_3}(\tau) \frac{1}{a}\frac{d}{d\tau}\left(\frac{1}{2}a^2\eta\varepsilon\right)\int_{\boldsymbol{q}_1\boldsymbol{q}_2\boldsymbol{q}_3} \mathcal{R}_{\boldsymbol{q}_1}(\tau')\mathcal{R}_{\boldsymbol{q}_2}(\tau')\partial_\tau\mathcal{R}_{\boldsymbol{q}_3}(\tau')\, e^{-i\boldsymbol{Q}\cdot\boldsymbol{x}}\Bigg)\Bigg]\Bigg\rangle, \tag{10.6.11}$$

where we have defined $\boldsymbol{Q} = \boldsymbol{q}_1 + \boldsymbol{q}_2 + \boldsymbol{q}_3$ and we switched to conformal time in the space-time volume measure as well. The momenta integrations and the factor of $e^{-i\boldsymbol{Q}\cdot\boldsymbol{x}}$ comes from transforming the interaction Hamiltonian to Fourier space as well. Notice that the temporal integral can be dropped together with the time derivative in the interaction Hamiltonian: this makes boundary terms straightforward to evaluate. Hence, we can write:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle_\eta = a^2\eta\varepsilon\ \mathrm{Im}\Bigg[\int d^3x \int_{\boldsymbol{q}_1\boldsymbol{q}_2\boldsymbol{q}_3} \langle 0|\mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\mathcal{R}_{\boldsymbol{q}_1}\mathcal{R}_{\boldsymbol{q}_2}\partial_\tau\mathcal{R}_{\boldsymbol{q}_3}|0\rangle\, e^{-i\boldsymbol{Q}\cdot\boldsymbol{x}}\Bigg]. \tag{10.6.12}$$

To proceed, we will first evaluate the vacuum expectation value of the string of field operators. According to Wick's theorem, the result is given by:

$$\begin{aligned}
&\langle 0|\mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\mathcal{R}_{\boldsymbol{q}_1}\mathcal{R}_{\boldsymbol{q}_2}\partial_\tau\mathcal{R}_{\boldsymbol{q}_3}|0\rangle \\
&\quad = 2\left([\mathcal{R}^+_{\boldsymbol{k}_1}, \mathcal{R}^-_{\boldsymbol{q}_1}][\mathcal{R}^+_{\boldsymbol{k}_2}, \mathcal{R}^-_{\boldsymbol{q}_2}][\mathcal{R}^+_{\boldsymbol{k}_3}, \partial_\tau\mathcal{R}^-_{\boldsymbol{q}_3}] + 2 \text{ perms.}\right) \\
&\quad = 2\big((f_{k_1}f^*_{q_1})(f_{k_2}f^*_{q_2})(f_{k_3}\partial_\tau f^*_{q_3}) \times (2\pi)^9\delta^{(3)}(\boldsymbol{k}_1 - \boldsymbol{q}_1)\delta^{(3)}(\boldsymbol{k}_2 - \boldsymbol{q}_2)\delta^{(3)}(\boldsymbol{k}_3 - \boldsymbol{q}_3) \\
&\quad + 2 \text{ perms.}\big)
\end{aligned} \tag{10.6.13}$$

where the additional factor of two is included due to the fact that we obtain two identical terms from the choice to contract either with $\mathcal{R}_{\boldsymbol{q}_1}$ or $\mathcal{R}_{\boldsymbol{q}_2}$ and we denoted the de Sitter mode functions (Eq. 10.3.14) by f_k rather than $\mathcal{R}_k$.

By insertion of the above result into Eq. 10.6.12, and performing the integration over $\boldsymbol{x}$, we obtain:

$$\begin{aligned}
\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle_\eta = 2a^2\eta\varepsilon\ (2\pi)^3\delta^{(3)}(\boldsymbol{Q}) \times \mathrm{Im}\Bigg[\int_{\boldsymbol{q}_1\boldsymbol{q}_2\boldsymbol{q}_3} (f_{k_1}f^*_{q_1})(f_{k_2}f^*_{q_2})(f_{k_3}\partial_\tau f^*_{q_3}) \\
\times (2\pi)^9\delta^{(3)}(\boldsymbol{k}_1 - \boldsymbol{q}_1)\delta^{(3)}(\boldsymbol{k}_2 - \boldsymbol{q}_2)\delta^{(3)}(\boldsymbol{k}_3 - \boldsymbol{q}_3) + 2 \text{ perms.}\Bigg].
\end{aligned} \tag{10.6.14}$$

Now integrating over using the de Sitter mode function and its first time derivative:

$$f_k = \frac{iH}{\sqrt{4\varepsilon k^3}}(1 + ik\tau)\, e^{-ik\tau}, \qquad \partial_\tau f_k = \frac{iH}{\sqrt{4\varepsilon k^3}}k^2\tau e^{-ik\tau}, \tag{10.6.15}$$

we find that:

$$\mathrm{Im}\left[(f_{k_1}f^*_{q_1})(f_{k_2}f^*_{q_2})(f_{k_3}\partial_\tau f^*_{q_3})\right] = \left(\frac{H^2}{4\varepsilon}\right)^3 \frac{1}{k^3_{123}}(1 + k_1^2\tau^2)(1 + k_2^2\tau^2)k_3^3\tau^2 \tag{10.6.16}$$

Finally using that $\tau = -1/aH$, and taking the limit $\tau \to 0$, the following expression is obtained for the contribution to the bispectrum due to the boundary term:

$$\begin{aligned}\lim_{\tau\to 0}\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle_\eta &= (2\pi)^3\delta^{(3)}(\boldsymbol{K}) \\ &\quad\times \lim_{\tau\to 0}\left[2\eta\varepsilon(aH)^2\left(\frac{H^2}{4\varepsilon}\right)^3\frac{1}{k_{123}^3}(1+k_1^2\tau^2)(1+k_2^2\tau^2)k_3^3\tau^2 + 2 \text{ perms.}\right] \\ &= (2\pi)^3\delta^{(3)}(\boldsymbol{K}) \times \frac{\eta}{2}\frac{H^4}{(4\varepsilon)^2}\frac{k_1^3+k_2^3+k_3^3}{k_{123}^3}, \end{aligned} \tag{10.6.17}$$

which coincides with the correction to the bispectrum (Eq. 10.5.14) when performing the field redefinition. This explicit computation shows that when treating boundary terms properly in the calculation, the result will be equivalent to neglecting all boundary terms and performing the field redefinition as proposed by Maldacena [64]. Below, we will show in detail why both approaches are equivalent.

10.6.2 Equivalent Approaches

Above, we have considered two approaches. First, following Maldacena [64], we performed a field redefinition (Eq. 10.5.3) and discarded all boundary terms. At the end of the calculation, we noted that a correction term (Eq. 10.5.14) should be added to the bispectrum of $\mathcal{R}_n$ in order to get the bispectrum of $\mathcal{R}$. Secondly, we treated all boundary terms of the cubic action explicitly and found that the contribution to the bispectrum due to those boundary terms (Eq. 10.6.17) is equivalent to the correction term related to the field redefinition.

When performing the field redefinition, we discarded all boundary terms. When *not* omitting the boundary terms, a third order boundary term is generated from the second order action, which cancels the boundary term (Eq. 10.6.9) of the third order action that cannot be neglected outside the horizon. This shows that both approaches are indeed equivalent. Here, we will derive this statement in detail.

The time derivative of the square of $\mathcal{R}$ in terms of the redefined field is given by:

$$\dot{\mathcal{R}}^2 = \dot{\mathcal{R}}_n^2 + 2\dot{\mathcal{R}}_n\dot{f}_n + \mathcal{O}(\dot{f}_n^2), \tag{10.6.18}$$

where the subscript on the function f denotes that it is evaluated using $\mathcal{R}_n$. Furthermore, the gradient term $(\partial\mathcal{R})^2$ is given by:

$$(\partial\mathcal{R})^2 = (\partial\mathcal{R}_n)^2 + 2\partial\mathcal{R}_n\partial f_n + \mathcal{O}((\partial f_n)^2). \tag{10.6.19}$$

Then, the quadratic action in terms of $\mathcal{R}_n$ becomes:

$$S_2[\mathcal{R}] = S_2[\mathcal{R}_n] + 2\int d^4x\, \varepsilon a^3\left[\dot{\mathcal{R}}_n\dot{f}_n - \frac{1}{a^2}\partial\mathcal{R}_n\partial f_n\right] + \mathcal{O}(\dot{f}_n^2, (\partial f_n)^2). \tag{10.6.20}$$

Performing the following partial integrations:

$$\begin{aligned}\int d^4x\, 2\varepsilon a^3\dot{\mathcal{R}}_n\dot{f}_n &= \int d^4x\, \partial_t(2a^3\varepsilon f_n\dot{\mathcal{R}}_n) - \int d^4x\, f_n\partial_t(2\varepsilon a^3\dot{\mathcal{R}}_n), \\ \int d^4x\, \varepsilon a\partial\mathcal{R}_n\partial f_n &= \int d^4x\, \varepsilon a f_n\partial\mathcal{R}_n - \int d^4x\, \varepsilon a(\partial^2\mathcal{R}_n)f_n, \end{aligned} \tag{10.6.21}$$

the expression for $S_2[\mathcal{R}]$ can be rewritten as follows:

$$S_2[\mathcal{R}] = S_2[\mathcal{R}_n] - \int d^4x\, 2f_n \frac{\delta S_2}{\delta \mathcal{R}_n} + \int d^4x\, \partial_t(2a^3\varepsilon f_n \dot{\mathcal{R}}_n). \tag{10.6.22}$$

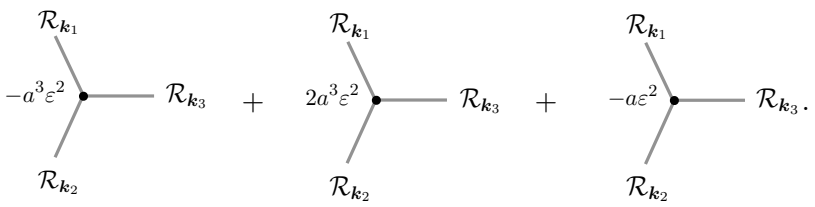
Comparing this result with the third order action S_3 (Eq. 10.3.20), we find that the term proportional to the quadratic equations of motion cancels the term proportional to the equations of motions in S_3. Finally, the last temporal boundary term exactly cancels the only boundary term in ∂S_3 that does not vanish outside the horizon (see Eq. 10.6.9). Therefore, we have shown that the field redefinition is equivalent to the approach in which the boundary terms are left explicitly.

10.7 Leading Bispectrum for Single-Field Inflation

Now that we have dealt with all the technical subtleties, we will finally derive the leading order bispectrum of single-field inflation. We will take the leading order terms in the interaction Hamiltonian and compute the three-point correlation function. Explicitly, the interaction Hamiltonian that we will consider in this section is:

$$\mathcal{H}_{\text{int}} = -a^3\varepsilon^2\mathcal{R}\dot{\mathcal{R}}^2 + 2a^3\varepsilon^2(\partial_i\mathcal{R})(\partial_i\partial^{-2}\dot{\mathcal{R}})\mathcal{R} - a\varepsilon^2\mathcal{R}(\partial\mathcal{R})^2, \tag{10.7.1}$$

which is, upon the substitution of $\partial^2\chi = \varepsilon a^2\dot{\mathcal{R}}$, equivalent to the first line of Eq. 10.4.14. Note that the above Hamiltonian is actually in terms of $\mathcal{R}_n$, rather than $\mathcal{R}$, but we have dropped the subscript for notational convinience. Taking the coupling constants to be the prefactors of each interaction term, we are computing the following three point function in terms of Feynman diagrams:

$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle =$ 

(10.7.2)

We will denote the contributions due to those three diagrams as:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle = \langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle_{\mathcal{R}\dot{\mathcal{R}}^2} + \langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle_{\dot{\mathcal{R}}\partial\mathcal{R}\partial\chi} + \langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle_{\mathcal{R}(\partial\mathcal{R})^2}. \tag{10.7.3}$$

In this section, we will explicitly compute the contribution of those terms to the bispectrum for single-field inflation by means of the In-In formalism. The final result will be:

$$\begin{aligned}\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle &= (2\pi)^3\delta^{(3)}(\boldsymbol{K}) \\ &\times \frac{H^4}{(4\varepsilon)^2}\frac{1}{k_{123}^3}\left[\frac{\eta}{2}\sum_i k_i^3 + \frac{\varepsilon}{2}\left(-\sum_i k_i^3 + \sum_{i\neq j} k_i k_j^2 + \frac{8}{K}\sum_{i>j}(k_i k_j)^2\right)\right],\end{aligned} \tag{10.7.4}$$

which is in exact agreement with Maldacena's result [64]. Below, we will derive this result in detail by considering each interaction term separately.

10.7.1 Interaction Term $\mathcal{R}\dot{\mathcal{R}}^2$

Here we will consider the contribution of the first term in the interaction Hamiltonian to the bispectrum. For notational convenience, we will use a prime to denote the three-point correlation function without an overall Dirac delta function for momentum conservation. For the bispectrum $\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle_A$ due to interaction term A, we define:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle_A \equiv (2\pi)^3\delta^{(3)}(\boldsymbol{K})\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle'_A \tag{10.7.5}$$

Hence we can write the contribution to the bispectrum as:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle'_{\mathcal{R}\dot{\mathcal{R}}^2} = 2\,\mathrm{Re}\left[i\int^{t_*} dt'\ \langle 0|\mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}(t)(a^3\varepsilon^2)\mathcal{R}_{\boldsymbol{q}_1}\dot{\mathcal{R}}_{\boldsymbol{q}_2}\dot{\mathcal{R}}_{\boldsymbol{q}_3}|0\rangle\right]. \tag{10.7.6}$$

The $\boldsymbol{k}$-modes of $\mathcal{R}$ are evaluated at the time of interest τ_* and therefore do not enter the temporal integration. Transforming the above integral to conformal time, taking the moment of evaluated to be the limit $\tau_* \to 0$ and performing the Wick contractions, we obtain:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle'_{\mathcal{R}\dot{\mathcal{R}}^2} = 4\,\mathrm{Re}\left[if_{k_1}f_{k_2}f_{k_3}(0)\int_{\infty^+}^{0} d\tau'(a\varepsilon)^2\left(f^*_{k_1}\partial_\tau f^*_{k_2}\partial_\tau f^*_{k_3}\right)\right], \tag{10.7.7}$$

where the additional factor of two is due to the fact that we get two identical terms from the choice to contract with either of the two $\partial_\tau\mathcal{R}$. Inserting the mode functions and the first temporal derivative (Eqs. 10.6.15), we find:

$$\begin{aligned}\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle'_{\mathcal{R}\dot{\mathcal{R}}^2} = 4\,\mathrm{Re}\Bigg[&\frac{H^3}{\sqrt{64\varepsilon^3k_{123}^3}}\\ &\int_{-\infty^+}^{0}\left(\frac{\varepsilon}{H\tau}\right)^2\frac{iH^3}{\sqrt{64\varepsilon^3k_{123}^3}}(k_2k_3)^2(1-ik_1\tau)\tau^2\ e^{+iK\tau}\\ &+2\text{ perms.}\Bigg].\end{aligned} \tag{10.7.8}$$

Notice that we substituted $\tau = -1/aH$ and the two factors of τ^2 neatly cancel. Performing the temporal integral including the $i\epsilon$-prescription to make the integral converge in the early time limit, we find that:

$$\int_{-\infty^+}^{0} d\tau'\ (1-ik_1\tau')e^{iK\tau'} = \left[-\frac{ik_1}{K^2}-\frac{i}{K}-\frac{k_1\tau}{K}\right]e^{iK\tau}\Bigg|_{-\infty^+}^{0} = -\frac{ik_1}{K^2}-\frac{i}{K}. \tag{10.7.9}$$

Substituting this result taking only the real part and writing out the explicit permutations, we obtain the contribution:

$$\begin{aligned}\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle'_{\mathcal{R}\dot{\mathcal{R}}^2} &= \frac{H^4}{(4\varepsilon)^2}\frac{\varepsilon}{k_{123}^3}\left[\frac{(k_2k_3)^2k_1}{K}+\frac{(k_2k_3)^2}{K}\right]+\ 2\text{ perms.}\\ &= \frac{H^4}{(4\varepsilon)^2}\frac{\varepsilon}{k_{123}^3}\left[\frac{1}{K}\sum_{i<j}(k_ik_j)^2+\frac{k_{123}}{K^2}\sum_{i<j}k_ik_j\right].\end{aligned} \tag{10.7.10}$$

10.7.2 Interaction Term $\dot{\mathcal{R}}\partial\mathcal{R}\partial\chi$

Next we consider the interaction term:

$$\mathcal{H}_{\text{int}} \supset 2a^3\varepsilon^2(\partial_i\mathcal{R})(\partial_i\partial^{-2}\dot{\mathcal{R}})\dot{\mathcal{R}}. \tag{10.7.11}$$

In Fourier space, the inverse gradient ∂^{-2} equals $-k^{-2}$. The Fourier counterpart of the above expression is:

$$\mathcal{H}_{\text{int}} \supset a^3\varepsilon^2\mathcal{R}_{\boldsymbol{q}_1}\partial_\tau\mathcal{R}_{\boldsymbol{q}_2}\partial_\tau\mathcal{R}_{\boldsymbol{q}_3}(\boldsymbol{q}_1\cdot\boldsymbol{q}_2)\left[\frac{1}{q_2^2}+\frac{1}{q_3^2}\right], \tag{10.7.12}$$

where the vector dot product arises from the $\partial_i\mathcal{R}\partial_i(\partial^{-2}\dot{\mathcal{R}})$ term and we distributed the factor of two over a term proportional to k_2^{-2} and a term proportional to k_3^{-2}. The dot product can be rewritten by using the fact that the total vector momentum is zero ($\boldsymbol{Q}=0$) and hence:

$$(\boldsymbol{q}_1+\boldsymbol{q}_2)^2=(-\boldsymbol{q}_3)^2=q_3^2=k_1^2+2\boldsymbol{k}_1\cdot\boldsymbol{k}_2+k_2^2, \qquad \boldsymbol{q}_1\cdot\boldsymbol{q}_2=\frac{1}{2}(k_3^2-k_1^2-k_2^2). \tag{10.7.13}$$

Performing computations as for the previous contributions, one eventually finds the intermediate expression:

$$\begin{aligned}\langle\mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle'_{\dot{\mathcal{R}}\partial\mathcal{R}\partial\chi} = &-\frac{H^4}{(4\varepsilon)^2}\frac{\varepsilon}{k_{123}^3}(k_2k_3)^2\left[\frac{1}{k_2^2}+\frac{1}{k_3^2}\right](\boldsymbol{k}_1\cdot\boldsymbol{k_2})\\ &\times\text{Re}\left[i\int_{-\infty^+}^0 d\tau'(1-ik_1\tau)e^{+iK\tau}+\text{ 2 perms.}\right].\end{aligned} \tag{10.7.14}$$

The real part of the temporal integral is given by:

$$\text{Re}\left[i\int_{-\infty^+}^0 d\tau'(1-ik_1\tau)e^{+iK\tau}\right]=\frac{K+k_1}{K^2} \tag{10.7.15}$$

so that we get the following expression for the contribution of the considered interaction term:

$$\begin{aligned}\langle\mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle'_{\dot{\mathcal{R}}\partial\mathcal{R}\partial\chi} &= -\frac{H^4}{(4\varepsilon)^2}\frac{\varepsilon}{k_{123}^3}\left[\frac{(k_2k_3)^2(K+k_1)}{K^2}\left(\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_2^2}+\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_3}{k_3^2}\right)+\text{ 2 perms.}\right]\\ &= -\frac{H^4}{(4\varepsilon)^2}\frac{\varepsilon}{k_{123}^3}\left[\sum_i\left(\frac{2k_{123}}{K^2}k_i^2-k_i^3-\frac{k_i^4}{K}\right)+\frac{6}{K}\sum_{i<j}(k_ik_j)^2\right].\end{aligned} \tag{10.7.16}$$

10.7.3 Interaction Term $\mathcal{R}(\partial\mathcal{R})^2$

Finally, the last interaction term that will be considered is given in momentum space by:

$$\mathcal{H}_{\text{int}} \supset a\varepsilon^2(\boldsymbol{q}_1\cdot\boldsymbol{q}_2)\mathcal{R}_{\boldsymbol{q}_1}\mathcal{R}_{\boldsymbol{q}_2}\mathcal{R}_{\boldsymbol{q}_3}. \tag{10.7.17}$$

The contribution of this term to the bispectrum can be calculated in a similar way to the two terms above and yields:

$$\langle\mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle'_{\mathcal{R}(\partial\mathcal{R})^2}=\frac{H^4}{(4\varepsilon)^2}\frac{\varepsilon}{k_{123}^3}\left[\sum_{i\neq j}k_i^2k_j+\frac{1}{K}\sum_i k_i^4-k_{123}\left(1+\frac{1}{K^2}\sum_i k_i^2\right)\right]. \tag{10.7.18}$$

10.7.4 Leading Bispectrum and Equivalence with Maldacena's Result

Adding the above contributions as well as the correction term due to the field redefinition (Eq. 10.5.14), we find the total leading order bispectrum for single-field inflation to be:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle = (2\pi)^3\delta^{(3)}(\boldsymbol{K})\frac{1}{M_{\rm pl}^4}\frac{H^4}{(4\varepsilon)^2}\frac{1}{k_{123}^3}\times \\ \left[\frac{\eta}{2}\sum_i k_i^3 + \frac{\varepsilon}{2}\left(-\sum_i k_i^3 + \sum_{i\neq j} k_i k_j^2 + \frac{8}{K}\sum_{i>j}(k_i k_j)^2\right)\right], \tag{10.7.19}$$

where we have re-instated the Planck mass for completeness: below it will be set to unity again. The above equation provides the main result of the part on non-gaussianity in single-field scenario. We will now show that this expression is equivalent to the one given by Maldacena [64]:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle = (2\pi)^3\delta^{(3)}(\boldsymbol{K})\frac{H^8}{\dot\phi^4}\frac{1}{8k_{123}^3}\mathcal{A}, \tag{10.7.20}$$

where the momentum-dependent function $\mathcal{A}$ is given by:

$$\mathcal{A} = 2\frac{\phi}{H\dot\phi}\sum_i k_i^3 + \frac{\dot\phi^2}{2H^2}\sum_i k_i^3 + \frac{1}{2}\frac{\dot\phi^2}{H^2}\sum_{i\neq j} k_i k_j^2 + \frac{1}{2}\frac{\dot\phi^2}{H^2}\frac{8}{K}\sum_{i>j}(k_i k_j)^2. \tag{10.7.21}$$

To show the equivalence between out result (Eq. 10.7.19) and Maldacena's form, we use the fact that we have set the Planck mass to unity, so that the prefactor of Maldacena's result can be written as:

$$(2\pi)^3\delta^{(3)}(\boldsymbol{K})\frac{H^8}{\dot\phi^4}\frac{1}{8k_{123}^3} = (2\pi)^3\delta^{(3)}(\boldsymbol{K})\frac{H^4}{(4\varepsilon)^2}\frac{1}{2k_{123}^3}. \tag{10.7.22}$$

In addition, the momentum function $\mathcal{A}$ can be written as:

$$\mathcal{A} = (-2\delta+\varepsilon)\sum_i k_i^3 + \varepsilon\sum_{i\neq j} k_i k_j^2 + \frac{8\varepsilon}{K}\sum_{i>j}(k_i k_j)^2, \tag{10.7.23}$$

where we used the slow-roll parameters defined in section 2.6. In particular, the first combination can be rewritten as:

$$\varepsilon - 2\delta = \eta - \varepsilon. \tag{10.7.24}$$

Substituting this expression and combining the prefactor and the rewritten form of $\mathcal{A}$, we find that Maldacena's result is equivalent to our form as given in Eq. 10.7.19.

10.7.5 Visual Representation of Momentum Dependence

Here we will discuss the momentum dependence of the leading order bispectrum in more detail. In the limit of scale-invariant power spectra, we know that the bispectrum $B_{\mathcal{R}}(k_1,k_2,k_3)$ (Eq. 7.2.3) behaves as follows under a rescaling of the momenta ($\boldsymbol{k}_i \to \lambda\boldsymbol{k}_i$):

$$B_{\mathcal{R}}(\lambda k_1,\lambda k_2,\lambda k_3) = \lambda^{-6}B_{\mathcal{R}}(k_1,k_2,k_3). \tag{10.7.25}$$

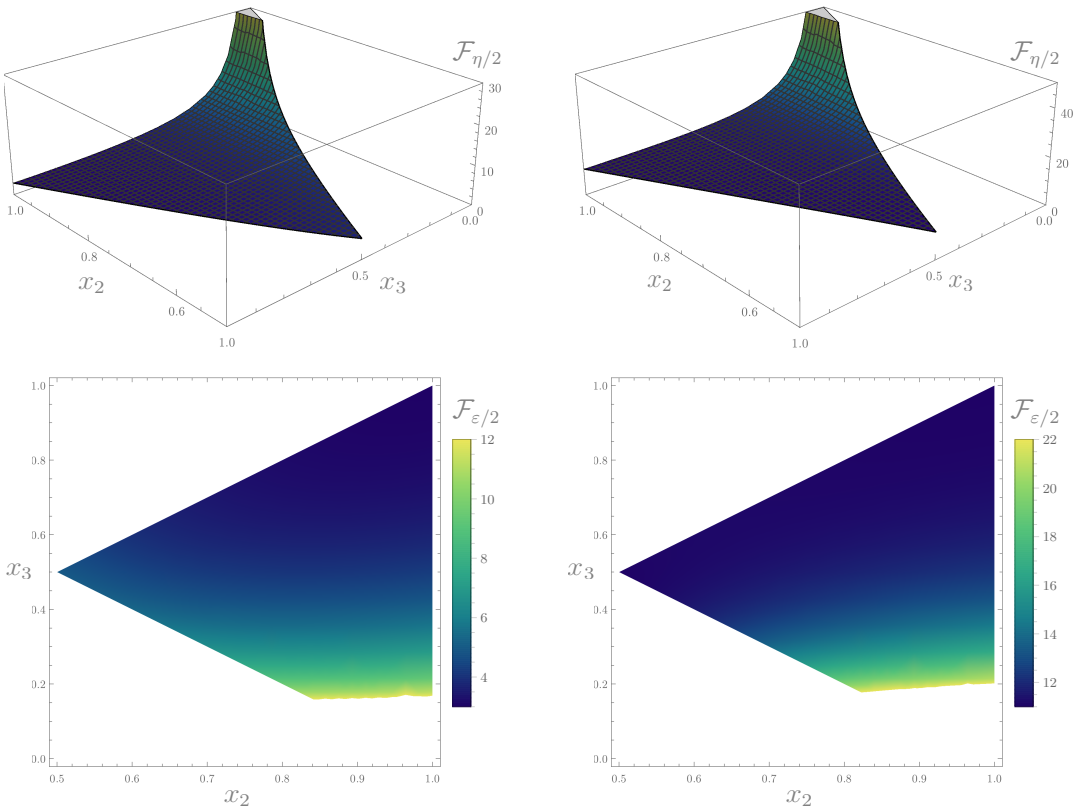


Figure 10.4: Three dimensional as well as density plots of the quantities $x_2^2 x_3^2\ \mathcal{F}(1, x_2, x_3)_A$ with $A = (\eta/2, \varepsilon/2)$. Notice that both functions maximize in the limit $x_3 \to 0$ and $x_2 \to 1$, corresponding to the squeezed limit $k_3 \to 0$ and $k_1 \simeq k_2$.

On account of rotational invariance, the number of independent variables reduces to two, for instance the ratios $x_{2,3} \equiv k_{2,3}/k_1$ and $x_1 = 1$. We define the momentum function $\mathcal{F}(k_1, k_2, k_3)$ in relation to the bispectrum:

$$B_{\mathcal{R}}(k_1, k_2, k_3) \equiv \frac{1}{M_{\rm pl}^2} \frac{H^4}{(4\varepsilon)^2} \mathcal{F}(k_1, k_2, k_3), \tag{10.7.26}$$

so that from Eq. 10.7.19 we find:

$$\mathcal{F}(k_1, k_2, k_3) = \frac{1}{k_{123}^3} \left[\frac{\eta}{2} \sum_i k_i^3 + \frac{\varepsilon}{2} \left(-\sum_i k_i^3 + \sum_{i \neq j} k_i k_j^2 + \frac{8}{K} \sum_{i>j} (k_i k_j)^2 \right) \right]. \tag{10.7.27}$$

Notice that this expression for the momentum dependence still involves slow-roll parameters, which are not related to the momentum dependence. Therefore, we will factor out these dependencies by defining:

$$\begin{aligned} \mathcal{F}_{\eta/2} &\equiv \frac{1}{k_{123}^3} \sum_i k_i^3, \\ \mathcal{F}_{\varepsilon/2} &\equiv \frac{1}{k_{123}^3} \left[-\sum_i k_i^3 + \sum_{i \neq j} k_i k_j^2 + \frac{8}{K} \sum_{i>j} (k_i k_j)^2 \right], \end{aligned} \tag{10.7.28}$$

so that we obtain two contributions to the momentum function, related to the parts involving η and ε, respectively.

Now we aim to obtain some insight on the momentum dependence via a visual approach. Following [11], we will plot the combinations:

$$x_2^2 x_3^2\, \mathcal{F}(1, x_2, x_3)_A, \qquad A = (\eta/2, \varepsilon/2). \tag{10.7.29}$$

In order not to show the equivalent configurations twice, we set the momentum function to zero when the condition $1 - x_2 \leq x_3 \leq x_2$ is not satisfied. In Fig. 10.4, we show the functions contained in Eq. 10.7.29 in terms of a three-dimensional plots as well as a density plots both as function of the ratios x_2 and x_3. Observe that both the functions $\mathcal{F}_{\eta/2}$ and $\mathcal{F}_{/2}$ become maximized in the squeezed limit (in this case $k_3 \to 0$ and $k_1 \simeq k_2$). This behavior is very closely related to the consistency relation derived in section 7.5, as we will show in the next section.

10.8 Consistency Relation

By inspection of the leading bispectrum in the form of Eq. 10.6.17, we find conclude that the bispectrum peaks in the squeezed limit where one of the modes has vanishing momentum. Taking k_1 to be vanishing or at least much smaller than the other two, we define the squeezed limit as:

$$k_1 \ll k_2 \simeq k_3. \tag{10.8.1}$$

In this limit, we find that $K = 2k_2$ and $\Sigma_i k_i^3 = k_2^3 + k_3^3$, so that the momentum-dependent part of the bispectrum can be written as:

$$\frac{\eta}{2}\sum_i k_i^3 + \frac{\varepsilon}{2}\bigg(-\sum_i k_i^3 + \sum_{i\neq j} k_i k_j^2 + \frac{8}{K}\sum_{i>j}(k_i k_j)^2\bigg) \simeq \frac{\eta + 2\varepsilon}{2}\sum_i k_i^3. \tag{10.8.2}$$

The squeezed limit of the bispectrum can thus reads (setting the Planck mass to unity again):

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle = (2\pi)^3\delta^{(3)}(\boldsymbol{K})\frac{H^4}{(4\varepsilon)^2}\frac{1}{k_{123}^3}\bigg(\frac{\eta + 2\varepsilon}{2}\sum_i k_i^3\bigg). \tag{10.8.3}$$

Notice that the momentum dependence of the squeezed limit bispectrum is of the same form as the momentum-dependence in the local bispectrum (Eq. 7.4.7). Hence, equating those the bispectra, we can extract an explicit expression for $f_{\rm NL}^{\rm local}$ in terms of slow-roll parameters. Inserting, the power spectrum $P(k_1)$, which is given by (using $M_{\rm pl} = 1$):

$$P(k_1) = \frac{H^2}{4\varepsilon k_1^3}, \tag{10.8.4}$$

into the local bispectrum Eq. 7.4.7, we obtain:

$$\langle \mathcal{R}_{\boldsymbol{k}_1}\mathcal{R}_{\boldsymbol{k}_2}\mathcal{R}_{\boldsymbol{k}_3}\rangle = (2\pi)^3\delta^{(3)}(\boldsymbol{K}) \times \frac{6}{5} f_{\rm NL}^{\rm local}\frac{H^4}{(4\varepsilon)^2}\frac{1}{k_{123}^3}\sum_i k_i^3. \tag{10.8.5}$$

Comparing this equation to the squeezed limit of the leading bispectrum Eq. 10.8.3, we find that $f_{\rm NL}^{\rm local}$ is given by:

$$f_{\rm NL}^{\rm local} = \frac{5}{12}(\eta + 2\varepsilon) = \frac{5}{12}(1 - n_{\rm s}), \tag{10.8.6}$$

where the second equality is obtained on account of Eq. 5.4.33. This result coincides with the consistency relation derived in section 7.5 (Eq. 7.5.2).

Part V

Non-Gaussianity in the Multi-Field Scenario

Chapter 11

Multi-Field Inflation and Quantum Effects

"The absence of evidence is not the evidence of absence."

— Carl Sagan

From now on, we relax the assumption that one scalar field, the inflaton, was dynamically relevant during inflation and consider a theory in which n scalar fields are responsible for the inflationary expansion. From the perspective of high energy theories, the existence of multiple light scalar fields is more natural than only one. The reason is that, within the framework of string theory, multiple scalar fields arise naturally in the dimensional compactification of the higher dimensional theory to obtain a 4-dimensional effective theory [5, 18, 20].

In this chapter, we will discuss the classical dynamics and quantum effects of multi-field inflation and emphasize the differences as compared to the single-field scenario. In particular, we will derive the equations of motion for n canonical scalar fields minimally coupled to gravity and examine the evolution of the comoving curvature perturbation on large scales for multiple fields. Furthermore, we will study the two-field scenario and compute the power spectrum and spectral index. In the next chapter, we will extend the framework developed here by including the computation of the bispectrum for multi-field inflation.

11.1 Multi-Field Action and Equations of Motion

In the remaining part of this work, we will consider a system of n canonical scalar fields minimally coupled to gravity. For scenarios with non-canonical kinetic terms and non-minimal couplings to gravity, see e.g. [53, 58, 83]. The action corresponding to the considered system reads:

$$S = \int d^4x\sqrt{-g}\left[\frac{1}{2}M_{\rm pl}^2 R - \frac{1}{2}G_{IJ}g^{\mu\nu}\partial_\mu\phi^I\partial_\nu\phi^J - V(\phi^I)\right], \tag{11.1.1}$$

where G_{IJ} is the metric of the field space spanned by the scalar fields ϕ^I. Notice that we have adopted a slightly different notation as compared to the multi-field Lagrangian (Eq. 4.8.40) in section 4.8.3; we have traded the labels (α) for I, adopted the convention that indices are contracted using G_{IJ} and take repeated indices to be implicitly summed over. Following [76], we consider the field space metric to be flat for simplicity, at least in the vicinity of the classical background trajectory through field space, so that we can write:

$$G_{IJ} = \delta_{IJ}. \tag{11.1.2}$$

For a more general discussion in which the field space geometry is left arbitrary, see [58].

11.1.1 Klein-Gordon Equations of Motion

The equations of motion for the scalar fields can be derived by varying the action with respect to ϕ^I and setting the result equal to zero: $\delta S = 0$. Due to minimal coupling, the gravitational sector does not involve ϕ^I and varying with respect to the Einstein-Hilbert term in the action consequently yields zero. We denote the second and third terms as $S_\phi^{(1,2)}$. Variation of those terms yields:

$$\begin{aligned}\delta S_\phi^{(1)} &= -\int d^4x\sqrt{-g}G_{IJ}g^{\mu\nu}\partial_\mu\phi^I\partial_\nu\phi^J,\\ \delta S_\phi^{(2)} &= -\int d^4x\sqrt{-g}\partial_K V(\phi^I)\delta\phi^K,\end{aligned} \tag{11.1.3}$$

where we used the fact that the field space metric $G_{IJ} = \delta_{IJ}$ is independent of the fields so that $\partial_K G^{IJ} = 0$. Integration the term $S_\phi^{(2)}$ by parts, we obtain:

$$\delta S_\phi^{(1)} = \int d^4x\partial_\nu(G_{IJ}\sqrt{-g}g^{\mu\nu}\partial_\mu\phi^I)\delta\phi^J, \tag{11.1.4}$$

where we omitted the boundary term proportional to $\delta\phi^I$, since it vanishes by the assumption the fluctuation tends to zero at the boundary of the space-time manifold. Requiring that the variation of the action vanishes for any $\delta\phi^I$, we find that the equation of motion for the field δ^I is given by:

$$\frac{1}{\sqrt{-g}}\partial_\nu(G_{IJ}\sqrt{-g}g^{\mu\nu}\partial_\mu\phi^I) = \partial_J V(\phi^I). \tag{11.1.5}$$

Contracting the field indices on the left-hand-side, i.e. $\phi_J \equiv G_{IJ}\phi^J$, and using the flat FRW metric Eq. 1.2.5 to compute the metric and metric determinant, we arrive at the Klein-Gordon equation for the scalar fields:

$$\ddot{\phi}_I + 3H\dot{\phi}_I + \partial_I V(\phi^J) = 0. \tag{11.1.6}$$

11.1.2 Energy-Momentum Tensor and Friedmann Equations

Now we consider the energy-momentum tensor of the system to derive the expressions for the pressure and energy density. Applying the prescription given in Eq. 2.6.1 to derive $T^{\mu\nu}$, we find:

$$T^{\mu\nu} = G_{IJ}\partial^\mu\phi^I\partial^\nu\phi^J - g^{\mu\nu}\left[\frac{1}{2}G_{IJ}\partial_\beta\phi^I\partial^\beta\phi^J + V(\phi^I)\right]. \tag{11.1.7}$$

The energy density and pressure of the matter can be found via the relations $\rho = T^{00}$ and $T^{ij} = P\delta^{ij}$, yielding:

$$\rho = T^{00} = \frac{1}{2}G_{IJ}\dot{\phi}^I\dot{\phi}^J + V(\phi^I), \qquad P = \frac{1}{2}G_{IJ}\dot{\phi}^I\dot{\phi}^J - V(\phi^I). \tag{11.1.8}$$

Now that we have the expressions for the energy density and pressure we can easily derive the Friedmann equations to be:

$$\begin{aligned}H^2 &= \frac{\rho}{3M_{\rm pl}^2} = \frac{1}{3M_{\rm pl}^2}\left(\frac{1}{2}\dot{\phi}_I\dot{\phi}^I + V(\phi^I)\right),\\ \frac{\ddot{a}}{a} &= H^2 + \dot{H} = -\frac{1}{6M_{\rm pl}^2}(3P+\rho) = -\frac{1}{3M_{\rm pl}^2}\left(\dot{\phi}_I\dot{\phi}^I - V(\phi^I)\right).\end{aligned} \tag{11.1.9}$$

Combining the two Friedmann equations, we find that we can write the time-derivative and field-derivative of the Hubble parameter H read:

$$\dot{H} = -\frac{1}{2M_{\rm pl}^2}\dot{\phi}_I\dot{\phi}^I, \qquad \partial_I H = \frac{\dot{H}}{\dot{\phi}^I} = -\frac{1}{2M_{\rm pl}^2}\dot{\phi}_I, \tag{11.1.10}$$

where in the first equality of the second equation we assumed monotonicity. The equations derived above will be used often in the remaining part of this chapter.

11.1.3 Slow-Roll Approximation

In exact analogy with the single-field scenario, we define the slow-roll parameters for the multi-field system and assume those parameters to be small, thereby restricting to the slow-roll approximation. In the slow-roll limit, the Klein-Gordon equation and the first Friedmann equation become:

$$3H\dot{\phi}_I + V_I = 0, \qquad H^2 = \frac{V}{3M_{\rm pl}^2}, \tag{11.1.11}$$

where we have used that the potential term dominates the kinetic term so that $\rho \simeq V$ in the first Friedmann equation. Combining the above two equations, we can derive the following expression for the time-derivative of the fields:

$$\frac{\dot{\phi}_I}{H} = -M_{\rm pl}^2\frac{V_I}{V}. \tag{11.1.12}$$

Therefore, in the slow-roll approximation, the background fields evolve down along the gradient of the potential V_I.

In analogy with the single-field case, the slow-roll approximation can be characterized by the slow-roll parameters. Setting the Planck mass to unity ($M_{\rm pl} \equiv 1$), the first slow-roll parameter can be written as:

$$\varepsilon^{IJ}\frac{\dot{\phi}^I\dot{\phi}^J}{2H^2} = \frac{2\partial^I H\partial^J H}{H^2} \equiv \varepsilon^I\varepsilon^J, \tag{11.1.13}$$

so that $\varepsilon^{IJ} \equiv \varepsilon^I\varepsilon^J$ and ε^I is defined as:

$$\varepsilon^I = \frac{\dot{\phi}^I}{\sqrt{2}H} = \frac{\sqrt{2}\partial^I H}{H}. \tag{11.1.14}$$

The total slow-roll parameter $\varepsilon \equiv -\dot{H}/H^2$ can be obtained by contracting ε^{IJ} with the field space metric: $\varepsilon \equiv G_{IJ}\varepsilon^{IJ}$. Except for models which contain finely tuned cancellations between the different ε^{IJ}, we expect in the slow-roll approximation that $|\varepsilon^{IJ}| \ll 1$. In terms of the total slow-roll parameter, we expect:

$$\varepsilon^{IJ} = \mathcal{O}(\varepsilon/n), \qquad \varepsilon^I = \mathcal{O}(\sqrt{\varepsilon/n}). \tag{11.1.15}$$

In order for the inflationary epoch to sustain for a sufficiently long period (to solve the flatness and horizon problems), we require the second slow-roll parameter η to be small as well. The multi-field generalization of η is denoted as η^{IJ} and is related to the time-derivative of ε^{IJ} in the following way:

$$\dot{\varepsilon}^{IJ} = 2\varepsilon H(\varepsilon^{IJ} - \eta^{IJ}), \qquad \eta^{IJ} = \frac{\ddot{\phi}^I\dot{\phi}^J + \dot{\phi}^I\ddot{\phi}^J}{4H\dot{H}}. \tag{11.1.16}$$

In the single-field scenario, η^{IJ} reduces to the usual parameter η, i.e. $\eta^{\phi\phi} = -\ddot{\phi}/H\dot{\phi}$. Finally, we introduce the a third slow-roll parameter associated with the potential:

$$\tilde{\eta}_{IJ} = \frac{\partial_I \partial_J V}{3H^2}, \tag{11.1.17}$$

This is comparable to the single-field potential slow-roll parameter $\eta_{\rm v} \equiv 2\varepsilon - \eta/2$, generalizing this relation to the multi-field case, we find:

$$\varepsilon^{IJ} + \eta^{IJ} = \tilde{\eta}_{MN} \frac{G^{M(I}\varepsilon^{J)N}}{\varepsilon}, \tag{11.1.18}$$

Notice that η^{IJ} and $\tilde{\eta}^{IJ}$ are of the same order in slow-roll.

Finally, we compute the following time-derivatives in terms of slow-roll parameters:

$$\begin{aligned} \frac{\ddot{\phi}_I}{H} &= \varepsilon \dot{\phi}_I - \tilde{\eta}_{IJ}\dot{\phi}^J, \\ \frac{\dot{\varepsilon}_{IJ}}{H} &= 4\varepsilon\varepsilon_{IJ} - \varepsilon_{IK}\varepsilon^K_J - \tilde{\eta}_{IK}\varepsilon^K_J \\ \frac{\dot{\varepsilon}}{H} &= 4\varepsilon^2 - 2\varepsilon_{IJ}\tilde{\eta}^{IJ}. \end{aligned} \tag{11.1.19}$$

Those relation will be relevant in later section of this chapter and are therefore listed here.

11.2 Evolution of the Comoving Curvature Perturbation

In exact analogy with the single-field scenario, we use the comoving curvature perturbation $\mathcal{R}$, or curvature perturbation on slices of uniform energy density ζ, in order to relate perturbations in the scalar fields, denoted as $\delta\phi^I$, to late time observables. On super-horizon scales, the variables $\mathcal{R}$ and ζ still coincide. This follows directly from the proof based on the field-equations approach presented in section 6.1, which does not rely on the specific matter content sourcing $\mathcal{R}$ and ζ, therefore applying to the multi-field scenario as well. In this section, we will consider the comoving curvature perturbation $\mathcal{R}$.

11.2.1 Multi-Field Expression for the Comoving Curvature Perturbation

First, a direct relation between the comoving curvature perturbation and the quantum fluctuations in the scalar fields will be constructed. From its definition for scalar perturbations (Eq. D.4.7), $\mathcal{R}$ reads:

$$\mathcal{R} = \Psi - H(v + B). \tag{11.2.1}$$

In order to rewrite this expression, we note that the combination $v + B$ can be related to the perturbation δT^0_i in the energy-momentum tensor (Eq. D.4.15):

$$\delta T^0_i \equiv \partial_i\, \delta q = (\rho + P)\partial_i(v + B), \tag{11.2.2}$$

where we defined the momentum perturbation δq. In terms of δq, the comoving curvature perturbation $\mathcal{R}$ can thus be rewritten as:

$$\mathcal{R} = \Psi - \frac{H}{\rho + P}\delta q. \tag{11.2.3}$$

For multiple fields, we know that the perturbation in the energy-momentum tensor (Eq. 4.8.42) is given by:

$$\delta T_i^0 = -\dot{\phi}_I \partial_i \delta\phi^I, \tag{11.2.4}$$

from which we can easily infer the momentum perturbation to be $\delta q = -\dot{\phi}_I \delta\phi^I$. The expression for $\mathcal{R}$ now becomes:

$$\mathcal{R} = \Psi + H\frac{\dot{\phi}_I \delta\phi^I}{\dot{\phi}_J \dot{\phi}^J}, \tag{11.2.5}$$

where we used the fact that the sum of the energy density and pressure of the system of fields equals $\dot{\phi}_I \dot{\phi}^I$. As we will be working mostly in the flat gauge, defined by $\Psi = 0$, it is convenient to introduce the gauge-invariant Mukhanov-Sasaki variable Q (Eq. D.4.4):

$$Q^I = \delta\phi^I + \frac{\dot{\phi}^I}{H}\Psi, \tag{11.2.6}$$

which indeed equals the fluctuations $\delta\phi^I$ in the flat gauge. Inserting this relation, we find that $\mathcal{R}$ can be connected to the scalar field fluctuations $\delta\phi^I$ in a gauge invariant way as follows:

$$\mathcal{R} = H\frac{\dot{\phi}_I Q^I}{\dot{\phi}_J \dot{\phi}^J}. \tag{11.2.7}$$

11.2.2 Evolution Outside the Horizon

Now we examine the evolution of the comoving curvature perturbation sourced by multiple scalar fields. Based on the energy-momentum approach, which applies to the multi-field scenario as well, we found in section 6.3 that the evolution $\mathcal{R}$ on super-horizon scales is given by:

$$\dot{\mathcal{R}} = -3c_s^2 H\mathcal{S}, \tag{11.2.8}$$

where the isocurvature perturbations are governed by $\mathcal{S}$, which reads (Eq. 4.6.11):

$$\mathcal{S} = H\left(\frac{\delta P}{\dot{P}} - \frac{\delta\rho}{\dot{\rho}}\right). \tag{11.2.9}$$

Therefore, examining the evolution of the comoving curvature perturbation outside the horizon amounts to computing $\mathcal{S}$ in the multi-field scenario. The fact that $\mathcal{S}$ is non-zero in the multi-field case is a direct consequence of the multiple trajectories of the fields in field space.

The expressions for the energy density and pressure fluctuations are given by (see section 4.8.3):

$$\begin{aligned}\delta\rho &= \dot{\phi}_I \delta\dot{\phi}^I - \Phi\dot{\phi}_I\dot{\phi}^I + V_I\delta\phi^I,\\ \delta P &= \dot{\phi}_I \delta\dot{\phi}^I - \Phi\dot{\phi}_I\dot{\phi}^I - V_I\delta\phi^I.\end{aligned} \tag{11.2.10}$$

Using the continuity equation, the equation of state $w = \rho/P$ and the speed of sound $c_s^2 = \dot{P}/\dot{\rho}$, we find that we can write the denominators in $\mathcal{S}$ as:

$$\dot{\rho} = -3H(1+w)\rho, \qquad \dot{P} = -3H(1+w)c_s^2\rho. \tag{11.2.11}$$

Hence, we can write $\mathcal{S}$ as:

$$\mathcal{S} = H\left(\frac{\dot{\rho}\delta P - \dot{P}\delta\rho}{\dot{P}\dot{\rho}}\right) = \frac{\dot{\rho}\delta P - \dot{P}\delta\rho}{9H(1+w)^2 c_s^2 \rho^2}. \tag{11.2.12}$$

Inserting the expressions for $\delta\rho$ and δP as well, we find that $\dot{\mathcal{R}}$ is given by:

$$\dot{\mathcal{R}} = -\frac{H}{(1+w)\rho}\Big[(c_s^2-1)(\dot{\phi}_I\delta\dot{\phi}^I - \Phi\dot{\phi}_I\dot{\phi}^I) + (c_s^2+1)V_I\delta\phi^I\Big]. \tag{11.2.13}$$

Finally, using the fact that we can write the sound speed as $c_s^2 = -1 - \ddot{H}/3H\dot{H}$ [57], the final expression for the time derivative of $\mathcal{R}$ becomes:

$$\dot{\mathcal{R}} = -\frac{\ddot{H}}{6\dot{H}}\delta\rho - \frac{H}{\dot{H}}(\dot{\phi}_I\delta\dot{\phi}^I - \Phi\dot{\phi}_I\dot{\phi}^I), \tag{11.2.14}$$

where we re-instated the expression for $\delta\rho$ to obtain a neat form. Notice that this expression is not suppressed in the limit $k/aH \to 0$. Hence in the multi-field scenario, evolution of the comoving curvature curvature perturbation outside the horizon is indeed present.

11.3 Quantum Effects

Here, we will consider the quantum fluctuations in the scalar fields ϕ^I, which are written as $Q^I \equiv \delta\phi^I$ in the spatially flat gauge, as specified by $\Psi = 0$. To study the quantum effects of this multi-field system, we will anticipate a result from the next chapter, where the action for quadratic fluctuations is derived using the ADM formalism. The result is given by (Eq. 12.4.1):

$$S_2 = \frac{1}{2}\int d^4x\, a^3\left[\delta_{IJ}\dot{Q}^I\dot{Q}^J - \frac{1}{a^2}\delta_{IJ}\partial Q^I\partial Q^J - \mathcal{M}_{IJ}Q^IQ^J\right], \tag{11.3.1}$$

where we have introduced the notation $\partial Q_I\partial Q^I \equiv \delta ij\partial_i Q_I\partial^j Q^I$. The so-called mass matrix $\mathcal{M}_{IJ}$ is defined as:

$$\mathcal{M}_{IJ} = \partial_I\partial_J V - \frac{1}{a^3}\frac{d}{dt}\left(\frac{a^3}{H}\dot{\phi}_I\dot{\phi}_J\right). \tag{11.3.2}$$

Notice that the mass matrix is principle not diagonal and hence couples different fields to each other. The equation of motion resulting from this action in momentum space (but omitting the Fourier labels) is given by:

$$\partial_{\tau\tau}f^I + \left(\left[k^2 - \frac{a''}{a}\right]\delta^I_J + a^2\mathcal{M}^I_J\right)f^J = 0, \tag{11.3.3}$$

where we have introduced conformal time and the canonical variable $f^I \equiv aQ^I$, in analogy with the single-field case. In the next chapter, we will show that the mass-matrix can be written to first order in slow-roll parameters as:

$$\mathcal{M}_{IJ} = H^2(3\tilde{\eta}_{IJ} - \varepsilon_{IJ}) + \mathcal{O}(\varepsilon^2), \tag{11.3.4}$$

where for later convenience we also define $\mathcal{W}_{IJ} \equiv \mathcal{M}_{IJ}/H^2$ Using this result as well as the following expressions for a''/a and $\mathcal{H}^2$:

$$\frac{a''}{a} = \mathcal{H}^2(2-\varepsilon) + \mathcal{O}(\varepsilon^2), \qquad \mathcal{H}^2 = \frac{1}{\tau^2}(1+2\varepsilon) + \mathcal{O}(\varepsilon^2), \tag{11.3.5}$$

we can rewrite the equation of motion for the canonical variable f^I as:

$$\partial_{\tau\tau}f^I + \left(k^2 - \frac{2}{\tau^2}\right)f^I = \frac{1}{\tau^2}\mathcal{C}^I_J f^J, \tag{11.3.6}$$

where we defined $\mathcal{C}_{IJ} \equiv 3\varepsilon\delta_{IJ} - \mathcal{W}_{IJ}$.

As mentioned before, the off-diagonal components of $\mathcal{C}_{IJ}$ will couple the evolution of the different field fluctuations to each other. However, since $\mathcal{C}_{IJ}$ is real and symmetric, we can diagonalize it by performing a rotation in field space. After the rotation, we will obtain independently evolving fluctuations which we denote as γ^I. The rotation matrix that diagonalizes the re-defined version of the mass-matrix $\mathcal{C}_{IJ}$ as U_{IJ}, so that:

$$\gamma^I = U^I_J f^J, \qquad U^{-1}\mathcal{C}U = \mathrm{diag}(\lambda_1, \ldots, \lambda_n), \tag{11.3.7}$$

where λ_I are the eigenvalues of the matrix $\mathcal{C}_{IJ}$. Since U is a rotation matrix, and in particular a representation of $SO(n)$, it satifies $U^{\mathrm{T}} = U^{-1}$ and $U_{IL}U_{LJ} = \delta_{IJ}$.

Acting with U on the equation of motion, i.e. diagonalizing it, we find:

$$\partial_{\tau\tau}\gamma^I + \left(k^2 - \frac{2}{\tau^2}\right)\gamma^I = \frac{1}{\tau^2}\lambda_I\gamma^I, \tag{11.3.8}$$

where there is no sum over the index I on the right side. Taking all terms to the left hand side and defining the terms proportional to γ^I/τ^2 to be equal to $\nu_I^2 - 1/4$, we find:

$$\partial_{\tau\tau}\gamma^I + \left(k^2 - \frac{\nu_I^2 - 1/4}{\tau^2}\right)\gamma^I = 0, \tag{11.3.9}$$

where there is again no sum over the field index and to first order in slow-roll parameters $\nu_I = 3/2 + \lambda_I/3$. This is exactly the same equation as we derived for the mode function f in the single-field scenario (Eq. 5.4.8). Assuming all the fields start out in the Bunch-Davies initial state, the mode function reads (equivalent to Eq. 5.4.27)

$$\gamma^I(\tau) = \sqrt{-\tau}\left[\frac{\sqrt{\pi}}{2}\Delta_I^{(+)}\right] H_\nu^{(1)}(-k\tau), \tag{11.3.10}$$

where $\Delta_I^{(+)} = \exp(+i\pi(\nu_I + 1/2)/2)$.

The above solution to the mode function is only valid in case the slow-roll parameters are small and time-independent. This approximation ceases to be valid near the end of inflation. Therefore, we switch from the rescaled fluctuation γ^I to the comoving curvature perturbation, which are related via Eq. 11.2.7:

$$\mathcal{R} = H\frac{\dot{\phi}_I Q^I}{\dot{\phi}_J\dot{\phi}^J} = -\frac{H^2}{\rho + P}\frac{V_I Q^I}{V}. \tag{11.3.11}$$

From this expression for $\mathcal{R}$, where $Q^I = af^I$, we find the comoving curvature perturbation is only sourced by the components of the field perturbations, which are tangential to the background solution, i.e. the direction of the vector $\dot{\phi}^I$. In the slow-roll approximation, this corresponds to the path of steepest descent along the potential.

In general, however, this direction does not coincide with any of the eigenvectors λ_I. Consequently, we have to perform yet another rotation in field space. We rotate the original field space basis:

$$Q^I = \begin{pmatrix} Q_1 \\ \vdots \\ Q_n, \end{pmatrix} \tag{11.3.12}$$

using the rotation matrix S to a new basis:

$$\tilde{Q}^I = \begin{pmatrix} Q_\sigma \\ \vdots \\ Q_{s(n-1)} \end{pmatrix} = S^{-1} \begin{pmatrix} Q_1 \\ \vdots \\ Q_n. \end{pmatrix} \tag{11.3.13}$$

Here we have defined an adiabatic field perturbation Q_σ, parallel to the classical trajectory, in the following way:

$$Q_\sigma \equiv \frac{V_I Q^I}{|\nabla V|}, \tag{11.3.14}$$

where $|\nabla V| \equiv \sum_I V_I$. The other $n-1$ field perturbations are called the entropy perturbations as they are constructed to be orthogonal to the background trajectory in field space.

In the above analysis, we did not consider the background fields, for which we have to take into account that the rotation matrix need not be constant over the field space. Formally speaking, the rotation as induced by S is a local rotation of the basis for the field perturbations, instead of a global rotation. To overcome this subtlety, we rotate the temporal derivatives of the background fields, rather than the fields themselves, yielding:

$$\begin{pmatrix} \dot{\sigma} \\ \dot{s}_1 \\ \vdots \\ s_{n-1} \end{pmatrix} = S^{-1} \begin{pmatrix} \dot{\phi}_1 \\ \dot{\phi}_2 \\ \vdots \\ \dot{\phi}_n \end{pmatrix}, \tag{11.3.15}$$

where σ is the background complement of the adiabatic perturbation Q_σ and the entropic fields, orthogonal to the background field space trajectory, are denoted as s_i with $i = 2, \ldots, n-1$.

11.4 Power Spectra for Two-Field Inflation

Here, we will apply the formalism developed in the previous sections to the case of two-field inflation in order to derive the power spectra and associated spectral indices. The analysis will be to first order in slow-roll parameters. We consider the fields, denoted by φ and χ, whose dynamics are governed by the action:

$$S \int d^4x \sqrt{-g} \left[\frac{1}{2} M_{\rm pl}^2 R - \frac{1}{2} g^{\mu\nu} (\partial_\mu \varphi \partial_\nu \varphi + \partial_\nu \chi \partial_\mu \chi) - V(\varphi, \chi) \right], \tag{11.4.1}$$

where the potential $V(\varphi, \sigma)$ is assumed to be suitable to provide a period of slow-roll inflation but is otherwise left unspecified.

11.4.1 Adiabatic and Entropy Perturbations

The direction of the background solution in the two-dimensional field space can be characterized by means of the field space angle θ, which is given by:

$$\tan\theta \equiv \frac{\dot{\chi}}{\dot{\varphi}}. \tag{11.4.2}$$

In order to make the meaning of the adiabatic and entropy perturbations explicit, we perform a rotation in field space. The adiabatic field σ is defined as the integrated path length along the background trajectory, which may be written as the function $\chi(\varphi)$. Now, the adiabatic field can be defined as the integrated path length along the background trajectory:

$$\sigma = \int_{\phi_i}^{\phi_f} d\phi \sqrt{1 + (\partial_\phi \chi)^2} = \int_{\phi_i}^{\phi_f} d\phi \sqrt{1 + \tan^2\theta}. \tag{11.4.3}$$

The initial field value ϕ_i is arbitrary, since only changes in σ are relevant. The entropy field s is defined as the orthogonal distance in field space away from the background trajectory, which thus vanishes by definition at background level: $s = \dot{s} = \ddot{s} = 0$. This rotation in field space defines a coordinate system (σ, s). We assume the field space to be flat for simplicity. Otherwise the introduced coordinate system is solely valid in the vicinity of the considered background trajectory.

As mention above for the n-field case, the time derivatives of the fields $\dot{\varphi}$ and $\dot{\chi}$ can be related to the adiabatic and entropy fields $\dot{\sigma}$ and s via the rotation matrix $S(\theta)$, yielding:

$$\begin{pmatrix} \dot{\sigma} \\ \dot{s} \end{pmatrix} = S^{-1}(\theta) \begin{pmatrix} \dot{\varphi} \\ \dot{\chi} \end{pmatrix}, \qquad S(\theta) \equiv \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{11.4.4}$$

The rotation matrix $S(\theta)$ constitutes a representation of the two-dimensional special orthogonal group $SO(2)$, so that $S^T = S^{-1}$ and $\det S = +1$. On account of the above transformation equation between the two bases, we find the following equations by using the fact that at background level $\dot{s} = 0$:

$$\dot{\sigma} = \dot{\varphi}\cos\theta + \dot{\chi}\sin\theta, \qquad \dot{\varphi}\sin\theta = \dot{\chi}\cos\theta. \tag{11.4.5}$$

From the inverse rotation and the fact that for the entropic field $\dot{s} = 0$,

$$\begin{pmatrix} \dot{\varphi} \\ \dot{\chi} \end{pmatrix} = S(\theta) \begin{pmatrix} \dot{\sigma} \\ \dot{s} \end{pmatrix}, \tag{11.4.6}$$

we find that we can write the square of the the time derivative of the adiabatic field and the second time derivative as:

$$\dot{\sigma}^2 = \dot{\varphi}^2 + \dot{\chi}^2, \qquad \ddot{\sigma} = \ddot{\varphi}\cos\theta + \ddot{\chi}\sin\theta. \tag{11.4.7}$$

The potential $V(\varphi, \chi)$ exists everywhere in the field space and its gradient may be expressed in the (φ, χ) or (σ, s) basis. The potential gradients in the different bases are related as:

$$\begin{pmatrix} V_\varphi \\ V_\chi \end{pmatrix} = S(\theta) \begin{pmatrix} V_\sigma \\ V_\chi \end{pmatrix}, \tag{11.4.8}$$

and the inverse transformation. Using the above relation for the time-derivatives of the field φ and χ, the time derivative of the potential reads:

$$\dot{V} = V_\varphi \dot{\varphi} + V_\chi \dot{\chi} = V_\sigma \dot{\sigma}. \tag{11.4.9}$$

In a similar manner, we can rotate the second field derivatives of the potential into the (σ, s) basis:

$$[V_{I'J'}] \equiv S^{-1}[V_{IJ}]S, \tag{11.4.10}$$

or explicitly:

$$\begin{pmatrix} V_{\sigma\sigma} & V_{\sigma s} \\ V_{\sigma s} & V_{ss} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} V_{\varphi\varphi} & V_{\varphi\chi} \\ V_{\varphi\chi} & V_{\chi\chi} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{11.4.11}$$

yielding the components:

$$\begin{aligned} V_{\sigma\sigma} &\equiv V_{\varphi\varphi}\cos^2\theta + 2\cos\theta\sin\theta V_{\varphi\chi} + \sin^2\theta V_{\chi\chi}, \\ V_{ss} &\equiv V_{\varphi\varphi}\cos^2\theta - 2\cos\theta\sin\theta V_{\varphi\chi} + \cos^2\theta V_{\chi\chi}, \\ V_{\sigma s} &\equiv -\sin\theta\cos\theta V_{\varphi\varphi} + (\cos^2\theta - \sin^2\theta)V_{\varphi\chi} + \cos\theta\sin\theta V_{\chi\chi}. \end{aligned} \tag{11.4.12}$$

11.4.2 Exact Field Equations

With the machinery to rotate to a basis of adiabatic and entropy field available, we now consider the field equations governing the background evolution of the field. In the basis (φ, χ), the Klein-Gordon equations are given in a neat form by:

$$\begin{pmatrix} \ddot{\varphi} \\ \ddot{\chi} \end{pmatrix} + 3H \begin{pmatrix} \dot{\varphi} \\ \dot{\chi} \end{pmatrix} + \begin{pmatrix} V_\varphi \\ V_\chi \end{pmatrix} = 0. \tag{11.4.13}$$

This equation can be transformed into the basis (σ, s) by acting from the left on the above system of equations by $S^{-1}(\theta)$ to give:

$$\ddot{\sigma} + 3H\dot{\sigma} + V_\sigma = 0, \tag{11.4.14}$$

and we know that at background level $\ddot{s} = 0$, so that the field s is dynamically irrelevant. Instead, the rol of the other degree of freedom is taken by the field space angle θ, which is related to the slope of the potential in the entropy direction V_s by:

$$\dot{\theta} = -\frac{V_s}{\dot{\sigma}}. \tag{11.4.15}$$

To obtain the second-order differential equation for θ we have to differentiate the above relation with respect to time again. Notice that the coordinates (σ, s) can in general be curved, even in the field space itself is flat.[1] Therefore, it proves convenient to work in the (φ, χ) basis and we obtain:

$$V_s = -V_\varphi \sin\theta + V_\chi \cos\theta, \qquad \dot{V}_s = V_\sigma\dot{\sigma} - \dot{\theta}V_\sigma. \tag{11.4.16}$$

Using this result we may properly take the time derivative of Eq. 11.4.15 to obtain the second order equation for θ, yielding:

$$\ddot{\theta} - 3H\dot{\theta} + V_{\sigma s} - 2\frac{V_\sigma}{\dot{\sigma}}\dot{\theta} = 0. \tag{11.4.17}$$

[1] This is consequence of the fact that the coordinate system (σ, s) is defined locally, i.e. at every point along the background trajectory curve and hence can vary its orientation if the curve is not straight, thereby constituting a so-called curvy-linear coordinate system.

11.4.3 Two-Field Slow-Roll Approximation

Using the slow-roll approximation for both fields, so that we neglect the second order time derivatives $\ddot{\varphi}$ and $\ddot{\chi}$ and we can write the field equations as (Eq. 11.1.12):

$$\frac{\dot{\varphi}}{H} = -M_{\rm pl}^2 \frac{V_\varphi}{V}, \qquad \frac{\dot{\chi}}{H} = -M_{\rm pl}^2 \frac{V_\chi}{V}. \tag{11.4.18}$$

Rotating those equations into the (σ, s) basis yields:

$$\frac{\dot{\sigma}}{H} = -M_{\rm pl}^2 \frac{V_\sigma}{V}, \qquad \frac{\dot{s}}{H} = -M_{\rm pl}^2 \frac{V_s}{V} = 0, \tag{11.4.19}$$

since at background level $\dot{s} = 0$ we know that $V_s = 0$ in the slow-roll regime. Consequently, the background trajectory is always in the direction of the gradient of the potential, such that there is no sideways component sourcing the gradient. We thus conclude that the total slow-roll parameter can be written as:

$$\varepsilon = \varepsilon^{\varphi\varphi} + \varepsilon^{\chi\chi} = \frac{M_{\rm pl}^2}{2}\left(\frac{V_\sigma}{V}\right)^2, \tag{11.4.20}$$

and we need not define $\varepsilon^{\sigma s}$, ε^{ss} or even a separate $\varepsilon^{\sigma\sigma}$. Instead we have the relations:

$$\varepsilon^{\varphi\varphi} = \varepsilon \cos^2\theta, \qquad \varepsilon^{\chi\chi} = \varepsilon \sin^2\theta. \tag{11.4.21}$$

However for $\tilde{\eta}_{IJ}$ we do define:

$$\tilde{\eta}_{\sigma\sigma} = M_{\rm pl}^2 \frac{V_{\sigma\sigma}}{V}, \quad \tilde{\eta}_{\sigma s} = M_{\rm pl}^2 \frac{V_{\sigma s}}{V}, \quad \tilde{\eta}_{ss=} M_{\rm pl}^2 \frac{V_{ss}}{V}, \tag{11.4.22}$$

from now on we will omit the tilde on $\tilde{\eta}_{IJ}$ and keep in mind that the η-parameter is always defined with respect to the potential in this section. Notice that the result $V_s = 0$ does in general not imply $V_{\sigma s} = 0$ since (σ, s) constitutes a curvy-linear coordinate system. The slow-roll parameter η_{IJ} in the basis (σ, s) are related to the basis (φ, χ) in the same way as the second field derivative of the potential, that is:

$$[\eta_{I'J'}] = S^{-1}[\eta_{IJ}]S, \tag{11.4.23}$$

and the explicit transformation equation becomes:

$$\begin{aligned}
\eta_{\sigma\sigma} &= \cos^2\theta\eta_{\varphi\varphi} + 2\cos\theta\sin\theta\eta_{\varphi\chi} + \sin^2\theta\eta_{\chi\chi}, \\
\eta_{ss} &= \sin^2\theta\eta_{\varphi\varphi} - 2\cos\theta\sin\theta\eta_{\varphi\chi} + \cos^2\eta_{\chi\chi}, \\
\eta_{\sigma s} &= -\sin\theta\cos\theta\eta_{\varphi\varphi} + (\cos^2\theta - \sin^2\theta)\eta_{\varphi\chi} + \cos\theta\sin\theta\eta_{\chi\chi}.
\end{aligned} \tag{11.4.24}$$

The above results show that $\eta_{\sigma\sigma} + \eta_{ss} = \eta_{\varphi\varphi} + \eta_{\chi\chi}$.

Using Eqs. 11.1.19 and the above results, we can now aim to write the second order equations for θ and σ in the slow-roll approximation. In particular, on account of Eqs. 11.1.19, the second time derivative of the adiabatic field σ can be written in terms of the first time derivative as:

$$\frac{\ddot{\sigma}}{H} = \dot{\sigma}(\varepsilon - \eta_{\sigma\sigma}). \tag{11.4.25}$$

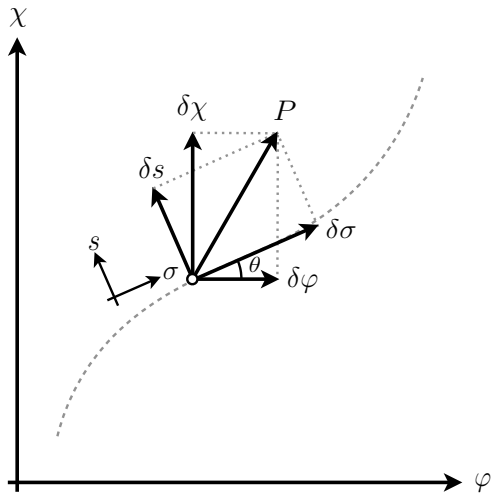


Figure 11.1: Pictorial definition of the bases (φ, χ) and (σ, s), their relation to the background trajectory (dotted grey) and the associated decomposition of the generic perturbation P. The field space angle θ is also indicated, as well as the orientation of the (σ, s) coordinate system. It is evident from this schematic that the orientation of this coordinate system indeed varies along the background trajectory and hence (σ, s) is a local coordinate system.

Similarly, by taking the time-derivative of the equation $\dot{s} = -\dot{\varphi}\sin\theta + \dot{\chi}\cos\theta = 0$, we find a slow-roll equation for θ, reading:

$$\frac{\dot{\theta}}{H} = -\eta_{\sigma s}. \tag{11.4.26}$$

It proves insightful to compare this equation for θ to the exact one (Eq. 11.4.15). In the slow-roll regime, where we neglected second order time derivatives of the fields, the background solution is completely determined by the topography of the potential. The classical background trajectories are paths of steepest descent on the potential surface. Sideway components V_s do not contribute to the gradient, as mentioned before. The curvature of potential is characterized by $\eta_{\sigma s}$, which measures how the potential tilts sideways as the field rolls in the direction of the adiabatic field σ. On the other hand, in the exact case (i.e. without the slow-roll approximation), the trajectory will not correspond to paths of steepest descent, since the centrifugal effect pushes the field away from the σ-direction (thus in the s-direction), causing the sideways component of the potential gradient, V_s, to cause a centripetal force inducing the background trajectory to curve.

11.4.4 Evolution of Perturbations

We now leave the background dynamics and will consider perturbations in the background fields φ and χ. In analogy with the background analysis, we will rotate the basis $(\delta\varphi, \delta\chi)$ to obtain a basis in which we have purely adiabatic and entropy perturbations:

$$\begin{pmatrix} \delta\varphi \\ \delta\chi \end{pmatrix} = S(\theta) \begin{pmatrix} \delta\sigma \\ \delta s \end{pmatrix}, \tag{11.4.27}$$

Here, the orientation of $\delta\sigma$ coincides with the direction of the (adiabatic) background solution at the point where the perturbation is defined and δs is orthogonal to this direction. In Fig.

11.1, the definition of the bases (φ, χ) and (σ, s) and the associated decomposition of a generic perturbation P are illustrated.

From the quadratic action S_2 (Eq. 11.3.1), we may find the equations of motion for the field perturbations in the basis (φ, χ). Rotating into the (σ, s) basis, we may rewrite those equations in the form [45]:

$$\delta\ddot{\sigma} + 3H\delta\dot{\sigma} + \left(\mathcal{M}_{\sigma\sigma} - \frac{\partial^2}{a^2} - \dot{\theta}^2\right)\delta\sigma = 2\frac{d}{dt}(\dot{\theta}\delta s) - 2\left(\frac{V_\sigma}{\dot{\sigma}} + \frac{\dot{H}}{H}\right)\dot{\phi}\delta s, \tag{11.4.28}$$

$$\delta\ddot{s} + 3H\delta\dot{s} + \left(V_{ss} + 3\dot{\theta}^2 - \frac{\partial^2}{a^2}\right) = -\frac{\dot{\theta}}{\dot{\sigma}}\frac{4}{a^2}\partial^2\Phi. \tag{11.4.29}$$

Here, we will not give a detailed derivation; we refer the reader to derivation contained in [45]. A few comments on the above evolution equations are in order. For a straight trajectory in field space ($\dot{\theta} = 0$), the equation of motion for $\delta\sigma$ will reduce to the fluctuation equation in the single field case: setting $\delta\sigma = \delta\phi$ and $\delta s = 0$, we obtain Eq. 5.2.7 by proper identification of $\mathcal{M}_{\phi\phi}$. In particular, to zeroth order in slow-roll (i.e. $\mathcal{M}_{\phi\phi} = V_{\phi\phi} + \mathcal{O}(\varepsilon)$), we obtain exactly Eq. 5.2.7 up to zeroth order in slow-roll.

Furthermore, notice that the right side of the equation of motion for the entropy perturbation contains a gradient $\partial^2\Phi$ and therefore vanishes outside the horizon. Consequently, if $\delta s = 0$ at horizon exit, no significant non-zero δs will be generated when the scales outside the horizon. This is essentially a manifestation of the statement that adiabatic perturbations remain so outside the horizon (see section 6.5 for the single-field case discussion).

Notice that the adiabatic and entropy perturbations are typically not defined in the basis which diagonalizes the mass matrix so that their evolution is coupled. In order to get to a basis where the two perturbations are independent, we perform a rotation $U(\vartheta)$ in field space of the form:

$$\gamma^I = U^{-1}(\vartheta)f^I, \qquad f^I \equiv a\begin{pmatrix}\delta\varphi \\ \delta\chi\end{pmatrix}, \qquad \gamma^I \equiv a\begin{pmatrix}\delta\tilde{\varphi} \\ \delta\tilde{\chi}\end{pmatrix}, \tag{11.4.30}$$

where we denote the angle with ϑ to avoid confusion with the angle θ associated with the transformation $S(\theta)$ defined earlier to get from (φ, χ) to the (σ, s) basis. The tilde on perturbations is used to indicate that they are defined in the basis where they evolve independently. The eigenvalue equation (Eq. 11.3.7) for two dimension becomes:

$$U^{-1}(\vartheta)\mathcal{C}U(\vartheta) = \begin{pmatrix}\lambda_\varphi & 0 \\ 0 & \lambda_\chi\end{pmatrix}. \tag{11.4.31}$$

To achieve independently evolving perturbations, the rotation matrix is chosen such that it diagonalizes the redefined mass-matrix $\mathcal{C}_{IJ}$, which is given by:

$$\mathcal{C}_{IJ} = 3\begin{pmatrix}\varepsilon + 2\varepsilon_{\varphi\varphi} - \eta_{\varphi\varphi} & 2\varepsilon_{\varphi\chi} - \eta_{\varphi\chi} \\ 2\varepsilon_{\varphi\chi} - \eta_{\varphi\chi} & \varepsilon + 2\varepsilon_{\chi\chi} - \eta_{\chi\chi}\end{pmatrix}, \tag{11.4.32}$$

for the two-field system (recall $\mathcal{C}_{IJ} \equiv 3\varepsilon\delta_{IJ} - \mathcal{W}_{IJ}$). From the requirement that $U^{-1}\mathcal{C}_{IJ}U$ is diagonal, we find the constraint:

$$(2\varepsilon_{\chi\chi} - 2_{\varphi\varphi} + \eta_{\varphi\varphi} - \eta_{\chi\chi})\sin\vartheta\cos\vartheta + 2(\varepsilon_{\varphi\chi} - \eta_{\varphi\chi})(\cos^2\vartheta - \sin^2\vartheta) = 0. \tag{11.4.33}$$

This constraint equation allows us to solve for the angle ϑ quantifying the rotation in field space to get from the basis $(\delta\varphi, \delta\chi)$ to the basis $(\delta\tilde{\varphi}, \delta\tilde{\chi})$ in which the two perturbations are

independent. The result reads:

$$\tan 2\vartheta = \frac{4\varepsilon_{\varphi\chi} - 2\eta_{\varphi\chi}}{2(\varepsilon_{\varphi\varphi} - \varepsilon_{\chi\chi}) - (\eta_{\varphi\varphi} - \eta_{\chi\chi})}. \tag{11.4.34}$$

In addition, we can solve for expression of the two eigenvalues from the eigenvalue equation (Eq. 11.4.31), i.e. $\det(\mathcal{C} - \lambda \mathbb{1}_{2\times 2}) = 0$. This yields the solutions:

$$\lambda_\pm = \frac{3}{2}\Big(4\varepsilon - (\eta_\varphi + \eta_{\chi\chi}) \pm \sqrt{(2\Delta\varepsilon - \Delta\eta)^2 + 4(2\varepsilon_{\varepsilon\chi} - \eta_{\varphi\chi})^2}\Big), \tag{11.4.35}$$

where $\Delta\varepsilon \equiv \varepsilon_{\varphi\varphi} - \varepsilon_{\chi\chi}$ and $\Delta\eta \equiv \eta_{\varphi\varphi} - \eta_{\chi\chi}$. Notice that the above expressions do not specify which one is λ_φ and which is λ_χ. Similarly, the solution for the angle ϑ only specifies its value up to a factor of $\pi/2$. Hence, the system of equations allows for the residual freedom to add $\pi/2$ to the angle ϑ, which interchanges $\lambda_{1,2}$.

In terms of the slow-roll parameters associated with the adiabatic and entropy fields, the results are:

$$\begin{aligned} \lambda_\pm &= \frac{3}{2}\Big(4\varepsilon - (\eta_{\sigma\sigma} + \eta_{ss}) \pm \sqrt{\omega^2 + 4\eta_{\sigma s}^2}\Big), \\ \tan 2\vartheta &= \frac{\omega \sin 2\theta - 2\eta_{\sigma s}\cos 2\theta}{\omega \cos 2\theta + 2\eta_{\sigma s}\sin 2\theta}, \end{aligned} \tag{11.4.36}$$

in which we defined $\omega \equiv 2\varepsilon - (\eta_{\sigma\sigma} - \eta_{ss})$. From the above expression, we find that $\vartheta = \theta$ solely in the special case of $\eta_{\sigma s} = 0$, corresponding to $V_{\sigma s} = 0$,[2] and hence the adiabatic and entropy perturbations are already independent. Otherwise, the perturbations in the (σ, s) are still coupled.

In conclusion, we will thus have to find a connection between the perturbations in the (σ, s) basis and those in the bases $(\tilde{\varphi}, \tilde{\chi})$, where the perturbations will evolve independent of each other. Using the inverse of Eq. 11.4.27 and Eq. 11.4.30, we find a relation between the bases (σ, s) and $(\tilde{\varphi}, \tilde{\chi})$, reading:

$$\begin{pmatrix} \delta\sigma \\ \delta s \end{pmatrix} = S^{-1}(\theta)U(\vartheta)\begin{pmatrix} \delta\tilde{\varphi} \\ \delta\tilde{\chi} \end{pmatrix}, \qquad S^{-1}(\theta)U(\vartheta) = \begin{pmatrix} \cos(\vartheta - \theta) & -\sin(\vartheta - \theta) \\ \sin(\vartheta - \theta) & \cos(\vartheta - \theta) \end{pmatrix}. \tag{11.4.37}$$

In the basis $(\tilde{\varphi}, \tilde{\chi})$, where the perturbations evolve independently, we can use the mode function (Eq. 11.3.10) to conveniently compute the primordial power spectra, which will be done explicitly in the section below.

11.4.5 Primordial Power Spectra at Horizon Crossing

Using the above results, we can compute the power spectra $\mathcal{P}_\sigma$, $\mathcal{P}_s$ and $\mathcal{P}_{\sigma s}$ at horizon exit by using the mode function for the independently evolving fields in the $(\tilde{\varphi}, \tilde{\chi})$ basis (conform Eq. 11.3.10). The first step in this computation is to relate the fluctuation amplitude of the adiabatic pertubation $|\delta\sigma|^2$, the entropy perturbation $|\delta s|^2$ and the cross correlation $\delta\sigma\delta s^*$, to the amplitudes of the independent fluctuations $\delta\tilde{\varphi}$ and $\delta\tilde{\chi}$. This can be done conveniently

[2] Notice that this would be case for a potential which is sum-separable in terms of σ and s, i.e. $V(\sigma, s) = U(\sigma) + W(s)$. We will consider a two-field model with sum-separable potential in section 12.8.

by means of the rotation matrix defined in Eq. 11.4.37, yielding:

$$\begin{aligned} a^2|\delta\sigma|^2 &= \cos^2(\vartheta-\theta)|\gamma_\varphi|^2 + \sin^2(\vartheta-\theta)|\gamma_\chi|^2, \\ a^2\delta\sigma\delta s^* &= \frac{1}{2}\sin 2(\vartheta-\theta)\big(|\gamma_\varphi|^2 - |\gamma_\chi|^2\big), \\ a^2|\delta s|^2 &= \frac{1}{2}\big(|\gamma_\varphi|^2 + |\gamma_\chi|^2\big) - \cos 2(\vartheta-\theta)\big(|\gamma_\varphi|^2 - |\gamma_\chi|^2\big), \end{aligned} \tag{11.4.38}$$

where $\gamma_{\varphi,\chi}$ denote the mode function for φ and χ, respectively. Those mode functions differ in the parameter $\nu_{\varphi,\chi}$, which is determined by the two different $\lambda_{\varphi,\chi}$.

However, recall that our final goal is to compute the spectral indices for the power spectra of $\mathcal{R}$, S and their cross-correlation to first order in slow-roll parameters, which can be calculated using the power spectra to zeroth order in slow-roll. Hence, for computing the spectral indices to first order, we can use the mode functions to zeroth order, i.e. we can set $\nu_{\varphi,\chi} \to 3/2$ and neglect the contribution of the eigenvalues, which enters at first order in slow-roll.[3]

In this limit, the two mode functions $\gamma_{\varphi,\chi}$ are equivalent, so that we can drop the field labels and the above equations for the amplitudes simplify significantly:

$$a^2|\delta\sigma|^2 = |\gamma|^2, \qquad a^2|\delta s|^2 = |\gamma|^2, \qquad \delta\sigma\delta s^* = 0. \tag{11.4.39}$$

In addition we can use the fact that in the limit of horizon crossing $k \to 0$, the mode function γ can be written as (conform Eq. 5.4.30):

$$\lim_{k\to 0}|\gamma| = \frac{C(\nu)}{\sqrt{2k}}\left(\frac{k}{aH}\right)^{1/2-\nu}, \qquad C(\nu) \equiv 2^{\nu-3/2}\frac{\Gamma(\nu)}{\Gamma(3/2)}, \tag{11.4.40}$$

with $C(\nu)$ tending to unity in the limit where $\nu \to 3/2$ (we also dropped the field label on ν). At horizon crossing, the power spectra then simply become:

$$\mathcal{P}^*_\sigma = \frac{k^3}{2\pi^2}\left|\frac{\gamma}{a}\right|^2 = \left(\frac{H}{2\pi}\right)^2 = \mathcal{P}^*_s, \qquad \mathcal{C}^*_{\sigma s} \equiv \frac{k^3}{2\pi^2}\delta\sigma\delta s = 0, \tag{11.4.41}$$

where we have used the superscript $*$ to denote the fact that those relations only hold at horizon exit and dropped the conjugate symbol on δs since it is manifestly real. Due to the fact that have multiple fields, which evolve on super-horizon scales, the power spectra will as well. In the next section we will discuss how to include this time-evolution, in order to compute the power spectra at some time $t > t_*$ during the radiation era. In this way, they can be related directly to observations of the CMB anisotropies.

11.4.6 Evolution of Spectra Outside the Horizon

As we have explained above, to compute the power spectra and associated spectral indices in the radiation era, i.e. at times that are accessible by observations, we will have to take the evolution of the comoving curvature perturbation outside the horizon into account.[4] For the two-field system, the comoving curvature perturbation is given by:

$$\mathcal{R} = H\frac{\dot\varphi\delta\varphi + \dot\chi\delta\chi}{\dot\varphi^2 + \dot\chi^2} = H\frac{\delta\sigma}{\dot\sigma}, \tag{11.4.42}$$

[3]Notice that this approach conveniently avoids the subtlety that Eq. 11.4.36 for $\lambda_\pm$ does not specify which of the two corresponds to λ_φ and which to λ_χ.

[4]We consider only modes which are still super-horizon at the moment of interest in the radiation era, so that the evolution equation simplify considerably.

from which we find that $\mathcal{R}$ is only sourced by perturbations which are parallel to the background trajectory (i.e. the adiabatic perturbation). The time-evolution of $\mathcal{R}$ in this model is given by [42, 43, 45]:

$$\dot{\mathcal{R}} = -\frac{H}{\dot{H}}\frac{1}{a^2}\partial^2\Psi + \frac{H}{2}\left(\frac{\delta\varphi}{\dot{\varphi}} - \frac{\delta\chi}{\dot{\chi}}\right) \times \frac{d}{dt}\left(\frac{\dot{\phi}^2 - \dot{\chi}^2}{\dot{\phi}^2 + \dot{\chi}^2}\right), \tag{11.4.43}$$

which in terms of the adiabatic and entropic field perturbations takes the simple form [45]:

$$\dot{\mathcal{R}} = -\frac{H}{\dot{H}}\frac{1}{a^2}\partial^2\Psi - 2H\frac{\dot{\theta}}{\dot{\sigma}}\delta s. \tag{11.4.44}$$

Outside the horizon, the gradient of the gravitational potential Ψ can be neglected so that the evolution equation for $\mathcal{R}$ simplifies to:

$$\dot{\mathcal{R}} = -2\dot{\theta}S, \qquad S \equiv H\frac{\delta s}{\dot{\sigma}}. \tag{11.4.45}$$

In terms of those variables, we can connect the power spectra for σ and s to those of $\mathcal{R}$ and S as follows:

$$\mathcal{P}^*_{\mathcal{R}} = \left(\frac{H}{\dot{\sigma}}\right)^2 \mathcal{P}^*_\sigma, \qquad \mathcal{P}^*_S = \left(\frac{H}{\dot{\sigma}}\right)\mathcal{P}^*_s, \qquad \mathcal{C}^*_{\mathcal{R}S} = 0, \tag{11.4.46}$$

at horizon exit.

Now the task is to evolve those power spectra to a time $t > t_*$ during the radiation era. In order to evolve momentum modes to the radiation era, we introduce a transfer function, denoted as $T_k(t_*, t)$, which evolves $\mathcal{R}_k$ and S_k from horizon exit t_* to some later time t in the radiation era, when the cosmological modes of interest are still outside the horizon. On those scales, the momentum dependence of the transfer vanishes since $k \to 0$ and we drop the momentum label:

$$T(t_*, t) \equiv \lim_{k\to 0} T_k(t_*, t). \tag{11.4.47}$$

The transfer function is defined in relation to $\mathcal{R}_k$ and S_k as:

$$\begin{pmatrix}\mathcal{R}_k(t)\\ S_k(t)\end{pmatrix} = T(t_*, t)\begin{pmatrix}\mathcal{R}_k(t_*)\\ S_k(t_*)\end{pmatrix}, \qquad T(t_*, t) = \begin{pmatrix}T_{\mathcal{R}\mathcal{R}} & T_{\mathcal{R}S}\\ T_{S\mathcal{R}} & T_{SS}\end{pmatrix}. \tag{11.4.48}$$

The transfer function can be simplified by noting that outside the horizon adiabatic perturbations remain adiabatic and constant, yielding the constraint $T_{S\mathcal{R}} = 0$ and $T_{\mathcal{R}\mathcal{R}} = 1$, respectively. In terms of the transfer function, we then get the evolution equations for $\mathcal{R}_k$ and S_k at time $t > t_*$, reading:

$$\mathcal{R}_k(t) = \mathcal{R}_k(t_*) + T_{\mathcal{R}S}(t_*, t)S_k(t_*), \qquad S_k(t) = T_{SS}(t_*, t)S_k(t_*). \tag{11.4.49}$$

In terms of those transfer equations, we can express the power spectra for $\mathcal{R}$, S and their cross correlation at time $t > t_*$ to the those at horizon exit (Eqs. 11.4.46):

$$\begin{aligned}\mathcal{P}_{\mathcal{R}}(t) &= \mathcal{P}_{\mathcal{R}}(t_*) + T^2_{\mathcal{R}S}(t_*, t)\mathcal{P}_S(t_*),\\ \mathcal{P}_S(t) &= T^2_{SS}(t_*, t)\mathcal{P}_S(t_*)\\ \mathcal{C}_{\mathcal{R}S} &= T_{SS}(t, t_*)\mathcal{C}_{\mathcal{R}S}(t_*) + T_{\mathcal{R}S}T_{SS}(t_*, t)\mathcal{P}_S(t_*).\end{aligned} \tag{11.4.50}$$

Notice that even if the cross correlation between $\mathcal{R}$ and S is zero at horizon crossing (i.e. $\mathcal{C}^*_{\mathcal{R}S} = 0$), cross correlation between the two will be generated during the evolution outside the horizon.

The next step is to derive an explicit form for the remaining transfer functions. On account of Eq. 11.4.45, we may write outside the horizon that :

$$\dot{\mathcal{R}} = H\alpha(t)S, \qquad \dot{S} = H\beta(t)S, \tag{11.4.51}$$

where the two introduced functions $\alpha(t)$ and $\beta(t)$ are related to the transfer functions $T_{\mathcal{R}S}$ and T_{SS}. To find the explicit relation, we note that $T_{SS}(t_*, t) = S(t)/S(t_*)$ and hence we can integrate the $\dot{S}$ equation to give:

$$\begin{aligned} \frac{dS(t')}{S(t')} &= \beta(t')H(t')dt' \\ \ln S(t) - \ln S(t_*) &= \int_{t_*}^{t} dt'\ \beta(t')H(t') \\ T_{SS}(t_*, t) &= \exp\left[\int_{t_*}^{t} dt'\ \beta(t')H(t')\right]. \end{aligned} \tag{11.4.52}$$

The expression for $T_{\mathcal{R}S}$ is now given in terms of the above formula for T_{SS} as:

$$T_{\mathcal{R}S}(t, t_*) = \int_{t_*}^{t} dt'\ \alpha(t')H(t')T_{SS}(t_*, t'). \tag{11.4.53}$$

For notational simplicity, we will drop the Fourier labels from now on, keeping in mind that $\mathcal{R} = \mathcal{R}_k$, $S = S_k$ and $t_* = t_*(k)$.

11.4.7 Computation of Spectral Indices

The functions $\alpha(t)$ and $\beta(t)$ depend on the details of the inflationary model and the subsequent reheating stage. However, in order to find the spectral indices to first order, it turns out we fortunately only require the time-derivatives of $T_{\mathcal{R}S}$ and T_{SS} at horizon exit:

$$\begin{aligned} \partial_{t_*} T_{SS} &= -\beta(t_*)H(t_*)T_{SS}(t, t_*), \\ \partial_{t_*} T_{\mathcal{R}S} &= -\alpha(t_*)H(t_*) - \beta(t_*)H(t_*)T_{\mathcal{R}S}(t, t_*). \end{aligned} \tag{11.4.54}$$

Those expressions involve $\alpha(t_*)$ and $\beta(t_*)$ evaluated at horizon exit. Assuming the slow-roll approximation still holds at the moment the modes of interest leave the horizon, we can use the slow-roll equations from section 11.4.3 to compute $\alpha(t_*)$ and $\beta(t_*)$.

In particular, combining Eqs. 11.4.26 and Eq. 11.4.51, we find the expression for $\alpha(t_*)$:

$$\alpha(t_*) = 2\eta_{\sigma s}. \tag{11.4.55}$$

To obtain the from for $\beta(t_*)$, we use the evolution equation for the entropy perturbation δs in the super-horizon limit:

$$\delta\ddot{s} + 3H\delta\dot{s} + (V_{ss} + 3\dot{\theta}^2)\delta s = 0. \tag{11.4.56}$$

Dropping the second order time derivative term and writing the $\dot{\theta}$ expression to first order in slow-roll, we obtain the equation:

$$\delta\dot{s} + H\eta_{ss}\delta s = 0. \tag{11.4.57}$$

In combination with the slow-roll equation for σ (Eq.

Now that we have obtained the explicit expressions for $\alpha(t_*)$ and $\beta(t_*)$, we can calculate the primordial spectral indices. We show the explicit calculation for the power spectrum of $\mathcal{R}$. Using that at zeroth order in slow-roll parameters the power spectra for $\mathcal{R}$ and S are the same at horizon exit, we find:

$$\mathcal{P}_{\mathcal{R}}(t) = \mathcal{P}_{\mathcal{R}}(t_*) + T^2_{\mathcal{R}S}(t_*, t)\mathcal{P}_S(t_*) = \frac{H^2}{\dot{\sigma}^2}\left(\frac{H}{2\pi}\right)^2 (1 + T^2_{\mathcal{R}S}(t_*, t)), \tag{11.4.58}$$

where all quantities in the prefactor are to be evaluated at horizon crossing. To compute the spectral index, defined as:

$$n_{\mathcal{R}} - 1 \equiv \frac{d\ln \mathcal{P}_{\mathcal{R}}}{d\ln k}, \tag{11.4.59}$$

we require the scaling of $H^4/\dot{\sigma}^2$, which can be obtained by using Eq. 11.4.19 and the slow-roll Friedmann equation $H^2 = V/3M^2_{\rm pl}$, we find:

$$\frac{H^4}{\dot{\sigma}^2} \propto \frac{V}{\varepsilon}. \tag{11.4.60}$$

On account of this scaling relation, the spectral index now becomes:

$$n_{\mathcal{R}} - 1 = \frac{d\ln \mathcal{P}_{\mathcal{R}}}{d\ln k} = \frac{d\ln V}{d\ln k} - \frac{d\ln \varepsilon}{d\ln k} + \frac{d\ln(1 + T^2_{\mathcal{R}S})}{d\ln k}. \tag{11.4.61}$$

The differential $d\ln k$ can be converted into a time differential at horizon crossing by using th the relation $k = aH$ at horizon crossing:

$$\frac{d\ln k}{dt_*} = \frac{d\ln aH}{dt} = (1 - \varepsilon)H, \tag{11.4.62}$$

where ε and H are to be evaluated at the time of horizon crossing t_*. On account of this relation, the first term contributing to the spectral index reads:

$$\frac{d\ln V}{d\ln k} = \frac{1}{(1-\varepsilon)H}\frac{V_\sigma \dot{\sigma}}{V} = -2\varepsilon + \mathcal{O}(\varepsilon^2), \tag{11.4.63}$$

where we have used that the time derivative of the potential can be written as $\dot{V} = V_\sigma \dot{\sigma}$ according to Eq. 11.4.9. Similarly, the second contribution becomes:

$$\frac{d\ln \varepsilon}{d\ln k} = \frac{1}{(1-\varepsilon)H}\frac{\dot{\varepsilon}}{\varepsilon} = 4\varepsilon - 2\eta_{\sigma\sigma}, \tag{11.4.64}$$

where we used the expression for $\dot{\varepsilon}/H = 4\varepsilon^2 - 2\varepsilon\eta_{\sigma\sigma}$ given in the last equation of Eqs. 11.1.19 (for two fields s and σ). Those two contributions add up to the single field result for the spectral index, in particular, compare to Eq. 5.4.33 by identifying $n_{\mathcal{R}} = n_{\rm s}$ and $\sigma = \phi$.

The difference between the single-field and the two-field result lies in the contribution coming from the third term in Eq. 11.4.61. This term can be evaluated as:

$$\begin{aligned}
\frac{d\ln(1 + T^2_{\mathcal{R}S})}{d\ln k} &= \frac{1}{1 + T^2_{\mathcal{R}}S}\frac{1}{H}\frac{d}{dt_*}(1 + T^2_{\mathcal{R}S}) \\
&= \frac{2T_{\mathcal{R}S}}{1 + T^2_{\mathcal{R}S}}(-\alpha(t_*) - \beta(t_*)T_{\mathcal{R}S}) \\
&= \frac{T^2_{\mathcal{R}S}}{1 + T^2_{\mathcal{R}S}}(4\varepsilon - 2\eta_{\sigma\sigma} + 2\eta_{ss}) - \frac{T_{\mathcal{R}S}}{1 + T^2_{\mathcal{R}S}}4\eta_{\sigma s}.
\end{aligned} \tag{11.4.65}$$

In general, we doe note have constraints on the transfer function $T_{\mathcal{R}S}$, as it strongly depends on the inflationary details as well as the subsequent era of reheating. However, the particular combinations in the above expressions have limit range:

$$0 \leq \frac{T_{\mathcal{R}S}^2}{1+T_{\mathcal{R}S}^2} \leq 1, \tag{11.4.66}$$

and its square root lies between -1 and 1. Therefore, we can conveniently define the angle Δ as:

$$\cos\Delta \equiv \frac{T_{\mathcal{R}S}}{\sqrt{1+T_{\mathcal{R}S}^2}}, \qquad \sin\Delta \equiv \frac{1}{\sqrt{1+T_{\mathcal{R}S}^2}}. \tag{11.4.67}$$

In terms of this angle, the last contribution to the spectral index reads:

$$\frac{d\ln(1+T_{\mathcal{R}S}^2)}{d\ln k} = -4\eta_{\sigma s}\cos\Delta\sin\Delta + (4\varepsilon - 2\eta_{\sigma\sigma} + 2\eta_{ss})\cos^2\Delta. \tag{11.4.68}$$

Finally, combining the three contributions above, we find that the spectral index for the power spectrum of the adiabatic mode $\mathcal{R}$ is time-dependent via the transfer function $T_{\mathcal{R}S}(t_*, t)$ and reads:

$$n_{\mathcal{R}} = -(6-4\cos^2\Delta)\varepsilon + 2\eta_{\sigma\sigma}\sin^2\Delta - 4\eta_{\sigma s}\cos\Delta\sin\Delta + 2\eta_{ss}\cos^2\Delta. \tag{11.4.69}$$

In a similar manner, the spectral indices for the spectrum of S and the cross correlation spectrum $\mathcal{C}_{\mathcal{R}S}$ between $\mathcal{R}$ and S can be found as well, yielding respectively:

$$\begin{aligned} n_S - 1 &= -2\varepsilon + 2\eta_{ss} \\ n_{\mathcal{R}S} - 1 &= -2\varepsilon - 2\eta_{\sigma s}\tan\Delta + 2\eta_{ss}. \end{aligned} \tag{11.4.70}$$

The above results provide the primary aim of this section: computing the spectral indices of the power spectra for two-field inflation to first order in slow-roll.

Chapter 12

Bispectrum for Multi-Field Inflation

"The recent developments in cosmology strongly suggest that the universe may be the ultimate free lunch."

— Alan Guth

In this chapter, we will compute the bispectrum for multi-field inflation with n canonical scalar fields in the slow-roll approximation. To derive the bispectrum, we closely follow the work by Seery and Lidsey [76]. As we will show, the bispectrum for the multi-field scenario differs in two important aspects from the single-field case. First of all, in multi-field models the comoving curvature perturbation is not conserved on super-horizon scales. This property of the multi-field dynamics complicates the calculation significantly. To conveniently compute the bispectrum, we will work with the so-called δN formalism, which provides an intuitive and simple prescription to compute three-point correlation functions. Secondly, we will find that, contrary to the single field scenario, the non-linearity parameter can become of order unity in the squeezed limit of the local template and can hence be detectable. This difference provides a distinct observational discrimination between the single-field and multi-field scenario. We will examine the conditions required for a large non-linearity parameter in the case of two fields.

This chapter is organized as follows. First of all, we will introduce the δN formalism used to relate the three-point correlation function of $\mathcal{R}$ in a convenient way to the three-point correlator of quantum fluctuations in the fields $\delta\phi^I$. In section 12.2, we will introduce the path integral formalism that we will use to compute the three-point correlation function of the quantum fluctuations $\delta\phi^I$. Subsequently, we will compute the multi-field action to third order in perturbations in a very similar way to the single-field case using the ADM formalism. In section 12.6, we will compute the leading order bispectrum for multi-field inflation and show that our result is equivalent to the result of Seery and Lidsey [76]. We consider the squeezed limit of multi-field inflation in section 12.7 and argue that the non-linearity parameter can in principle become large, in sharp contrast to the single-field scenario. Finally, in section 12.8, we examine the conditions required to get a significant level non-gaussianity in inflaton with two fields.

12.1 The δN Formalism

In the multi-field scenario, it is not trivial to find a direct relation between $\mathcal{R}$, the gravitational potential and the field fluctuations $\delta\phi^I$. Therefore, we take a different approach to relate $\mathcal{R}$ to the fluctuations in the fields, known as the δN formalism. This formalism states that, outside

the horizon, the curvature perturbation can be written as the variation of the number of e-foldings of expansion:

$$\mathcal{R}(\boldsymbol{x}, t) = \delta N(\boldsymbol{x}, t). \tag{12.1.1}$$

Below, we will derive this equation in detail, based on the so-called separate universe approach.

12.1.1 Separate Universe Approach

To derive the above equation, we assume that we can describe the actual universe on large scales by introducing a smoothing scale L, which is larger than the Hubble scale H^{-1}, so that:

$$LH \gg 1. \tag{12.1.2}$$

The existence of such a scale is required, as otherwise the notion of a background FRW universe on which the perturbations are defined, would not make sense. In term of this length scale, spatial gradients can be associated with the small parameter ξ, which we define as:

$$\xi \equiv \frac{1}{HL}. \tag{12.1.3}$$

In terms of comoving modes in Fourier space, the scale L corresponds to a/k, where k^{-1} is the inverse comoving momentum. Comparison to the comoving Hubble size $1/aH$, we find that we can express ξ as:

$$\frac{k}{aH} \ll 1, \tag{12.1.4}$$

which is exactly the condition for modes to be on super-horizon scales. During inflation, the comoving Hubble sphere will decrease exponentially so that the smallness of ξ is guaranteed soon after the mode exits the horizon $k = aH$.

In terms of ξ, we assume that in the limit $\xi \to 0$, corresponding to a sufficiently large smoothing scale, the universe becomes locally homogeneous and isotropic. By locally homogeneous and isotropic, we mean that we can model the universe as a collection of patches with size L_P much larger than the Hubble scale but smaller than the smoothing scale, hence the hierarchy in length scales is:

$$L \gg L_P \gg H^{-1} \tag{12.1.5}$$

The patches are themselves isotropic and homogeneous and can, therefore, be modeled as separate FRW universes, giving rise to the so-called separate universe approach. Notice that this approach becomes applicable when $\xi \ll 1$, so that modes are outside the horizon.

The local properties of the different patches are the same. However, in comparing the different patches, there will be small differences in the evolution induced by quantum fluctuations in the fields. This is most easily explained in the single-field scenario, but would also apply to the multi-field case. Consider two patches of size L_P, located around coordinates $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, which are separated by a distance:

$$L_S \equiv |\boldsymbol{x}_2 - \boldsymbol{x}_1| \gg L_P. \tag{12.1.6}$$

At some fixed time t_c, let the fluctuation of the inflaton in patch 1 and patch 2 be positive and negative, respectively. That is:

$$\delta\phi_1 \equiv \delta\phi(\boldsymbol{x}_1, t_c) > 0, \qquad \delta\phi_2 \equiv \delta\phi(\boldsymbol{x}_2, t_c) < 0. \tag{12.1.7}$$

This means that in patch 1, the inflaton will be slightly lower on the potential at time t_c compared to patch 2, which is slightly higher on the potential. Consequently, patch 2 will inflate slightly longer than patch 1, until inflaton ends when $\varepsilon \to 1$. In other the number of e-folds N in patch 2 will be larger than in patch 1. This difference in e-foldings can be related directly to the comoving curvature perturbation.

12.1.2 Relating $\mathcal{R}$ and δN

In order to find a direct relation between $\mathcal{R}$ and δN, we consider a local patch located at coordinate $\boldsymbol{x}$ with local (i.e. spatial dependent) scale factor:

$$\tilde{a}(t, \boldsymbol{x}) \equiv a(t) e^{\mathcal{R}(\boldsymbol{x}, t)}, \tag{12.1.8}$$

where the tilde denotes that the quantity is local and $a(t)$ is the background FRW scale factor. In the literature, ζ is often considered instead of $\mathcal{R}$. However, since we are interested in super-horizon dynamics, we can set $\zeta = \mathcal{R}$. The local Hubble parameter can be written as follows:

$$\tilde{H} = \frac{\dot{\tilde{a}}}{\tilde{a}} = H(t) + \dot{\mathcal{R}}(t, \boldsymbol{x}). \tag{12.1.9}$$

Using these definitions, we can compute the number of e-foldings of inflationary expansion between t_1 and t_2 as follows:

$$\begin{aligned} \tilde{N}(t_2, t_1, \boldsymbol{x}) = \int_{t_1}^{t_2} dt \tilde{H} &= \int_{t_1}^{t_2} dt \, \left(H(t) + \dot{\mathcal{R}}(t, \boldsymbol{x})\right) \\ &= \ln \left[\frac{a(t_2)}{a(t_1)}\right] + \mathcal{R}(t_2, \boldsymbol{x}) - \mathcal{R}(t_1, \boldsymbol{x}). \end{aligned} \tag{12.1.10}$$

To proceed, we take the initial slice at t_1 to be defined in the flat gauge, so that $\mathcal{R}(t_1, \boldsymbol{x}) \equiv 0$. At some later time t_1, we consider the slice to be in the uniform density gauge ($\mathcal{R}(t_2, \boldsymbol{x}) \neq 0$). Then we find from the above equation that the comoving curvature perturbation at t_2 can be written as:

$$\mathcal{R}(t_2) = \tilde{N}(t_2, t_1, \boldsymbol{x}) - N(t_2, t_1) \equiv \delta N, \tag{12.1.11}$$

which is the advocated result. Notice that the left-hand-side is independent of t_1 and hence we can take t_1 arbitrarily. This freedom reflects on the fact that the number of e-foldings between any two flat slices is homogeneous ($\delta N = 0$) and therefore equals the background expression.

By noting that the number of e-foldings N is a direct function of the fields, we can expand the above differential relation to second order in field fluctuations, which we denote as $Q^I \equiv \delta\phi^I$, as:

$$\mathcal{R} = \partial_I N \, Q^I + \frac{1}{2} \partial_I \partial_J N \, Q^I Q^J + \mathcal{O}(Q^3), \tag{12.1.12}$$

which is provides a direct relation between the comoving curvature perturbation and the field fluctuations. Below, we will use the above result to conveniently relate correlation functions of fluctuations in the fields to the correlators of the comoving curvature perturbation.

12.1.3 Two-Point Function

First, we focus on the two-point correlation function. In momentum space, the two-point function of Q^I, denoted as $\langle Q^I_{\boldsymbol{k}_1} Q^J_{\boldsymbol{k}_2}\rangle$, reads:

$$\langle Q^I_{\boldsymbol{k}_1} Q^J_{\boldsymbol{k}_2}\rangle = (2\pi)^3 \delta^{(3)}(\boldsymbol{k}_1 + \boldsymbol{k}_2)\, \frac{2\pi^2}{k_1^3} \mathcal{P}_Q \delta^{IJ}. \tag{12.1.13}$$

We will show that the power spectrum $\mathcal{P}_Q$ takes the form of the spectrum for a massless field in de Sitter space, that is:

$$\mathcal{P}_Q = \left(\frac{H}{2\pi}\right)^2. \tag{12.1.14}$$

Using the linear term in Eq. 12.1.12, we can relate the two-point functions of Q and $\mathcal{R}$ as follows:

$$\langle \mathcal{R}_{\boldsymbol{k}_1} \mathcal{R}_{\boldsymbol{k}_2}\rangle = \partial_I N \partial_J N \langle Q^I_{\boldsymbol{k}_1} Q^J_{\boldsymbol{k}_2}\rangle, \qquad \mathcal{P}_{\mathcal{R}} = \delta^{IJ} \partial_I N \partial_J N \mathcal{P}_Q. \tag{12.1.15}$$

12.1.4 Three-Point Function

In an analogous manner, we can derive the relation between the bispectra of $\mathcal{R}$ and Q^I. We parametrize the bispectrum of Q^I following the convention of Maladacena [64, 76]:

$$\langle Q^I_{\boldsymbol{k}_1} Q^J_{\boldsymbol{k}_2} Q^K_{\boldsymbol{k}_3}\rangle = (2\pi)^3 \delta^{(3)}(\boldsymbol{K})\, \frac{4\pi^4}{k_{123}^3} \mathcal{P}_Q^2\, \mathcal{A}^{IJK}, \tag{12.1.16}$$

where $\mathcal{A}^{IJK}$ encompasses the momentum dependence of the bispectrum, similar to $\mathcal{A}$ in Eq. 10.7.21 for the single-field case. By using the first two terms in Eq. 12.1.12, we can write the three-point function of $\mathcal{R}$ in terms Q^I as follows:

$$\begin{aligned}\langle \mathcal{R}_{\boldsymbol{k}_1} \mathcal{R}_{\boldsymbol{k}_2} \mathcal{R}_{\boldsymbol{k}_3}\rangle &= \partial_I N \partial_J N \partial_K N \langle Q^I_{\boldsymbol{k}_1} Q^J_{\boldsymbol{k}_2} Q^K_{\boldsymbol{k}_3}\rangle \\ &\quad + \frac{1}{2} \partial_I N \partial_K N \partial_M \partial_N N \langle Q^I_{\boldsymbol{k}_1} Q^K_{\boldsymbol{k}_2} [Q^M \star Q^N]_{\boldsymbol{k}_3}\rangle + \text{perms.} + \cdots,\end{aligned} \tag{12.1.17}$$

where we have truncated the series at the four-point function and $\cdots$ denote the cross terms from the expansion of $\mathcal{R}$ according in Eq. 12.1.12, which we do not display for brevity. The $\star$ notation denotes a convolution product:

$$[Q^M \star Q^N]_{\boldsymbol{k}_3} \equiv \int \frac{d^3\boldsymbol{k}'}{(2\pi)^3}\, Q^M(\boldsymbol{k}_3 - \boldsymbol{k}') Q^N(\boldsymbol{k}'). \tag{12.1.18}$$

Let us now discuss the different terms in Eq. 12.1.17. The first term is just the connected bispectrum since we assume tadpole contributions to vanish (see end next section). This is equivalent to making the assumption that the vacuum is perturbatively stable. Under this assumption, we can rewrite the second (four-point) contribution as products of two-point functions:

$$\begin{aligned}\langle Q^I_{\boldsymbol{k}_1} Q^J_{\boldsymbol{k}_2} [Q^M \star Q^N]_{\boldsymbol{k}_3}\rangle &= \langle Q^I_{\boldsymbol{k}_1} Q^M_{\boldsymbol{k}_3}\rangle \langle Q^J_{\boldsymbol{k}_2} Q^N_{\boldsymbol{k}_3}\rangle + \text{perms.} \\ &= \delta^{IM}\delta^{JN} \times (2\pi)^3 \delta^{(3)}(\boldsymbol{K})\, \frac{4\pi^4}{k_{123}^3} \mathcal{P}_Q^2 \sum_i k_i^3.\end{aligned} \tag{12.1.19}$$

In the above calculation, we have implicitly assumed that the connected part of the four-point function is negligible, thereby validating the application of Wick's theorem to rewrite the four-point correlator in terms of two-point functions.

Comparing Eq. 12.1.17 to the three-point correlation function of the comoving curvature perturbation:

$$\langle \mathcal{R}_{\boldsymbol{k}_1} \mathcal{R}_{\boldsymbol{k}_2} \mathcal{R}_{\boldsymbol{k}_3} \rangle = (2\pi)^3 \delta^{(3)}(\boldsymbol{K}) \frac{4\pi^4}{k_{123}^3} \mathcal{P}_{\mathcal{R}}^2 \mathcal{A}_{\mathcal{R}}, \tag{12.1.20}$$

we find that we can relate $\mathcal{A}_{\mathcal{R}}$ to $\mathcal{A}^{IJK}$ as follows:

$$\begin{aligned} \mathcal{A}_{\mathcal{R}} &\equiv \mathcal{A}_{\mathcal{R}}^{(1)} + \mathcal{A}_{\mathcal{R}}^{(2)}, \\ \mathcal{A}_{\mathcal{R}}^{(1)} &= \frac{\partial_I N \partial_J N \partial_K N \, \mathcal{A}^{IJK}}{(\partial^L N \partial_L N)^2}, \\ \mathcal{A}_{\mathcal{R}}^{(2)} &= \frac{\partial_I N \partial_J N \partial^I \partial^J N \sum_i k_i^3}{(\partial^L N \partial_L N)^2}, \end{aligned} \tag{12.1.21}$$

where $\mathcal{A}^{(1)}$ and $\mathcal{A}^{(2)}$ are the contributions to the momentum function due to the connected three point function and the four point function defined on the first and second line in Eq. 12.1.17, respectively. Notice furthermore that we have written the power spectrum of $\mathcal{R}$ instead of Q, introducing the terms involving $\partial^I N \partial_I N$.

Finally, we introduce the momentum dependent non-linearity parameter $f_{\rm NL}$ in terms of $A_{\mathcal{R}}$. Since we know how to relate $A_{\mathcal{R}}$ to the momentum dependence $\mathcal{A}^{IJK}$ of the three-point function of the field fluctuation Q, we can directly relate $f_{\rm NL}$ to the fluctuations in the fields Q^I as follows:

$$f_{\rm NL} \equiv -\frac{5}{6} \frac{\mathcal{A}_{\mathcal{R}}}{\sum_i k_i^3} = -\frac{5}{6} \frac{\partial_I N \partial_J N \partial_K N}{(\partial^L N \partial_L N)^2} \frac{\mathcal{A}^{IJK}}{\sum_i k_i^3} - \frac{5}{6} \frac{\partial_I N \partial_J N \partial^I \partial^J N}{(\partial^L N \partial_L N)^2} + \cdots, \tag{12.1.22}$$

where the dots again denote the undisplayed cross terms and we contracted the field space indices in the second term. We will define the first and second term in the above expressions as $f_{\rm NL}^{(1)}$ and $f_{\rm NL}^{(2)}$, respectively. In the literature, there exist other definitions of the non-linearity parameter $f_{\rm NL}$, which are all equivalent to the one given above [76]. The non-linearity parameter represents the key observable, since it can be related directly to the non-linearity $f_{\rm NL}^{\delta T/T}$ in the CMB anisotropies. For a discussion on the procedure to relate $f_{\rm NL}$ to $f_{\rm NL}^{\delta T/T}$, see [13]. We conclude that in order to find $f_{\rm NL}$, the main task for subsequent sections is to determine the momentum function $\mathcal{A}^{IJK}$ by means of computing the three-point function of the field fluctuation Q^I as given in Eq. 12.1.16.

12.2 Path Integral Formalism for the Three-Point Function

Contrary to the calculation of the bispectrum in the single-field case, we will compute the three-point function (Eq. 12.1.16) in the multi-field scenario by invoking the path integral formalism of quantum field theory, following the approach in [28, 77]. Since we are solely dealing with scalar fields, we will discuss the path integral formalism of QFT for scalar fields here. In this formalism, originally due to Feynman, the central object is the so-called transition amplitude $U(\phi_a, \phi_b, T)$ for a state $|\phi_a\rangle$ to become a state $|\phi_b\rangle$ after some time T,

where ϕ denotes the considered scalar field.[1] Feynman showed that this transition amplitude can be written as a sum (integral) over all possible trajectories in field space between the configurations ϕ_a and ϕ_b. Such an integral is called a functional integral and can be written as:

$$U(\phi_a, \phi_b, T) \equiv \langle \phi_b | e^{-iHT} | \phi_a \rangle = \int_a^b \mathcal{D}\phi(t) \, \exp\left[i \int_0^T d^4x \, \mathcal{L} \right], \tag{12.2.1}$$

where $\mathcal{D}\phi$ denotes the functional integration measure over all field configurations. Notice that the amplitude is a function of the Lagrangian, rather than the Hamiltonian, so that there is no need to transform between the Lagrangian into the Hamiltonian, which was the case for the In-In formalism.

In the path integral formalism, the n-point correlation function denoted as $\langle Q_n \rangle \equiv \langle \phi(t_1, \boldsymbol{x}_1) \cdots \phi(t_n, \boldsymbol{x}_n) \rangle$ can be written as:

$$\boxed{\langle Q_n(t_*) \rangle = \frac{1}{\mathcal{A}} \int \mathcal{D}\phi \, Q_n \exp\left[i \int_{\mathcal{C}}^{t_*} dt \, d^3x \, \mathcal{L} \right],} \tag{12.2.2}$$

where $\mathcal{A}$ is given by:

$$\mathcal{A} \equiv \int \mathcal{D}\phi \, \exp\left[i \int_{\mathcal{C}}^{t_*} dt \, d^3x \, \mathcal{L} \right], \tag{12.2.3}$$

and $\mathcal{C}$ denotes the temporal contour of integration including the time of evaluation t_*, which should be constructed such that at early times the interaction vacuum is effectively turned off (this can be achieved by including the $i\varepsilon$-prescription conform Eq. 8.1.4).

12.2.1 The Generating Functional

In order to derive the above expression for the n-point correlation function, we start by defining the so-called functional derivatives $\delta/\delta J(x)$ as follows:

$$\frac{\delta}{\delta J(x)} J(y) \equiv \delta^{(4)}(x - y), \qquad \frac{\delta}{\delta J(x)} \int d^4y \, J(y)\phi(y) = \phi(x). \tag{12.2.4}$$

Next, we define the generating functional of correlation functions, denoted by $Z[J]$, in the following way:

$$Z[J] \equiv \int \mathcal{D}\phi \, \exp\left[i \int d^4x \, (\mathcal{L} + J(x)\phi(x)) \right], \tag{12.2.5}$$

where the quantity $J(x)$ is called the source term and will play an essential role in generating correlation functions using $Z[J]$. The generating function evaluated in the absence of the source term $J(x) \equiv 0$ is written as $Z_0 \equiv Z[J = 0]$. In terms of the generating functional, the n-point correlation function becomes:

$$\langle Q_n \rangle = \frac{1}{Z_0} \prod_{i=1}^{n} \left(-i \frac{\delta}{\delta J(x_i)} \right) Z[J] \bigg|_{J=0}. \tag{12.2.6}$$

Each functional derivative $\delta/\delta J(x_i)$ brings down a factor $\phi(x_i)$ in the numerator. Performing this procedure n times and setting the source term J to zero in the end exactly yields path integral form of the correlation function.

[1] In this section we will constrain to discussing only one scalar field for (notational) simplicity, although the formalism can be easily extended to several fields, as in the case of multi-field inflation.

12.2.2 Two- and Three-Point Function of Multi-Field Quantum Fluctuations

Now we consider the scalar field to be the multi-field quantum fluctuation $Q^I \equiv \delta\phi^I$ and we aim to obtain a path integral expression for the two- and three-point functions denoted $\langle Q^I Q^J \rangle$ and $\langle Q^I Q^J Q^K \rangle$, respectively. For notational simplicity, we will drop the labels $I, J, K, \cdots$ and consider a single field Q. For a flat field space metric $G_{IJ} = \delta_{IJ}$, the analysis of one field is easily extendable to multiple fluctuation fields Q^I.

In exact analogy with the preceding section, we start by defining the generation functional for Q, which we denote by $Z^{(Q)}[J]$. The Lagrangian for this quantity consists of a quadratic Gaussian[2] part $\mathcal{L}_2$ governing the leading evolution and a cubic interaction contribution $\mathcal{L}_3$. Typically, the cubic Lagrangian is one order higher in slow-roll parameters, so that it is much smaller than the quadratic part (in the slow-roll approximation). The generating functional is thus given by:

$$Z[J]^{(Q)} \equiv \int \mathcal{D}Q \ \exp\left[i \int d^4x \left(\mathcal{L}_2 + \mathcal{L}_3 + J(x)Q(x) \right) \right]. \tag{12.2.8}$$

Two-Point Correlator and Propagator

To start, we will derive the two-point function using the generating functional. In principle, the two-point function receives contributions from both S_2 and S_3, however since the two-point correlation function involves an even number of Q's that are integrated over, the contribution related to e^{iS_2} does not vanish (see Eq. 12.2.7). Using this observation and the fact that $S_3 \ll S_2$ we can approximate the two-point function to reasonable accuracy by considering only the quadratic action. The generating function now reads:

$$Z_0^{(Q)}[J; \mathcal{L}_2] = \int \mathcal{D}Q \ e^{iS_2}, \tag{12.2.9}$$

where denote the fact the generating functional only depends on the quadratic Lagrangian by explicitly writing $\mathcal{L}_2$ between square brackets. This approach is convenient, since it is now possible to construct a propagator for Q [28, 77] (in the de Sitter limit), which can be used to easily compute higher order correlation functions.

Considering only the quadratic action S_2, it is possible to write the generating function in a very explicit form involving the field propagator $D_Q(x_1 - x_2)$ (this propagator will be derived below, with the explicit final form given by Eq. 12.4.13). Performing the field shift:

$$\tilde{Q}(x) \equiv Q(x) - i \int d^4x D_Q(x - x_2) J(x_2), \tag{12.2.10}$$

we can rewrite the generating functional $Z_0^{(Q)}[J; \mathcal{L}_2]$ as follows:

$$Z_0^{(Q)}[J; \mathcal{L}_2] = Z_0[\mathcal{L}_2] \ \exp\left[-\frac{1}{2} \int d^4x \int d^4y \ J(x) D_Q(x - y) J(y) \right]. \tag{12.2.11}$$

[2]The quadratic action is Gaussian in the sense that integrating an odd number of Q's against e^{iS_2} yields zero:

$$\int \mathcal{D}Q \ Q^n \ e^{iS_2} = 0, \tag{12.2.7}$$

for n odd. This is a direct consequence of Wick's or Isserlis' theorem, see section 3.2.

For a derivation of this result, see section 9.2 in [72]. Now applying Eq. 12.2.6 and using the above result for the generating functional, we can perform the two-point correlation function as follows:

$$\begin{aligned}\langle Q(x_1)Q(x_2)\rangle &= -\frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_2)}\exp\left[-\frac{1}{2}\int d^4x\int d^4y\; J(x)D_Q(x-y)J(y)\right]\\ &= D_Q(x_1-x_2),\end{aligned} \tag{12.2.12}$$

where we have used the properties of the functional derivative (Eq. 12.2.4) and defined $x_1 \equiv (t_1, \boldsymbol{x}_1)$. This verifies that the two-point function is indeed equal to the propagator. In terms of multiple fields and a flat field space metric, the generalized propagator becomes:

$$\langle Q^I(x_1)Q^J(x_2)\rangle = \delta^{IJ}D_Q(x_1-x_2). \tag{12.2.13}$$

Three-Point Correlator

The three-point function of Q, evaluated at the time of interest t_*, can then be found by application of Eq. 12.2.6, yielding:

$$\begin{aligned}\langle Q(t_*,\boldsymbol{x}_1)Q(t_*,\boldsymbol{x}_2)Q(t_*,\boldsymbol{x}_3)\rangle &= \frac{1}{Z_0^{(Q)}}\prod_{i=1}^{3}\left(-i\frac{\delta}{\delta J(x_i)}\right)Z^{(Q)}[J]\Big|_{J=0}\\ &= \frac{i}{Z_0^{(Q)}}\int \mathcal{D}Q\; Q(t_*,\boldsymbol{x}_1)Q(t_*,\boldsymbol{x}_2)Q(t_*,\boldsymbol{x}_3)\\ &\quad\times\exp\left[i\int_{\mathcal{C}}^{t_*}dt'\; d^3x\;\big(\mathcal{L}_2(t')+\mathcal{L}_3(t')\big)\right].\end{aligned} \tag{12.2.14}$$

Notice that the exponential could also be written using the shorthand notation:

$$e^{iS} = e^{i(S_2+S_3)} = e^{iS_2}(1+iS_3+\mathcal{O}(S_3^2)), \tag{12.2.15}$$

where we have expanded the S_3 term to linear order since we know that in the slow-roll approximation $S_3 \ll S_2$. Notice that this expansion implicitly assumes that the slow-roll approximation is still valid at the moment of evaluation of the bispectrum t_*. This not emphatically the case at the end of inflation, when the slow-roll parameters become of order unity and the quadratic and cubic action are of the same order and hence the above expansion to first order in S_3 ceases to be valid.

Having mentioned this subtlety, the above expansion is very convenient as the zeroth order term in the expansion of e^{iS_3}, which just becomes e^{iS_2}, will vanish when it is integrated over in combination with an odd number of Q's (Eq. 12.2.7). In particular, for the three-point function, we obtain the vanishing term:

$$\int \mathcal{D}Q\; Q(t_*,\boldsymbol{x}_1)Q(t_*,\boldsymbol{x}_2)Q(t_*,\boldsymbol{x}_3)\; e^{iS_2} = 0. \tag{12.2.16}$$

Since the functional integral in the denominator $Z_0^{(Q)}$ has no field-dependence contained in the integrand except for the intrinsic Q-dependence in the action $S[Q]$, the above reasoning is not applicable. Instead we use the simple observation that $S_3 \ll S_2$, so that to leading order we can approximate the total generating function to be equal to the generating function constructed from solely the quadratic Lagrangian (Eq. 12.2.11):

$$Z^{(Q)}[J] \simeq Z^{(Q)}[J;\mathcal{L}_2]. \tag{12.2.17}$$

Using the above results, we can write the three-point function (Eq. 12.2.14) as:

$$\langle Q(t_*,\boldsymbol{x}_1)Q(t_*,\boldsymbol{x}_2)Q(t_*,\boldsymbol{x}_3)\rangle = \frac{i}{Z_0^{(Q)}[\mathcal{L}_2]}\int \mathcal{D}Q\; Q(t_*,\boldsymbol{x}_1)Q(t_*,\boldsymbol{x}_2)Q(t_*,\boldsymbol{x}_3) \times \int_{\mathcal{C}}^{t_*} dt'\; d^3x\; \mathcal{L}_3(t')\; e^{iS_2(t')} \equiv \frac{\mathcal{A}}{Z_0^{(Q)}[\mathcal{L}_2]}, \tag{12.2.18}$$

where $Z_0^{(Q)}[\mathcal{L}_2]$ is the generating function containing solely the quadratic action evaluated at $J=0$ (vanishing source term). For later convenience, the numerator is denoted as $\mathcal{A}$.

Now we aim to write the above expression for the three-point function in terms of the propagator D_Q. However, we have to be careful with the time-dependence, as we will illustrate below by means of an example. Consider a term in the cubic Lagrangian of the form:

$$\mathcal{L}_3(t',\boldsymbol{x}) \supset g\dot{Q}^2\partial^{-2}\dot{Q}, \tag{12.2.19}$$

where g is a coupling constant, typically proportional to slow-roll parameters. Substituting the expression for $\mathcal{L}_3$ in the functional integral in the numerator of Eq. 12.2.18, denoted as $\mathcal{A}$, we can explicitly perform the integral as follows:

$$\mathcal{A} = ig \times Z_0^{(Q)}[\mathcal{L}_2]\int_{\mathcal{C}}^{t_*} dt'\; d^3x\; \langle Q_{123}(t_*)\; \partial_{t'}Q(t',\boldsymbol{x})\; \partial_{t'}Q(t',\boldsymbol{x})\; \partial_{\boldsymbol{x}}^{-2}\partial_{t'}Q(t',\boldsymbol{x})\rangle, \tag{12.2.20}$$

where Q_{123} is defined as:

$$Q_{123}(t_*) \equiv Q(t_*,\boldsymbol{x}_1)Q(t_*,\boldsymbol{x}_2)Q(t_*,\boldsymbol{x}_3). \tag{12.2.21}$$

Now, we obtain the following explicit expression for the contribution of the term $\mathcal{L}_3(t',\boldsymbol{x}) \supset g\dot{Q}^2\partial^{-2}\dot{Q}$ to the bispectrum:

$$\langle Q(t_*,\boldsymbol{x}_1)Q(t_*,\boldsymbol{x}_2)Q(t_*,\boldsymbol{x}_3)\rangle_{\dot{Q}^2\partial^{-2}\dot{Q}} = ig\int_{\mathcal{C}}^{t_*} dt'\; d^3x\; \mathcal{G}, \tag{12.2.22}$$

where the integrand $\mathcal{G}$ reads:

$$\mathcal{G} \equiv \langle Q_{123}(t_*)\; \partial_{t'}Q(t',\boldsymbol{x})\; \partial_{t'}Q(t',\boldsymbol{x})\; \partial_{\boldsymbol{x}}^{-2}\partial_{t'}Q(t',\boldsymbol{x})\rangle. \tag{12.2.23}$$

Using Wick's theorem, we can subsequently expand the six-point correlator in $\mathcal{G}$ in terms of propagators (i.e. normal ordered two-point functions). The result is:

$$\begin{aligned}\mathcal{G} &= 2\sum_{\text{perms}} \partial_{t'}\langle Q(t_*,\boldsymbol{x}_1)Q(t',\boldsymbol{x})\rangle\; \partial_{t'}\langle Q(t_*,\boldsymbol{x}_2)Q(t',\boldsymbol{x})\rangle\; \partial_{\boldsymbol{x}}^{-2}\; \partial_{t'}\langle Q(t_*,\boldsymbol{x}_3)Q(t',\boldsymbol{x})\rangle \\ &= 2\sum_{\text{perms}} \partial_{t'}D_Q(t_*-t')\; \partial_{t'}D_Q(t_*-t')\; \partial_{\boldsymbol{x}}^{-2}\partial_{t'}D_Q(t_*-t'),\end{aligned} \tag{12.2.24}$$

where the second line is in terms of the propagator and we have suppressed its spatial dependence in the argument. The factor of two in front arises from the possibility to form the product of two-point functions with the $\partial_{t'}Q_I$ and $\partial_{t'}Q^I$ terms separately. In principle, we should sum over all permutations, i.e. over all ways of pairing the so-called *external* fields $Q(t_*,\boldsymbol{x}_i)$ with the *vertex* fields contained in the term contributing to $\mathcal{L}_3$. However, we only sum over pairings involving one of the external and one of the vertex fields in the $\mathcal{L}_3$-term [76, 77]. Other pairings give rise to tadpoles, governing vacuum processes in which Q-quanta spontaneously emerge form the vacuum. Since we assume the vacuum to be stable, such tadpoles vanish and we can neglect the corresponding pairings. This is equivalent to saying that we solely work with the connected three-point function.

12.3 Multi-Field Action in the ADM Formalism

For the same reasons as discussed in section 9.1, we will write the multi-field action:

$$S = \int d^4x\sqrt{-g}\left[\frac{1}{2}M_{\rm pl}^2 R - \frac{1}{2}\delta_{IJ}g^{\mu\nu}\partial_\mu\phi^I\partial_\nu\phi^J - V(\phi^I)\right], \tag{12.3.1}$$

in the ADM formalism, where the index I runs from 1 to n, where n is the number of fields. Recall that the line element of the ADM formalism is given by (Eq. 9.2.11):

$$ds^2 = -N^2\,dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt). \tag{12.3.2}$$

For a single field, the ADM decomposition of the action yields (Eq. 9.5.10):

$$\begin{aligned} S = &\frac{1}{2}\int d^4x\; N\sqrt{h}\Big[M_{\rm pl}^2 R^{(3)} - \partial_i\phi\partial^i\phi - 2V\Big] \\ &+ \frac{1}{2}\int d^4x\; N^{-1}\sqrt{h}\Big[E_{ij}E^{ij} - E^2 + (\dot{\phi} - N^i\partial_i\phi)^2\Big]. \end{aligned} \tag{12.3.3}$$

In exact analogy, we will decompose the multi-field action. The only difference is that we will have generalize the single-field to multiple field by including the field space metric $G_{IJ} = \delta_{IJ}$ in order to sum over the field. The result is as follows:

$$\begin{aligned} S = &\frac{1}{2}\int d^4x\; N\sqrt{h}\Big[M_{\rm pl}^2 R^{(3)} - \delta_{IJ}h^{ij}\partial_i\phi^I\partial_j\phi^J - 2V\Big] \\ &+ \frac{1}{2}\int d^4x\; N^{-1}\sqrt{h}\Big[E_{ij}E^{ij} - E^2 + G_{IJ}v^I v^J\Big], \end{aligned} \tag{12.3.4}$$

where we have defined $v^I \equiv \dot{\phi}^I - N^i\partial_i\phi^I$, so that the field conjugate momentum becomes:

$$\pi^2 \equiv N^{-2}\delta_{IJ}v^I v^J, \tag{12.3.5}$$

which generalizes the variable π_ϕ used in the single-field analysis.

In order to solve for the non-dynamical Lagrange multipliers N (lapse) and N^i (shift), we vary the above action with respect to them to obtain the constraint equations. Imposing N to be stationary under variations, we find the constraint:

$$-\delta_{IJ}h^{ij}\partial_i\phi^I\partial_j\phi^I - 2V - N^{-2}(E_{ij}E^{ij} - E^2 + \delta_{IJ}v^I v^J) = 0. \tag{12.3.6}$$

Similarly, the constraint for the shift N^i becomes:

$$\mathcal{D}_j\Big[N^{-1}(E_i^j - E\delta_j^i)\Big] = N^{-1}\delta_{IJ}v^I\partial_i\phi^J, \tag{12.3.7}$$

where $\mathcal{D}_j$ is the covariant derivative associated with the induced metric h_{ij}.

We will follow Seery and Lidsey [76] and work in the flat gauge, defined for scalar perturbations via:

$$\delta\phi^I \equiv Q^I, \qquad h_{ij} = a^2\delta_{ij}, \tag{12.3.8}$$

i.e. we have set curvature perturbation to zero ($\Psi \equiv 0$) in order to obtain spatially flat slices. The solutions for N and N^i are parametrized to first order, considering scalar perturbations only, as follows:

$$N = 1 + N^{(1)}, \qquad N^i = \partial^i\psi. \tag{12.3.9}$$

To find the expressions for $N^{(1)}$ and $\partial^i\psi$, we will perturb the constraint equations to first order by splitting the field into a background part and a perturbation:

$$\phi^I = \bar{\phi}^I + Q^I, \tag{12.3.10}$$

where the fluctuation is denoted by $Q^I \equiv \delta\phi^I$. From now on, there will be no need to distinguish between the total and background field, so we drop the bar on the background field. Since the details are explained for the single-field calculation, we will not explicitly derive the corresponding expressions for the multi-field case here, but instead mention the results:

$$N^{(1)} = \frac{1}{2H}\delta_{IJ}\phi^I Q^J, \qquad \partial^2\psi = \frac{a^2}{2H}\left(Q_I\ddot{\phi}^I - \dot{Q}_I\dot{\phi}^I - \frac{\dot{H}}{H^2}\dot{\phi}_I Q^I\right), \tag{12.3.11}$$

Notice that in the slow-roll limit, the second term in $\partial^2\psi$ dominates the other two. The above solutions can be substituted in the action in order to recast the action solely in terms of physical degrees of freedom.

12.4 Second-Order Action

In exact analogy to the single-field case, we can plug the first order solutions to N and N^i in the ADM action and obtain the second order and third order action. In this section, we focus on the quadratic action, which is given by (after a few integrations by parts) [76]:

$$S_2 = \frac{1}{2}\int d^4x\, a^3\left[\delta_{IJ}\dot{Q}^I\dot{Q}^J - \frac{1}{a^2}\delta_{IJ}\partial Q^I\partial Q^J - \mathcal{M}_{IJ}Q^IQ^J\right], \tag{12.4.1}$$

where we re-introduced the notation $\delta^{ij}\partial_i Q^I\partial_j Q^J \equiv \partial Q^I\partial Q^J$. The so-called mass matrix $\mathcal{M}_{IJ}$ is defined as:

$$\mathcal{M}_{IJ} = \partial_I\partial_J V - \frac{1}{a^3}\frac{d}{dt}\left(\frac{a^3}{H}\dot{\phi}_I\dot{\phi}_J\right). \tag{12.4.2}$$

Notice that the mass matrix is principle not diagonal and hence couples different fields to each other. Changing to conformal time, momentum space and introducing the canonical Mukhanov-Sasaki variable $f^I \equiv aQ^I$, the equation of motion resulting from the quadratic action reads:

$$\partial_{\tau\tau}f^I + \left(\left[k^2 - \frac{a''}{a}\right]\delta^I_J + a^2\mathcal{M}^I_J\right)f^J = 0, \tag{12.4.3}$$

where we dropped the subscript k denoting that we are in working in Fourier space for notational convenience. Again, since $\mathcal{M}_{IJ}$ is typically not diagonal, the above equations of motion for f^I do not factorize into n copies of the Mukhanov-Sasaki equation (one for each field).

However, it can be shown that in the de Sitter limit, where the slow-roll parameters all tend to zero, the above equations of motion, in fact, do reduce to n copies of the Mukhanov-Sasaki equation. In order to verify this statement, we will have to show that the mass matrix can be expressed entirely in terms of slow-roll parameters and hence $\mathcal{M}_{IJ}$ vanishes in the de Sitter limit. Evaluating the time derivative in the definition of the mass matrix and invoking

the definitions of the multi-field slow-roll parameters, it can be written as:

$$\begin{aligned}\frac{1}{H^2}\mathcal{M}_{IJ} &= \frac{1}{H^2}\partial_I\partial_J V - \frac{1}{H^2a^3}\frac{d}{dt}\left(\frac{a^3}{H}\dot{\phi}_I\dot{\phi}_J\right)\\ &= 3\tilde{\eta}_{IJ} - \frac{1}{H^2a^3}\left(\frac{3a^2\dot{a}}{H}\dot{\phi}_I\dot{\phi}_J - \frac{\dot{H}a^3}{H^2}\dot{\phi}_I\dot{\phi}_J + \frac{a^3}{H}(\dot{\phi}_I\ddot{\phi}_J + \dot{\phi}_J\ddot{\phi}_I)\right)\\ &= 3\tilde{\eta}_{IJ} - 6\varepsilon_{IJ} - 2\varepsilon\varepsilon_{IJ} + 4\varepsilon\eta_{IJ}.\end{aligned} \tag{12.4.4}$$

Hence, we conclude that the mass matrix is indeed proportional to slow-roll parameters and therefore it is a sub-leading effect which can be neglected in the de Sitter limit. The equations of motion then become n copies of the Mukhanov-Sasaki equation of motion, one for each field:

$$\partial_{\tau\tau}f_I + \left(k^2 - \frac{a''}{a}\right)f_I = 0, \tag{12.4.5}$$

where using the relation $\tau = -1/aH$ we can replace $a''/a = 2/\tau^2$. Assuming all fields start in the Bunch-Davies vacuum (see Eq. 5.4.23), we can write the solutions as:

$$f_k^I(\tau) = \frac{e^{-ik\tau}}{\sqrt{2k}}\left(1 - \frac{i}{k\tau}\right). \tag{12.4.6}$$

By invoking the above result for the mode function in the de Sitter limit, we can construct a propagator which allows us to conveniently evaluate higher-order correlation function (see section 12.2.2). In order to construct this propagator, first quantize the f^I as usual:

$$f^I(\tau) = (f^+)^I + (f^-)^I = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\left[f_k^I(\tau)\hat{a}_{\boldsymbol{k}} + f_k^{I*}(\tau)\hat{a}_{\boldsymbol{k}}^\dagger\right]e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{12.4.7}$$

with the creation and annihilation operators satisfying the commutation relation defined in Eq. 5.3.9. Now, we consider the two point correlation function and drop the labels denoted different fields:

$$\langle f(\tau_1, \boldsymbol{x}_1)f(\tau_2, \boldsymbol{x}_2)\rangle. \tag{12.4.8}$$

In evaluating $\langle f_k(\tau_1)f_k(\tau_2)\rangle$ and constructing the corresponding propagator, we have to be careful with the time-ordening. We will apply normal time ordening by placing all operators evaluated at a later time to the left:

$$T\ f(\tau_1)f(\tau_2) = \begin{cases} f(\tau_1)f(\tau_2) & (\tau_1 > \tau_2),\\ f(\tau_2)f(\tau_1) & (\tau_2 > \tau_1).\end{cases} \tag{12.4.9}$$

On account of Wick's theorem, we can now define the the following propagator:

$$D_f(\tau_1 - \tau_2) = [f^+(\tau_1), f^-(\tau_2)] = (2\pi)^3\ \delta^{(3)}(\boldsymbol{k}_1 - \boldsymbol{k}_2) \times f_{k_1}(\tau_1)f_{k_2}^*(\tau_2). \tag{12.4.10}$$

In terms of this propagator, the above time ordering operator acting on two fields can be written as:

$$T\ f(\tau_1)f(\tau_2) = \begin{cases} f(\tau_1)f(\tau_2) = :f(\tau_1)f(\tau_2): + D_f(\tau_1 - \tau_2) & (\tau_1 > \tau_2),\\ f(\tau_2)f(\tau_1) = :f(\tau_1)f(\tau_2): + D_f(\tau_2 - \tau_1) & (\tau_2 > \tau_1).\end{cases} \tag{12.4.11}$$

Inserting this result into the two-point function and integrating over momentum $\boldsymbol{k}_2$ and setting $\boldsymbol{k} \equiv \boldsymbol{k}_1$, we obtain:

$$\langle T\, f(\tau_1) f(\tau_2)\rangle = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \times \begin{cases} f_k(\tau_1) f_k^*(\tau_2) & (\tau_1 > \tau_2), \\ f_k(\tau_2) f_k^*(\tau_1) & (\tau_2 > \tau_1). \end{cases} \tag{12.4.12}$$

Now using the form of $f_k(\tau)$ as given by Eq. 12.4.6, using $Q = f/a$ and $\tau = -1/aH$, we obtain following propagator for Q in momentum space:

$$D_Q(\tau_1 - \tau_2) = \frac{H^2}{2k^3} \times \begin{cases} (1 + ik\tau_1)(1 - ik\tau_2) e^{-ik(\tau_1 - \tau_2)} & (\tau_1 > \tau_2), \\ (1 + ik\tau_2)(1 - ik\tau_1) e^{+ik(\tau_1 - \tau_2)} & (\tau_1 > \tau_2), \end{cases} \tag{12.4.13}$$

Introducing the labels I and J again and using the flat field space metric δ^{IJ}, we can relate the two-point function in real space to the above propagator in momentum space via:

$$\langle Q^I(\tau_1, \boldsymbol{x}_1) Q^J(\tau_2, \boldsymbol{x}_2)\rangle = \delta^{IJ} \int \frac{d^3\boldsymbol{k}}{(2\pi)^3}\, D_Q(\tau_1 - \tau_2) \tag{12.4.14}$$

Taking the coincidence limit of the above two-point function and corresponding propagator outside the horizon (where $k \to 0$), we obtain the familiar result in momentum space:

$$\langle Q^I_{\boldsymbol{k}_1}(\tau_1) Q^J_{\boldsymbol{k}_2}(\tau_2)\rangle = \delta^{IJ} (2\pi)^3 \delta^{(3)}(\boldsymbol{k}_1 - \boldsymbol{k}_2) \frac{H^2}{2k^3}, \tag{12.4.15}$$

so that the power spectrum P_Q (setting $I = J$) becomes:

$$\mathcal{P}_Q = \left(\frac{H}{2\pi}\right)^2, \tag{12.4.16}$$

which corresponds to the spectrum of a massless scalar field in de Sitter space. This result verifies the statement given in Eq. 12.1.14 and concludes our discussion on the second-order theory.

12.5 Third-Order Action

Now we move away from the quadratic action, describing the leading evolution of the fields and instead consider the third order action, governing interaction. Inserting the first-order solutions for the lapse and shift in the ADM action and expanding the action up to third order, we find that to leading order the cubic action S_3 is given by [76]:

$$\begin{aligned} S_3 = \int d^4x\, a^3 \Big[& -\frac{1}{a^2} \delta_{IJ} \dot{Q}^I \partial\psi \partial Q^J - \frac{1}{4H} \delta_{IJ} \delta\dot{\phi}^M Q^N \dot{Q}^I \dot{Q}^J \\ & - \frac{1}{a^2} \frac{1}{4H} \delta_{IJ} \delta_{MN} \dot{\phi}^M Q^N \partial Q^I \partial Q^J \Big]. \end{aligned} \tag{12.5.1}$$

Since we work only up to leading order in the slow-roll parameters, we can approximate ψ as (see Eq. 12.3.11):

$$\partial^2 \psi \simeq -\frac{a^2}{2H} \delta_{IJ} \dot{\phi}^I Q^J. \tag{12.5.2}$$

12.5.1 Terms Related to the Equations of Motion

The aim is to recast the cubic action in a form that resembles Eq. 10.3.20, which includes terms that are directly related to the quadratic equations of motion. As we have seen in the single-field calculation, such terms represent a field redefinition [64, 76]. In the present case, we will find that we can rewrite the cubic action including terms that are proportional to the first order equations of motion for Q^I. Subsequently, those terms will be removed by means of a field redefinition similar to the one used in the single-field calculation.

We start by integrating the first term in S_3 by parts to obtain:

$$-\int d^4x\, a\delta_{IJ}\dot{Q}^I\partial\psi\partial Q^J = \int d^4x\Big(-\frac{aH}{2}\delta_{IJ}\psi\partial Q^I\partial Q^J - \frac{1}{2}a\dot{\psi}\delta_{IJ}\partial Q^I\partial Q^J + a\delta_{IJ}\dot{Q}^I\psi\partial^2 Q^J\Big). \tag{12.5.3}$$

Computing the time derivative of ψ and keeping only leading order terms in the slow-roll parameters, we find:

$$\dot{\psi} = -a^2\delta_{IJ}\dot{\phi}^I\partial^{-2}\dot{Q}^J - \frac{a^2}{2H}\delta_{IJ}\dot{\phi}^I\partial^{-2}\ddot{Q}^J. \tag{12.5.4}$$

Notice that the leading order expression for $\dot{\psi}$ involves terms containing a second-order time derivative $\ddot{Q}^J$ of the field fluctuation. Therefore, substituting $\dot{\psi}$ directly into the cubic action would change the order of the field equations for Q^I. Instead, we will use the first order perturbed Klein-Gordon equation to eliminate the second order time derivative of Q^I.

The Klein-Gordon equation for field Q^I reads:

$$\frac{1}{\sqrt{-g}}\partial_\nu\left(\delta_{IJ}\sqrt{-g}g^{\mu\nu}\partial_\mu\phi^I\right) = \partial_J V(\phi), \tag{12.5.5}$$

where we can write the metric determinant in the ADM form by using $\sqrt{-g} = N\sqrt{h}$. Since we work in the flat gauge, the determinant of the induced metric is simply $\sqrt{h} = a^3$. Therefore, only the field and lapse N should be expanded to first order to find the equation of motion for the fluctuation Q^I. Moving the term $1/\sqrt{-g}$ to the right-hand-side and perturbing this side first, we find:

$$\begin{aligned} N\sqrt{h}\,\partial_J V(\phi) &= a^3\left(1 + \frac{1}{2H}\delta_{MN}\dot{\phi}^N Q^M\right)\partial_J V(\phi) \\ &= a^3\left(1 + \frac{1}{\sqrt{2}}\delta_{MN}\varepsilon^N Q^M\right)\partial_J V(\phi) \\ &= a^3\partial_J V(\phi) + \mathcal{O}(\varepsilon), \end{aligned} \tag{12.5.6}$$

where we substituted the first order expression for N up to leading order terms. Similarly, expanding the left-hand-side yields:

$$\begin{aligned} \partial_\nu\left(N\sqrt{h}g^{\mu\nu}\partial_\mu\phi_J\right) &= -\partial_0(Na^3\partial\phi_J) + \partial_j(Na^3 g^{ij}\partial_i\phi_J) \\ &= -\partial_0\left(a^3(\dot{\bar{\phi}}_J + \dot{Q}_J)\right) + a^3\nabla^2 Q_J + \mathcal{O}(\varepsilon), \end{aligned} \tag{12.5.7}$$

where we have expanded the field into a background part and fluctuation. For the lapse $N = 1 + N^{(1)}$, the first-order contribution $N^{(1)}$ is proportional to slow-roll parameters and

therefore sources the neglected term $\mathcal{O}(\varepsilon)$. Equating both sides and using the background Klein-Gordon equation, we find that the equation of motion for the fluctuation is given by:

$$-\ddot{Q}_J - 3H\dot{Q}_J + \nabla^2 Q_J = 0 + \mathcal{O}(\varepsilon). \tag{12.5.8}$$

Expanding the ADM action to first order in Q_I, we would find that the variation $\delta\mathcal{L}/\delta Q_I$ is related to the above equation of motion as [76]:

$$\frac{1}{a^3}\frac{\delta\mathcal{L}_1}{\delta Q^I} = -\ddot{Q}_I - 3H\dot{Q}_I + \frac{1}{a^2}\partial^2 Q_I + \mathcal{O}(\varepsilon). \tag{12.5.9}$$

Solving the above equation for $\ddot{Q}_I$, substituting this expression into Eq. 12.5.4 and using this equation in turn in Eq. 12.5.3 for the first term of S_3, we find the following expression for S_3 after a few integrations by parts:

$$S_3 = \int d^4x\, a^3 \Big[-\frac{1}{4H}\dot{\phi}^J Q_J \dot{Q}_I \dot{Q}^I - \frac{1}{2H}\dot{\phi}^J \partial^{-2}\dot{Q}_J \dot{Q}_I \partial^2 Q^J + \frac{1}{a^3}\frac{\delta\mathcal{L}_1}{\delta Q_I} G_{IJ} F^J \Big], \tag{12.5.10}$$

where the function F is defined as:

$$F^I(Q) = \frac{1}{4H}\dot{\phi}^I \partial^{-2}(Q^J \partial^2 Q_J) - \frac{1}{8H}\dot{\phi}^I Q_J Q^J. \tag{12.5.11}$$

12.5.2 Field Redefinition

In exact analogy to the single-field calculation [64], it is possible to remove the terms proportional to the equations of motion $\delta\mathcal{L}_1/\delta Q_I$ via a field redefinition. The required field redefinition reads:

$$Q^I = \mathcal{Q}^I - F^I(\mathcal{Q}). \tag{12.5.12}$$

Applying this field redefinition to the cubic action induces only new terms at fourth order in $\mathcal{Q}^I$. However, we also have to apply this shift to the quadratic action (see section 10.6.2 for the single-field analog). The result of the field redefinition reads:

$$S_2[Q^I] = S_2[\mathcal{Q}^I] - \int d^4x\, a^3 \delta_{IJ} \dot{\mathcal{Q}}^I F^J + \int d^4x\, a\delta_{IJ}\partial\mathcal{Q}^I \partial F^I + \int d^4x\, a^3 \mathcal{M}_{IJ}\mathcal{Q}^I F^J, \tag{12.5.13}$$

where we neglect the term proportional to $\mathcal{M}_{IJ}$ from now on since it is directly proportional to slow-roll parameters and hence is not a leading order effect. Partially integrating the second and third term in the above expansion omitting boundary terms, we find:

$$\begin{aligned} \int d^4x\, a^3 \delta_{IJ} \dot{\mathcal{Q}} \dot{F}^J &= -\int d^4x\, \delta_{IJ} \partial_t (a^3 \dot{\mathcal{Q}}^I) F^J, \\ \int d^4x\, a\delta_{IJ}\partial\mathcal{Q}^I \partial F^J &= -\int d^4x\, a(\partial^2 \mathcal{Q}^I) F^J. \end{aligned} \tag{12.5.14}$$

Inserting these results, we find that the action $S_2[Q^I]$ splits into the quadratic action for the redefined field $\mathcal{Q}^I$ and first order equation of motion:

$$\begin{aligned} S_2[Q^I] &= S_2[\mathcal{Q}^I] + \int d^4x\, \delta_{IJ} \Big[\partial_t (a^3 \dot{\mathcal{Q}}^I - a\partial^2 \mathcal{Q}^I) \Big] F^J \\ &= S_2[\mathcal{Q}^I] - \int d^4x\, \delta_{IJ} F^J \frac{\delta\mathcal{L}_1}{\delta Q_I}. \end{aligned} \tag{12.5.15}$$

The term proportional to the linear equation of motion exactly cancels the term last term S_3. The cubic action for the field $\mathcal{Q}^I$ thus reads:

$$S_3[\mathcal{Q}^I] = -\int d^4x \frac{a^3}{2H}\left[\frac{1}{2}\dot{\phi}^J \mathcal{Q}_J \dot{\mathcal{Q}}_I \dot{\mathcal{Q}}^I + \dot{\phi}^J \partial^{-2}\dot{\mathcal{Q}}_J \dot{\mathcal{Q}}_I \partial^2 \mathcal{Q}^I\right], \tag{12.5.16}$$

which is the final form that we will use to compute the leading order bispectrum for multi-field inflation in the next section.

Ultimately, we are interested in the three-point correlation function of Q^I, rather than the redefined field $\mathcal{Q}^I$. To take this field redefinition into account, we first compute the three-point correlation function of $\mathcal{Q}^I$ and then relate it to the one for Q^I via the relation:

$$\langle Q^I Q^J Q^K\rangle = \langle \mathcal{Q}^I \mathcal{Q}^J \mathcal{Q}^K\rangle - \langle F^I \mathcal{Q}^J \mathcal{Q}^K\rangle + 2 \text{ perms.} \tag{12.5.17}$$

Now we proceed by computing the leading order three-point function for multi-field inflation in the next section.

12.6 Leading Bispectrum for Multi-Field Inflation

In this section, we will finally compute the leading order bispectrum of multi-field inflation with n minimally coupled canonical scalar fields, i.e. they all have a standard kinetic term and are only coupled via gravity. The cubic action that will be considered is the one presented in Eq. 12.5.16. For notational convenience, we will express S_3 in terms of Q^I instead of $\mathcal{Q}^I$, keeping in mind that we will have to add the correction terms conform Eq. 12.5.17 in order to obtain the correct leading order bispectrum. The cubic action under consideration thus reads (in conformal time):

$$S_3 = \int d\tau\, d^3x \Big[g^J_{(1)} Q_J \partial_\tau Q_I \partial_\tau Q^I + g^I_{(2)} (\partial^{-2}\partial_\tau Q_I)\partial_\tau Q_J (\partial^2 Q^J)\Big], \tag{12.6.1}$$

where the effective coupling constants $g^I_{(1,2)}$ are defined as:

$$g^I_{(1)} \equiv -\frac{a^2}{4H}\dot{\phi}^I, \qquad g^I_{(2)} \equiv 2g^I_{(1)}. \tag{12.6.2}$$

12.6.1 Interaction Term g^I_1

We will first calculate the contribution due to the first term in S_3 explicitly, corresponding to the cubic Lagrangian term:

$$\mathcal{L}_3 \supset g^M_{(1)}\, \delta^{NL} Q_M \partial_\tau Q_N \partial_\tau Q_L. \tag{12.6.3}$$

Following the example given in section 12.2.2, we find that we can write the contribution of this term to three-point correlation function as:

$$\langle Q^I_{\boldsymbol{k}_1} Q^J_{\boldsymbol{k}_2} Q^K_{\boldsymbol{k}_3}\rangle_{(1)} = i\int_{-\infty^+}^{\tau_*} d\tau'\, g^M_{(1)} \delta^{NL}\, \mathcal{G}_{(1)}, \tag{12.6.4}$$

where the time of evaluation τ_* will be taken to tend to zero and we effectively turned of the interaction Lagrangian by means of the $i\varepsilon$ in the early time limit. In momentum space, the integrand is given by:

$$\mathcal{G}_{(1)} = \delta^I_M \delta^J_N \delta^K_L \langle Q_{\boldsymbol{k}_1}(\tau_*) Q_{\boldsymbol{k}}(\tau')\rangle\, \partial_{\tau'}\langle Q_{\boldsymbol{k}_2}(\tau_*) Q_{\boldsymbol{k}}(\tau')\rangle\, \partial_{\tau'}\langle Q_{\boldsymbol{k}_3}(\tau_*) Q_{\boldsymbol{k}}(\tau')\rangle + 2 \text{ perms.} \tag{12.6.5}$$

Now taking the limit $\tau_* \to 0$ and using the de Sitter propagator (Eq. 12.4.13), we find that the different two-point functions in the integrand $\mathcal{G}_{(1)}$ are given by:

$$\langle Q_{\boldsymbol{k}_1}(0)Q_{\boldsymbol{k}}(\tau')\rangle = \frac{H^2}{2k_1^3}(1-ik_1'\tau')e^{+ik_1\tau'}, \qquad \partial_{\tau'}\langle Q_{\boldsymbol{k}_2}(0)Q_{\boldsymbol{k}}(\tau')\rangle = \frac{H^2}{2k_2^3}k_2^2\tau'\, e^{ik_2\tau'}. \tag{12.6.6}$$

Using these results in the integrand $\mathcal{G}_{(1)}$, we can finally write the time-integral expression for contribution of the first term to the bispectrum in the form:

$$\langle Q^I_{\boldsymbol{k}_1}Q^J_{\boldsymbol{k}_2}Q^K_{\boldsymbol{k}_3}\rangle_{(1)} = i\int_{-\infty^+}^{\tau_*} d\tau'\, g^I_{(1)}\mathcal{G}^{JK}\frac{H^6}{4k_{123}^3}(k_2k_3)^2(1-ik_1\tau')\tau'^2\, e^{+iK\tau'} \tag{12.6.7}$$

Performing the time integral and inserting the explicit expression for the coupling constant $g^I_{(1)}$, we find the contribution to bispectrum reads:

$$\langle Q^I_{\boldsymbol{k}_1}Q^J_{\boldsymbol{k}_2}Q^K_{\boldsymbol{k}_3}\rangle_{(1)} = -\frac{H^3}{2}\frac{1}{8k_{123}^3}\sum_{\text{perms}}\dot{\phi}^I\delta^{JK}\left[\frac{(k_2k_3)^2}{K}+\frac{k_1(k_2k_3)^2}{K^2}\right]. \tag{12.6.8}$$

12.6.2 Interaction Term g^I_2

Next, we consider the interaction term proportional to the coupling constant $g^I_{(2)}$. The corresponding cubic Lagrangian term is given in conformal time by:

$$\mathcal{L}_3 \supset g^I_{(2)}(\partial^{-2}\partial_\tau Q_I)\partial_\tau Q_J(\partial^2 Q^J). \tag{12.6.9}$$

Performing the computations similar to the ones given for the first interaction term, one finds the intermediate and final expressions:

$$\begin{aligned}\langle Q^I_{\boldsymbol{k}_1}Q^J_{\boldsymbol{k}_2}Q^K_{\boldsymbol{k}_3}\rangle_{(2)} &= -i\frac{H^3}{8k_{123}^3}\int_{-\infty^+}^{\tau_*} d\tau'\, \dot{\phi}^I\delta^{JK}(k_1k_2)^2(1-ik_2\tau')\, e^{+iK\tau'}\\ &= -\frac{H^3}{8k_{123}^3}\sum_{\text{perms}}\dot{\phi}^I\delta^{JK}\left[\frac{(k_2k_3)^2}{K}+\frac{k_2^2k_3^3}{K^2}\right].\end{aligned} \tag{12.6.10}$$

12.6.3 Correction Terms

Finally, we consider the contribution to the bispectrum arising from the field redefinition. According to the prescription of Eq. 12.5.17, the effect of the field redefinition can be included by adding the terms:

$$-\langle F^I Q^J Q^K\rangle + 2 \text{ perms.}\,, \tag{12.6.11}$$

to the bispectrum. We split the function $F^I(Q)$ (Eq. 12.5.11) according to the two terms as follows:

$$F^I(Q) = F^I_{(a)} + F^I_{(b)}, \tag{12.6.12}$$

where:

$$F^I_{(a)}(Q) \equiv \frac{1}{4H}\dot{\phi}^I\partial^{-2}(Q^J\partial^2 Q_J), \qquad F^I_{(b)}(Q) \equiv -\frac{1}{8H}\dot{\phi}^I Q_J Q^J. \tag{12.6.13}$$

In real space, the three-point correlator involving $F^I_{(a)}$ can be written in terms of two-point functions is given by:

$$\langle F^I_{(a)}Q^JQ^K\rangle = \frac{1}{4H}\dot{\phi}^I\delta^{MN}\,\partial^{-2}_{\boldsymbol{x}_2}\Big[\langle Q^M(\boldsymbol{x}_1)Q^J(\boldsymbol{x}_2)\rangle\,\partial^2_{\boldsymbol{x}_1}\langle Q^N(\boldsymbol{x}_1)Q^K(\boldsymbol{x}_3)\rangle\Big]. \tag{12.6.14}$$

Going to momentum space now amount to replacing $\partial_{\boldsymbol{x}_2}^{-2}$ and $\partial_{\boldsymbol{x}_1}^{2}$ by $-k_2^2$ and $-k_1^{-2}$, respectively. Using the propagator (Eq. 12.4.13) in the coincidence limit at $\tau_* \to 0$ outside the horizon (so that the momenta vanish), we obtain:

$$\langle F^I_{(a)} Q^J Q^K \rangle = \frac{H^3}{2} \frac{1}{8k_{123}^3} \sum_{\text{perms}} \dot{\phi}^I \delta^{JK} (k_1 k_2^2). \tag{12.6.15}$$

Similarly, the contribution to the second term involving $F^I_{(b)}$ can be written as:

$$\langle F^I_{(b)} Q^J Q^K \rangle = -\frac{H^3}{2} \frac{1}{8k_{123}^3} \sum_{\text{perms}} \dot{\phi}^I \delta^{JK} \frac{k_1^3}{2}. \tag{12.6.16}$$

12.6.4 Leading Order Bispectrum Multi-Field Inflation

Combining the contributions given above, i.e. we sum over Eqs. 12.6.8, 12.6.10, 12.6.15 and 12.6.16, we obtain the leading order bispectrum:

$$\langle Q^I_{\boldsymbol{k}_1} Q^J_{\boldsymbol{k}_2} Q^K_{\boldsymbol{k}_3} \rangle = (2\pi)^3 \delta^{(3)}(\boldsymbol{K}) \, \frac{H^3}{2} \frac{1}{8k_{123}^3} \, \mathcal{B}^{IJK}. \tag{12.6.17}$$

where the momentum-dependence is governed in $\mathcal{B}^{IJK}(k_1, k_2, k_3)$, defined as:

$$\begin{aligned} \mathcal{B}^{IJK} &= \sum_{\text{perms}} \dot{\phi}^I \delta^{JK} \left[-3\frac{(k_2 k_3)^2}{K} - \frac{(k_2 k_3)^2}{K^2}(k_1 + 2k_3) + \frac{1}{2}k_1^3 - k_1 k_2^2 \right] \\ &= \sqrt{2} H \sum_{\text{perms}} \varepsilon^I \delta^{JK} \left[-3\frac{(k_2 k_3)^2}{K} - \frac{(k_2 k_3)^2}{K^2}(k_1 + 2k_3) + \frac{1}{2}k_1^3 - k_1 k_2^2 \right], \end{aligned} \tag{12.6.18}$$

where the Planck mass is still set to unity $M_{\text{pl}} \equiv 1$.

Two comments on the above result are in order. First of all, the above three-point function for Q^I does not represent the primordial non-gaussianity. Instead, the three-point function governs the cubic interactions of Q-modes around the moment of horizon crossing. In order to calculate the non-Gaussianity using the above result (Eq. 12.6.17), the formalism briefly described in the section 12.1 and [13] should be used to compute the bispectrum of the comoving curvature perturbation.

Secondly, the temporal integrals over the de Sitter propagators that lead to the above result are insensitive to the behavior of the field modes both deep inside the horizon and well after horizon-crossing. The dominant contribution arises around the epoch of horizon-crossing. In other words, Eq. 12.6.17 is the bispectrum of Q^I a short time after horizon crossing. The assumption that we have implicitly made is that the wavenumbers of the three momentum modes are of the same order. If they are not, the modes differ significantly at the moment they leave the horizon and the approach above, in which we evaluated all modes at the same moment in time ($\tau_* \to 0$), would cease to be valid.

This result is in exact agreement with Eq. 68 of Seery and Lidsey [76]. To show the equivalence, define $\mathcal{A}^{IJK}$ in relation to $\mathcal{B}^{IJK}$ as follows:

$$\mathcal{A}^{IJK} \equiv \frac{1}{4H} \mathcal{B}^{IJK} \equiv \frac{1}{4H} \dot{\phi}^I \delta^{JK} \sum_{\text{perms}} \mathcal{M}, \tag{12.6.19}$$

where we have introduced and defined a third form of the momentum-dependence which is independent of the field labels I, J and K. The appearence of the various different quantities governing the momentum-dependence of the bispectrum is a consequence of the different notations in the literature. In particular, the function $\mathcal{B}^{IJK}$ is defined in this work, $\mathcal{A}^{IJK}$ is used in [64, 76] and the index-free variable $\mathcal{M}$ is defined in [85].

Now we find that Eq. 12.6.17 can be rewritten, using the fact the spectrum of a massless scalar field in de Sitter space can be written conform Eq. 12.4.16, in the following way (see Eq. 12.1.16):

$$\langle Q^I_{\boldsymbol{k}_1} Q^J_{\boldsymbol{k}_2} Q^K_{\boldsymbol{k}_3} \rangle = (2\pi)^3 \delta^{(3)}(\boldsymbol{K}) \, \frac{4\pi^4}{k_{123}^3} \mathcal{P}_Q^2 \, \mathcal{A}^{IJK}. \tag{12.6.20}$$

Explicitly, and in exact agreement with Eq. 69 in [76], the momentum function $\mathcal{A}^{IJK}$ can be written as:

$$\begin{aligned} \mathcal{A}^{IJK} &= \sum_{\text{perms}} \frac{\dot{\phi}^I}{4H} \delta^{JK} \left[-3 \frac{(k_2 k_3)^2}{K} - \frac{(k_2 k_3)^2}{K^2} (k_1 + 2k_3) + \frac{1}{2} k_1^3 - k_1 k_2^2 \right] \\ &= \frac{1}{2\sqrt{2}} \sum_{\text{perms}} \varepsilon^I \delta^{JK} \left[-3 \frac{(k_2 k_3)^2}{K} - \frac{(k_2 k_3)^2}{K^2} (k_1 + 2k_3) + \frac{1}{2} k_1^3 - k_1 k_2^2 \right]. \end{aligned} \tag{12.6.21}$$

12.7 Squeezed Limit of Multi-Field Inflation

For single-field inflation, we have shown in sections 7.5 and 10.8 that the non-linearity parameter in the squeezed limit ($k_1 \ll k_2 \simeq k_3$) of the local template is suppressed by the spectral index (Eq. 7.5.2):

$$\lim_{k_1 \to 0} f_{\text{NL}}^{\text{local}} = \frac{5}{12}(1 - n_{\text{s}}), \tag{12.7.1}$$

thereby enforcing the bispectrum for single field inflation to be small in this limit. In this section, we will show that in the same limit (squeezed limit of the local template), the bispectrum for multi-field inflation instead gets very large. Therefore, the squeezed limit of the local template provides a clear distinction between the single- and multi-field scenario. Observing the non-linearity parameter to be of order unity would be in accordance with the multi-field scenario, and the single-field scenario would be ruled out.

To derive this statement, we start by considering the different leading ordercontributions to the non-linearity parameter f_{NL} as parametrized in Eq. 12.1.22. In particular, we will consider the contribution denoted as $f_{\text{NL}}^{(1)}$ and argue that it is supressed by the tensor-to-scalar ratio r defined in Eq. 3.4.6. Starting from its definition, $f_{\text{NL}}^{(1)}$ reads:

$$-\frac{6}{5} f_{\text{NL}}^{(1)} \equiv \frac{\partial_I N \partial_J N \partial_K N}{(\partial_L N \partial^L N)^2} \frac{\mathcal{A}^{IJK}}{\sum_i k_i^3}. \tag{12.7.2}$$

Evaluating the numerator using $\partial_I N \dot{\phi}^I = -H$ gives:

$$\partial_I N \partial_J N \partial_K N \mathcal{A}^{IJK} = \partial_I N \partial_J N \partial_K N \frac{\dot{\phi}^I \delta^J K}{4H} \sum_{\text{perms}} \mathcal{M} = -\frac{1}{4} (\partial_J N \partial^J N) \sum_{\text{perms}} \mathcal{M}. \tag{12.7.3}$$

Therefore, we find that we write $f_{\text{NL}}^{(1)}$ as follows:

$$-\frac{6}{5} f_{\text{NL}}^{(1)} = -\frac{\mathcal{P}_Q}{4\mathcal{P}_{\mathcal{R}}} \frac{\sum \mathcal{M}}{\sum_i k_i^3}, \tag{12.7.4}$$

where we have used Eq. 12.1.15 to replace the denominator $\partial_L N \partial^L N$ by the ratio of power spectra.

The trick is now as follows: we know that the power spectrum of Q equals $(H/2\pi)^2$, which has a similar form to the gravitational wave power spectrum $\mathcal{P}_E$ (Eq. 5.6.28) as derived in section 5.6.[3] In the limit of scale-invariant power spectra, the relationship between the gravitational wave power spectrum and the spectrum of Q is given by:

$$\mathcal{P}_E = \frac{1}{8}\left(\frac{H}{2\pi}\right) = \frac{\mathcal{P}_Q}{8}. \tag{12.7.5}$$

Using this relation, we can write the non-linearity parameter $f_{\rm NL}^{(1)}$ in terms of the tensor-to-scalar ratio r, yielding:

$$-\frac{6}{5} f_{\rm NL}^{(1)} = -\frac{r}{32}\frac{\sum \mathcal{M}}{\sum_i k_i^3}. \tag{12.7.6}$$

Since r is constrained by observations to be small ($r < 0.11$ [3]), we find that $f_{\rm NL}^{(1)}$ is small in case the momentum-dependent part is of order unity for all triangular shapes.

We will show that this is the case by first explicitly performing the permutation sum of $\mathcal{M}$, which gives the result:

$$\sum_{\rm perms} \mathcal{M} = -2\left[\frac{1}{2}\sum_{i\neq j} k_i k_j^2 + \frac{4}{K}\sum_{i>j}(k_i k_j)^2 - \frac{1}{2}\sum_i k_i^3\right]. \tag{12.7.7}$$

Using this result, we can recast the ratio $\sum \mathcal{M}/\sum_i k_i^3$ in the following form:

$$\frac{\sum \mathcal{M}}{\sum_i k_i^3} = -2(f+1). \tag{12.7.8}$$

The reason to rewrite the ratio in this form is that the bounds on the momentum-dependent function f are easily found. Explicity, the function f reads:

$$f(k_1,k_2,k_3) \equiv \frac{1}{2\sum_i k_i^3}\left[\sum_{i\neq j} k_i k_j^2 + \frac{8}{K}\sum_{i>j}(k_i k_j)^2\right] - \frac{3}{2}. \tag{12.7.9}$$

The lower bound on f is found by considering the squeezed limit $k_1 \to 0$ and $k \equiv k_2 \simeq k_3$, yielding:

$$\lim_{k_1\to 0} f(k_1,k,k) = 0. \tag{12.7.10}$$

The upper bound is obtained in the equilateral triangular configuration, where $k_{1,2,3} \equiv k$ and the limit on f becomes:

$$\lim_{k_{1,2,3}\to k} f(k_1,k_2,k_3) = \frac{5}{6}. \tag{12.7.11}$$

Combining those results, the function f is constrained to attain values in the range $0 \leq f \leq 5/6$ [64]. Therefore, the ratio in Eq. 12.7.8 will always be of order unity $\mathcal{O}(1)$, independent

[3] The gravitational wave power spectrum (Eq. 5.6.28) was derived in the context of single-field inflation. However, note that the derivation is independent of the number of fields and therefore also apply to multi-field inflation.

of the triangular configuration. Hence the non-linearity parameter $f_{\rm NL}^{(1)}$ is constrained to be of the order $r/16$, yielding:

$$-\frac{6}{5}f_{\rm NL}^{(1)} = \mathcal{O}(r/16) < \mathcal{O}(10^{-3}), \tag{12.7.12}$$

on account of the observational bound $r < 0.11$ [3].

In conclusion, the momentum-dependent contribution to the non-linearity parameter $f_{\rm NL}$ much less than unity and can be neglected in case the momentum-independent part $f_{\rm NL}^{(2)}$ is large than unity. Moreover, the ideal CMB experiment is expected to only reach an accuracy of $f_{\rm NL}$ of order unity and therefore observations will be insensitive to the momentum-dependent part. Therefore, we will set the non-linearity parameter to be equal to the momentum-independent contribution from now:

$$f_{\rm NL} \simeq f_{\rm NL}^{(2)} = -\frac{5}{6}\frac{\partial_I N \partial_J N \partial^I \partial^J N}{(\partial^L N \partial_L N)^2}. \tag{12.7.13}$$

The bispectrum of the comoving curvature perturbation $B_{\mathcal{R}}$ is related to the non-linearity parameter as follows:

$$\begin{aligned} B_{\mathcal{R}}(k_1,k_2,k_3) &= \frac{4\pi^4}{k_{123}^3}\mathcal{P}_{\mathcal{R}}^2\left(\mathcal{A}_{\mathcal{R}}^{(1)} + \mathcal{A}_{\mathcal{R}}^{(2)}\right) = -\frac{4\pi^4}{k_{123}^3}\mathcal{P}_{\mathcal{R}}^2 \times \frac{6}{5}(f_{\rm NL}^{(1)} + f_{\rm NL}^{(2)})\sum_i k_i^3 \\ &\simeq -\frac{6}{5}f_{\rm NL}^{(2)}(4\pi^4\mathcal{P}_{\mathcal{R}}^2)\left[\frac{1}{(k_1k_2)^3} + \frac{1}{(k_1k_3)^3} + \frac{1}{(k_2k_3)^3}\right]. \end{aligned} \tag{12.7.14}$$

In case the non-linearity parameter $f_{\rm NL} \simeq f_{\rm NL}^{(2)}$ is of order unity or larger, the bispectrum can become large in the squeezed limit ($k_1 \to 0$), which is in sharp contrast to the single-field scenario, where the local non-linearity parameter is always suppressed by the spectral index. In the next chapter, we will show that it is indeed possible to achieve $f_{\rm NL}^{(2)} = \mathcal{O}(1)$ already in the case of two-field and hence a significant amount of non-gaussianity is contained in the bispectrum for multi-field inflation.

12.8 Large Non-Gaussianities in Two-Field Inflation

In this section, we will make the statement that the large non-gaussianity ($f_{\rm NL} \sim \mathcal{O}(1)$) can be generated in multi-field models more precise by examining a generic two-field model with a sum-separable potential [85]. In particular, we will derive the dynamical conditions required for two-field models to induce a high level of non-gaussianity.

Consider two fields, denoted as φ and χ, such that the field space is spanned by the vector $\phi^I = (\varphi, \chi)$. Assuming the two field have canonical kinetic terms, the Lagrangian density $\mathcal{L}(\varphi,\chi)$ for the fields reads:

$$\mathcal{L}(\varphi,\chi) = -\frac{1}{2}g^{\mu\nu}\Big(\partial_\mu\varphi\partial_\nu\varphi + \partial_\mu\chi\partial_\nu\chi\Big) - W(\varphi,\chi), \tag{12.8.1}$$

where the potential of the two-field system is given by $W(\varphi,\chi)$. For simplicity, we take the potential to be sum-separable, such that:

$$W(\varphi,\chi) = U(\varphi) + V(\chi), \tag{12.8.2}$$

for the case of a product-separable potential, see [29]. Because of the sum-separable potential, the Klein-Gordon equations for the fields can be written as:

$$\ddot{\varphi} + 3H\dot{\varphi} + U_\varphi = 0, \qquad \ddot{\chi} + 3H\dot{\chi} + V_\chi = 0. \tag{12.8.3}$$

In addition, the Friedmann equations can be written as follows:

$$H^2 = \frac{1}{3M_{\rm pl}^2}\left[\frac{1}{2}\dot{\varphi}^2 + \frac{1}{2}\dot{\chi}^2 + W(\varphi,\chi)\right], \qquad \dot{H} = -\frac{1}{2M_{\rm pl}^2}(\dot{\varphi}^2 + \dot{\chi}^2). \tag{12.8.4}$$

To conveniently describe the dynamics of the fields, we will assume the slow-roll approximation is valid throughout the inflationary epoch for both fields separately. In that case, the Klein-Gordon and Friedmann equations for the two fields become:

$$3H\dot{\varphi} = -U_\varphi, \qquad 3H\dot{\chi} = -V_\chi, \qquad 3M_{\rm pl}^2 H^2 = W(\varphi,\chi). \tag{12.8.5}$$

We will characterize the slow-roll regime by means of the potential slow-roll perameters, instead of the more general Hubble parameters ε^{IJ} and $\tilde{\eta}^{IJ}$ defined earlier. Since we are dealing with a sum-separable potential, the mixed slow-roll parameters vanish and we find:

$$\begin{aligned} &\varepsilon_\varphi^W \equiv \frac{M_{\rm pl}^2}{2}\left(\frac{U_\varphi}{W}\right)^2, \qquad \varepsilon_\chi^W \equiv \frac{M_{\rm pl}^2}{2}\left(\frac{V_\chi}{W}\right)^2, \qquad \varepsilon^W = \varepsilon_\varphi^W + \varepsilon_\chi^W, \\ &\eta_{\varphi\varphi}^W \equiv M_{\rm pl}^2\frac{U_{\varphi\varphi}}{W}, \qquad \eta_{\chi\chi}^W \equiv M_{\rm pl}^2\frac{U_{\chi\chi}}{W}. \end{aligned} \tag{12.8.6}$$

In order to connect to the notations ε^{IJ} and $\tilde{\eta}^{IJ}$, notice that we have made the notational redefinitions $\varepsilon_\varphi \equiv \varepsilon^{\varphi\varphi}$ and $\eta_{\varphi\varphi}^W \equiv \tilde{\eta}_{\varphi\varphi}$. From now on, we set the Planck mass to unity and will often omit the label W and the slow-roll parameters with labels as sub- or superscripts are taken to be equivalent in terms of notation. Finally, we assume that throughout inflation $U_\varphi, V_\chi \geq 0$, so that we can elimate first field derivatives of the potential by the square root of the first slow-roll parameter via the relation:

$$U_\varphi = W\sqrt{2\varepsilon_\varphi}, \tag{12.8.7}$$

and a similar relationship for the field χ.

The number of e-foldings $N(t_c, t_*)$, from an initial flat slice at t_* to a slice of constant energy density at t_c, can be written in field space using the first Friedmann equation as:

$$N(t_c, t_*) \equiv \int_{t_*}^{t_c} H\, dt = -\frac{1}{M_{\rm pl}^2}\left[\int_{\varphi_*}^{\varphi_c}\frac{U}{U_\varphi}d\varphi + \int_{\chi_*}^{\chi_c}\frac{V}{V_\chi}d\chi\right]. \tag{12.8.8}$$

In terms of the Fourier modes k, we take the time t_* as the time at which the considered k-mode crosses the horizon (i.e. $k = aH$ at $t = t_*$).

For single-field models, the background dynamics of the inflaton can be described by a unique classical trajectory in phase space. However, in the multi-field scenario, an infinite number of classical trajectories are possible. In order to label those different trajectories, we introduce the following dimensional integral of motion:

$$C \equiv -M_{\rm pl}^2\int_{\varphi_*}^{\varphi_c}\frac{d\varphi}{U_\varphi} + M_{\rm pl}^2\int_{\chi_*}^{\chi_c}\frac{d\chi}{V_\chi}, \tag{12.8.9}$$

which yields a different value for each classical trajectory. Notice that the integral of motion C parametrizes the motion of the classical trajectory, whereas N characterizes the evolution along a given trajectory.

12.8.1 Two-Field Non-Linearity Parameter

In order to compute the non-linearity parameter, we will have to expand dN to second order in field perturbations. Therefore, we compute the derivatives N_φ and N_χ, by differentiating Eq. 12.8.8 with respect to φ_* and χ_*. Notice that the integrals in Eq. 12.8.8 depend on both fields, since they are coupled via the integral of motion $C(\varphi_*, \chi_*)$. The result can be expressed as:

$$\begin{aligned} dN = &\left[\frac{U}{U_\varphi}\bigg|_{t_*} - \frac{\partial\chi_c}{\partial\varphi_*} \times \frac{V}{V_\chi}\bigg|_{t_c} - \frac{\partial\varphi_c}{\partial\varphi_*} \times \frac{U}{U-\varphi}\bigg|_{t_c}\right] d\varphi_* \\ &+ \left[\frac{V}{V_\chi}\bigg|_{t_*} - \frac{\partial\varphi_c}{\partial\chi_*} \times \frac{U}{U_\varphi}\bigg|_{t_c} - \frac{\partial\chi_c}{\partial\chi_*} \times \frac{V}{V_\chi}\bigg|_{t_c}\right] d\chi_*, \end{aligned} \tag{12.8.10}$$

recall that we have set the Planck mass to unity ($M_{\rm pl} \equiv 1$). The first terms on both lines are trivial, while the second and third terms in both line arise form the coupling between the two fields by means of C.

To compute the derivatives between the square brackets, we use the relations:

$$\begin{aligned} d\varphi_c &= \frac{d\varphi_c}{dC} dC = \frac{d\varphi_c}{dC}\left(\frac{\partial C}{\partial\varphi_*} d\varphi_* + \frac{\partial C}{\partial\chi_*} d\chi_*\right) \\ d\chi_c &= \frac{d\chi_c}{dC} dC \frac{d\chi_c}{dC}\left(\frac{\partial C}{\partial\varphi_*} d\varphi_* + \frac{\partial C}{\partial\chi_*} d\chi_*\right). \end{aligned} \tag{12.8.11}$$

From the definition of the integral of motion, the partial derivatives of C can be calculated in trivial way, yielding:

$$\frac{\partial C}{\partial\varphi_*} = -\frac{1}{U_\varphi}\bigg|_{t_*}, \qquad \frac{\partial C}{\partial\chi_*} = \frac{1}{V_\chi}\bigg|_{t_*}. \tag{12.8.12}$$

In order to proceed, we invoke the fact that the slice at t_c is taken to be of uniform energy density, such that in the slow-roll regime the total energy $\rho \simeq W$ equals a constant ω. This can be expressed by means of the following condition:

$$\rho(t_c) \simeq W(t_c) = U(t_c) + V(t_c) \equiv \omega. \tag{12.8.13}$$

Differentiating this result with respect to C and using the chain rule, we find:

$$\frac{d\varphi_c}{dC} U_\varphi\bigg|_{t_c} + \frac{d\chi_c}{dC} V_\chi\bigg|_{t_c} = 0. \tag{12.8.14}$$

In addition, differentiating the definition C with respect to C, we obtain two equations for two unknowns ($d\varphi_c/dC$ and $d\chi_c/dC$). Solving those equations yields:

$$\begin{aligned} \frac{d\varphi_c}{dC} &= -\left[U_\varphi\left(\frac{1}{U_\varphi^2} + \frac{1}{V_\chi^2}\right)\bigg|_{t_c}\right]^{-1} \equiv \mathcal{G}_\varphi^{-1}, \\ \frac{d\chi}{dC} &= \left[V_\chi\left(\frac{1}{U_\varphi^2} + \frac{1}{V_\chi^2}\right)\bigg|_{t_c}\right]^{-1} \equiv \mathcal{G}_\chi^{-1}, \end{aligned} \tag{12.8.15}$$

it should be emphasized that the quantities $\mathcal{G}_{\varphi,\chi}$ are evaluated at time t_c.

Upon substituting the above relations in the differential of $d\varphi_c$, we obtain the following result:

$$d\varphi_c = \frac{1}{\mathcal{G}_\varphi}\left(\frac{1}{U_\varphi}\bigg|_{t_*} d\varphi_* - \frac{1}{V_\chi}\bigg|_{t_*} d\chi_*\right), \tag{12.8.16}$$

and a similar expression for the differential $d\chi_c$. Hence, we can identify that:

$$\frac{\partial\varphi_*}{\partial\varphi_c} = \frac{1}{\mathcal{G}_\varphi} \times \frac{1}{U_\varphi}\bigg|_{t_*} = \frac{W_c}{W_*}\frac{\varepsilon_c^\chi}{\varepsilon_c^\varphi + \varepsilon_c^\chi}\left(\frac{\varepsilon_c^\varphi}{\varepsilon_*^\varphi}\right)^{1/2}. \tag{12.8.17}$$

In the second equality we used Eq. 12.8.7 to eliminate the field derivatives of the potentials in favor of the potential slow-roll parameters. Furthermore, we use subscripts $*$ and c to denote evaluation of the considered quantity at t_* and t_c, respectively. Similarly, the other differentials can be found to be:

$$\frac{\partial\varphi_c}{\partial\varphi_*} = -\frac{W_c}{W_*}\frac{\varepsilon_c^\chi}{\varepsilon_c}\left(\frac{\varepsilon_c^\varphi}{\varepsilon_*^\chi}\right)^{1/2}, \quad \frac{\partial\chi_c}{\partial\varphi_*} = -\frac{W_c}{W_*}\frac{\varepsilon_c^\varphi}{\varepsilon_c}\left(\frac{\varepsilon_c^\chi}{\varepsilon_*^\varphi}\right)^{1/2}, \quad \frac{\partial\chi_c}{\partial\chi_*} = \frac{W_c}{W_*}\frac{\varepsilon_c^\varphi}{\varepsilon_c}\left(\frac{\varepsilon_c^\chi}{\varepsilon_*^\varphi}\right)^{1/2}. \tag{12.8.18}$$

Now we can compute the field derivatives of the number of e-foldings N, to be used to compute $f_{\rm NL}$ by means of the δN formalism. Substituting the above results into Eq. 12.8.10 and comparing to the differential relation:

$$dN = \frac{\partial N}{\partial\varphi_*}d\varphi_* + \frac{\partial N}{\partial\chi_*}d\chi_*, \tag{12.8.19}$$

we find the following expressions for the field derivatives of N:

$$\frac{\partial N}{\partial\varphi_*} = \frac{1}{\sqrt{2\varepsilon_*^\varphi}}\frac{U_* + Z_c}{W_*}, \qquad \frac{\partial N}{\partial\chi_*} = \frac{1}{\sqrt{2\varepsilon_*^\chi}}\frac{V_* - Z_c}{W_*}, \tag{12.8.20}$$

where we have defined the function Z_c as follows:

$$Z_c \equiv \frac{V_c\varepsilon_c^\varphi - U_c\varepsilon_c^\chi}{W_*}. \tag{12.8.21}$$

We can differentiate again with respect to the fields to obtain the second field derivatives of N. The results are:

$$\begin{aligned}
\frac{\partial^2 N}{\partial\varphi_*^2} &= 1 - \frac{\eta_*^\varphi}{2\varepsilon_*^\varphi}\frac{U_* + Z_c}{W_*} + \frac{W_*}{\sqrt{2\varepsilon_*^\varphi}}\frac{\partial Z_c}{\partial\varphi_*},\\
\frac{\partial^2 N}{\partial\chi_*^2} &= 1 - \frac{\eta_*^\chi}{2\varepsilon_*^\chi}\frac{V_* - Z_c}{W_*} - \frac{1}{W_*\sqrt{2\varepsilon_*^\chi}}\frac{\partial Z_c}{\partial\chi_*},\\
\frac{\partial^2 N}{\partial\varphi_*\chi_*} &= \frac{1}{W_*\sqrt{2\varepsilon\varphi*}}\frac{\partial Z_c}{\partial\chi_*} = -\frac{1}{W_*\sqrt{2\varepsilon_*^\chi}}\frac{\partial Z_c}{\partial\varphi_*}.
\end{aligned} \tag{12.8.22}$$

The only remaining quantity to compute is the derivative of Z_c with respect to the field values at time t_*. The results are compactly given by:

$$\sqrt{\varepsilon_*^\varphi}\frac{\partial Z_c}{\partial\varphi_*} = -\sqrt{\varepsilon_*^\chi}\frac{\partial Z_c}{\partial\chi_*} \equiv \sqrt{2}W_*\mathcal{A}, \tag{12.8.23}$$

where we have defined:

$$\mathcal{A} \equiv -\frac{W_c^2}{W_*^2}\frac{\varepsilon_c^\varphi\varepsilon_c^\chi}{\varepsilon_c}\left(1 - \frac{\eta_c^{ss}}{\varepsilon_c}\right), \tag{12.8.24}$$

where $\eta^{ss} \equiv (\varepsilon^\chi\eta^\varphi + \varepsilon^\varphi\eta^\chi)/\varepsilon$ describes the effective mass of isocurvature perturbations orthogonal to the classical trajectory of the fields in phase space.

Finally, using the above results for the first and second field derivatives of the number of e-foldings, we can explicitly derive the non-linearity parameter $f_{\rm NL}$. Since the momentum-dependent contribution $f_{\rm NL}^{(1)}$ is constrained to be much less than unity and therefore lies out of experimental reach, we solely compute the momentum-independent contribution denoted by $f_{\rm NL}^{(2)}$. From its definition, we find:

$$-\frac{6}{5}f_{\rm NL}^{(2)} \equiv 2\frac{(\partial_\varphi N)^2\partial^\varphi\partial^\varphi N + (\partial_\chi N)^2\partial^\chi\partial^\chi N}{((\partial_\varphi N)^2 + (\partial_\chi N)^2)^2}. \tag{12.8.25}$$

Introducing the variables $u \equiv (U_* + Z_c)/W_*$ and $v \equiv (V_* - Z_c)/W_*$ (subject to the combined constraint $u + v = 1$), we find that, by invoking the results above, $f_{\rm NL}^{(2)}$ can be written as:

$$\boxed{-\frac{6}{5}f_{\rm NL}^{(2)} = 2\left[\frac{u^2}{\varepsilon_*^\varphi} + \frac{v^2}{\varepsilon_*^\chi}\right]^{-2} \times \mathcal{C},} \tag{12.8.26}$$

where we defined the function $\mathcal{C}$ to be:

$$\mathcal{C} \equiv \frac{u^2}{\varepsilon_*^\varphi}\left(1 - \frac{\eta_*^{\varphi\varphi}}{2\varepsilon_*^\varphi}u\right) + \frac{v^2}{\varepsilon_*^\chi}\left(1 - \frac{\eta_*^{\chi\chi}}{2\varepsilon_*^\chi}v\right) + \left(\frac{u}{\varepsilon_*^\varphi} - \frac{v}{\varepsilon_*^\chi}\right)\mathcal{A}. \tag{12.8.27}$$

12.8.2 Conditions for Large Non-Gaussianity

With the explicit expression for the non-linearity parameter at hand, we can now examine the conditions required for large non-gaussianity, i.e. $f_{\rm NL}^{(2)} \sim \mathcal{O}(1)$. To do so, we will introduce the field space angle θ as follows:

$$\cos\theta = \frac{\dot\varphi}{\sqrt{\dot\varphi^2 + \dot\chi^2}}, \qquad \sin\theta = \frac{\dot\chi}{\sqrt{\dot\varphi^2 + \dot\chi^2}}. \tag{12.8.28}$$

In terms of the field space angle, we find that the first slow-roll parameters can be written as:

$$\frac{\varepsilon_\varphi}{\varepsilon} = \cos^2\theta, \qquad \frac{\varepsilon_\chi}{\varepsilon} = \sin^2\theta, \tag{12.8.29}$$

and we can identify the dimensionless parameters u and v with the angle θ_c at time t_c as $u = \cos^2\theta_c$ and $v = \sin^2\theta_c$, respectively [29]. Using these relations, we find that:

$$\eta^{ss} = \eta^{\varphi\varphi}\sin^2\theta + \eta^{\chi\chi}\cos^2\theta. \tag{12.8.30}$$

In terms of the field angle θ, we can re-express the non-linearity parameter (Eq. 12.8.25) by means of auxiliary harmonic functions, yielding:

$$-\frac{6}{5}f_{\rm NL}^{(2)} = \left[2j_s\varepsilon^* - f_s\eta_{\varphi\varphi}^* - g_s\eta_{\chi\chi}^* - 2h_s\frac{W_c^2}{W_*^2}(\varepsilon_c - \eta_c^{ss})\right]. \tag{12.8.31}$$

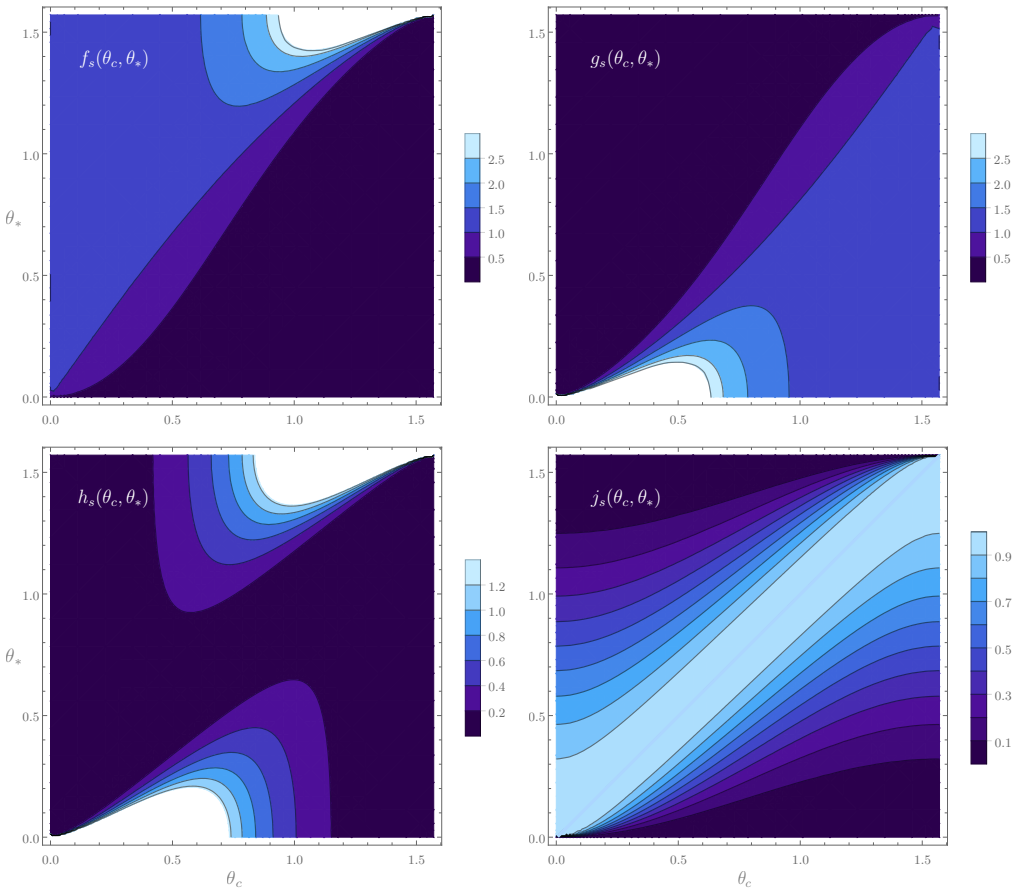


Figure 12.1: Contour plots of the harmonic functions (j_s, g_s, h_s and f_s) in the range $0 \leq \theta_{c,*} \leq \pi/2$. The white areas are understood to be regimes in which the functions become larger than the highest color scale indicated. Those are the regimes in which large NG can potentially be generated. Notice that, contrary to all others, the function j_s is bounded to values smaller than unity over the entire parameter range and hence can never induce large NG.

Since $u + v = 1$, the auxiliary harmonic functions are only a function of two variables, we choose them to be u and θ_*. Explicitly, the functions read:

$$
\begin{aligned}
f_s(\theta_c, \theta_*) &= \frac{u^3 \sin^3 \theta_*}{(u^2 \sin^2 \theta_* + v^2 \cos^2 \theta_*)^2} \\
g_s(\theta_c, \theta_*) &= \frac{v^3 \cos^4 \theta_*}{(u^2 \sin^2 \theta_* + v^2 \cos^2 \theta_*)^2} \\
h_s(\theta_c, \theta_*) &= \sin^2 \theta_c \cos^2 \theta_c \times \frac{(u \sin^2 \theta_* - v \cos^2 \theta_*)^2}{(u^2 \sin^2 \theta_* + v^2 \cos^2 \theta_*)^2} \\
j_s(\theta_c, \theta_*) &= \frac{u^2 \sin^4 \theta_* \cos^2 \theta_* + v^2 \cos^4 \theta_* \sin^2 \theta_*}{(u^2 \sin^2 \theta_* + v^2 \cos^2 \theta_*)^2}.
\end{aligned}
\tag{12.8.32}
$$

In order to get large non-gaussianity, we need to find the parameter regimes (θ_*, θ_c) in which the harmonic prefactors become large. To find those regimes, we plotted the contours

of the harmonic functions in Fig. 12.1. Based on the contour plots, we distinguish two different regimes in which large non-gaussianity can potentially be generated:

▷ *Region A.*—Observe that the harmonic function f_s becomes large in the limit $\theta_{*,c} \to \pi/2$ (and consequently the function g_s becomes small). This corresponds to the limit $\cos^2\theta_{*,c} \to 0$ so that the condition on the first slow-roll parameter becomes:

$$\varepsilon_\chi^* \gg \varepsilon_\varphi^*. \tag{12.8.33}$$

This hierarchy of the slow-roll parameters corresponds to a field configuration in which the field χ rolls much faster down its potential compared to the field φ (still both fields have to roll slowly, for the slow-roll approximation to be valid). In other words, the field χ dominates φ in terms of kinetic energy.

▷ *Region B.*—The opposite limit is obtained when the angles $\theta_{c,*}$ both tend to zero. In that case, g_s becomes very large and f_s tends to zero. The corresponding condition on the slow-roll parameter is given by:

$$\varepsilon_\varphi^* \gg \varepsilon_\chi^*, \tag{12.8.34}$$

such that the field φ dominates χ in terms of kinetic energy.

In conclusion, the prefactors g_s and f_s become large in case one of the fields dominates over the other in terms of the kinetic energy. Furthermore, notice that the function j_s is smaller than unity over the whole examined parameter space and can hence never induce large NG. Therefore, we will neglect this term in further analysis.

Due to the symmetry of both regions, we choose to focus solely on region B here and based on the analysis below similar results for region A could easily be derived as well. In region B, the function g_s becomes large and f_s negligibly small. In the limit $\theta_{c,*} \to 0$, the functions g_s and h_s can be written as:

$$g_s(\theta_*, \theta_c) = \frac{\sin^6\theta_c}{(\sin^2\theta_* + \sin^2\theta_c)^2}, \qquad h_s(\theta_*, \theta_c) = g_s(\theta_*, \theta_c) \times \cos^2\theta_c. \tag{12.8.35}$$

Since the other terms cannot produce large contributions to the non-linearity parameter in this regime, we find that:

$$\begin{aligned} f_{\rm NL}^{(2)} &= \frac{5}{6}\left[g_s\eta_{\chi\chi}^* + 2h_s\frac{W_c^2}{W_*^2}(\varepsilon_c - \eta_c^{ss})\right] \\ &= \frac{5}{6}g_s\left[\eta_{\chi\chi}^* + 2\frac{W_c^2}{W_*^2}\cos^2\theta_c(\varepsilon_c - \eta_c^{ss})\right]. \end{aligned} \tag{12.8.36}$$

Hence, for $f_{\rm NL}^{(2)}$ to be large, we conclude that $g_s(\theta_c, \theta_*)$ must be large than the inverse of the slow-roll suppressed terms between brackets. This can be formalized by the condition:

$$g_s \gtrsim \left(\eta_{\chi\chi}^* + 2\frac{W_c^2}{W_*^2}(\varepsilon_c - \eta_c^{\chi\chi})\right)^{-1}, \tag{12.8.37}$$

where we have used that in the limit of vanishing $\theta_{*,c}$, we have $\cos^2\theta_c \simeq 1$ and $\eta_c^{ss} \simeq \eta_c^{\chi\chi}$. A similar condition could be find for f_s in region A.

Afterword: Summary and Outlook

This thesis aimed to fulfill three main objectives. First of all, we have given a detailed introduction into the mechanism by which inflation dynamically solves the shortcomings of conventional Big Bang theory, such as flatness and horizon problem. Secondly, we have discussed how quantum effects during inflation can generate the primordial seeds for structure formation and the anisotropies in the CMB. Finally, we compared the single- and multi-field inflationary scenarios, focussing in particular on the non-gaussianity signals in the CMB generated by both scenario.

Summary of this Thesis

Below, we will list the most important predictions of the inflationary paradigm, thereby summarizing the most relevant aspects discussed in the thesis.

- ▷ *Solution to Flatness and Horizon Problem.*—Inflation provides a natural solution to the flatness and horizon problems arising in the conventional Big Bang theory. Even if the universe was highly curved shortly after the Big Bang, inflation can still drive the universe to the extremely flat geometrical state we observe today. In addition, inflation explains why seemingly disconnected patches of the CMB are observed to have the same temperature to a very high degree, which violates the notion of causality in the framework of conventional Big Bang theory. Due to the superluminal expansion of space-time during inflation, i.e. before the release of the CMB, patches in the CMB that appear to be causally disconnected have in fact be in causal contact before inflation. As a result, the uniformity of the CMB background is not a surprise, but merely a logical consequence of the inflationary stage.

- ▷ *Violation of the Strong-Energy-Condition.*—As we have shown, a period of inflationary expansion requires the sourcing matter constituent to have an equation of state $w = P/\rho$ smaller than $-1/3$. This so-called violation of the strong-energy-condition is only satisfied by non-trivial forms of matter, the most intuitive and simple candidate being a scalar field, called the inflaton. In addition, we have considered the case of multiple scalar fields sourcing the inflationary era, which is more natural from the perspective of high energy theories, such as string theory.

- ▷ *Primordial Seeds and CMB Anisotropies.*—Due to the Heisenberg uncertainty principle, quantum fluctuations in the inflaton field(s) are necessarily present. During the exponential expansion, those perturbations are rapidly stretched to cosmological scales, became classical and froze in, thereby forming the primordial seeds for the formation of large-scale structure (LSS) and the temperature anisotropies in the CMB.

- ▷ *Adiabicity of Inflationary Perturbations.*—Outside the horizon, primordial perturbations are predicted to be adiabatic in both the single- and multi-field scenarios of inflation. That is, the perturbations are perturbations in the total energy density of the universe, corresponding to scalar perturbations to the FRW metric of the universe. Adiabicity of those perturbations implies that the ratio of number densities of any two particle species in the universe should be the same everywhere throughout the universe, although the absolute number densities may vary over space.

- ▷ *Scale-Invariance.*—The power spectrum of inflationary perturbations is predicted to be nearly scale-invariant, as measured by the spectral index n_{s}, which is predicted to be close to unity. The deviation from exact scale-invariance arises due to the fact that the inflaton field(s) are not necessarily massless and because the Hubble parameter need not be exactly constant during inflation. (Recall that for a massless scalar field in de Sitter space, we have shown that the power spectrum is completely scale-invariant.)

- ▷ *Primordial Gravitational Waves.*—Inflation predicts primordial gravitational waves to be generated. Similar to the power spectrum for scalar perturbations, the power spectrum for gravitational waves is predicted to be nearly scale invariant. Observing those primordial gravitational waves would provide a direct measure of the energy scale during inflation and therefore constitutes one of the most interesting observational probes for the near future.

- ▷ *Gaussianity of Perturbations.*—Inflation predicts the primordial perturbations to be Gaussian to a very high degree. In the single-field scenario, the gaussianity has a clear origin. In order to get a phase of inflationary expansion, the potential should very flat. As a result of this requirement, (self-)interactions and other sources of non-linearity are constrained to be small. Therefore, only the necessary coupling to gravity provides a small source of non-gaussianity. The same reasoning applies to multi-field inflation in case the fields all adhere to the slow-roll condition.

- ▷ *Primordial Non-Gaussianity.*—Although deviations from purely gaussian perturbations are expected to be small for inflationary perturbations, informative imprints are contained in the non-gaussian signal, which can be probed using the CMB. In particular, for single field inflation, the level of non-gaussianity contained in the CMB should be small (in the squeezed limit of the local template), as measured by $f_{\mathrm{NL}}^{\mathrm{local}}$, which should be smaller than $\mathcal{O}(10^{-3})$. This provides a distinct observational probe for inflation: in case the non-linearity parameter is observed to be much larger, i.e. $f_{\mathrm{NL}}^{\mathrm{local}} = \mathcal{O}(1)$, the complete single-field scenario would be ruled out. At the same time, a significant non-linearity parameter would be in accordance with the generic predictions of multi-field inflation. Therefore, future observations have the potential to discriminate between the single- and multi-field scenarios.

Possible Extensions to this Thesis

We will finish this afterword by considering possible directions for future research. Below we give a few suggestions for further research, but it should be noted that the list of possible extensions given below is certainly not complete.

First of all, the current work deals with generic classes of single and multi-field inflation but does not consider well motivated explicit models of inflation. Therefore, as a straightforward

extension to this work, the formalism developed in this work could be applied to specific models of single- and multi-field inflation. Examples of such explicit inflation models could be Dirac-Born-Infeld inflation or assisted inflation.

Secondly, notice that the current work focusses solely on scalar perturbations. The analysis could be extended to include tensor modes as well, i.e. primordial gravitational waves. Additional observational predictions could be derived by considering combined correlation functions of scalars and tensors [64, 78]. A related extension is concerned with the fact that we only compared single- and multi-field inflation. In addition, there is the possibility of quasi single-field inflation, for which we could compute the level of non-gaussianity as an observational probe [35].

Finally, an interesting direction to proceed in is to use the observational window of non-gaussianities to derive imprints of new particles in the CMB, an approach which is referred to as cosmological collider physics [7]. In particular, the inflationary background is well approximated as a quasi de-Sitter background and, as such, allows for the existence of a class of partially massless particles, which have no flat space analog and only exist in a (quasi) de Sitter space-time [19]. Therefore, observing their imprints in higher order cosmological correlation functions (encoding non-gaussianities) via the CMB would provide strong evidence that the early universe went through an inflationary phase of exponential expansion.

In conclusion, the inflationary paradigm is a promising and dynamically intuitive candidate for the physics of the early universe and explains the origin of all structure in the universe, in addition to its uniformity and flatness. However, the microscopic mechanism behind inflation is yet to be revealed. Solving this puzzle will require a continuous interplay and synergy between theoretical advances and precise cosmological observations to be performed in the (near) future.

Part VI

Appendices

Appendix A

Geometry and Kinematics of the Universe

A.1 Spatial Metric of the Universe

Here we will derive the advocated 3-metric γ_{ij} (Eq. A.1.10) of the universe from first principles. On account of spatial homogeneity and isotropy, the (background evolution of the) universe may be represented by a sequence of constant time hypersurfaces $\mathcal{M}_t$. A hypersurface is a sub-manifold of dimension $n-1$ embedded in a manifold of dimension n. The hypersurfaces of constant time $\mathcal{M}_t$ may be regarded as spatial slices, each of which is homogeneous and isotropic by the CP. These spatial hypersurfaces are maximally symmetric 3-spaces, which, again by homogeneity and isotropy, are characterized by a constant 3-curvature K. Based on the constant 3-curvature, there are three options for the 3-spaces: negative curvature $K=-1$, zero curvature $K=0$ and positive curvature $K=+1$. Figure 1.2 visualizes the time-ordered spatial slices for the three options.

Below, the three possible symmetric 3-spaces will be described. The 3-spaces are characterized by the differential line segment $d\ell$ and, for negatively and positively curved space, the embedding condition.

- ▷ *Flat space.*—The line element $d\ell$ of flat ($K=0$) three-dimensional Euclidean space E^3 is given by:
$$d\ell^2 = d\boldsymbol{x}^2 = \delta_{ij}dx^i dx^j, \tag{A.1.1}$$
where δ_{ij} is the Kronecker delta function: 0 for $i \neq j$ and 1 for $i=j$.

- ▷ *Positively curved space.*—A 3-space with constant positive curvature ($K>0$) can be represented as a 3-sphere S^3 embedded in four-dimensional Euclidean space E^4. The differential line segment $d\ell$ and embedding condition are:
$$d\ell^2 = d\boldsymbol{x}^2 + dw^2, \qquad \boldsymbol{x}^2 + w^2 = a^2, \tag{A.1.2}$$
respectively. The radius of the 3-sphere is given by a.

- ▷ *Negatively curved space.*—A 3-space with constant negative curvature ($K<0$) may be described by a 3-hyperboloid H^3 embedded in a four-dimensional Lorentzian space $\mathbb{R}^{1,3}$. The corresponding differential line element plus embedding condition read:
$$d\ell^2 = d\boldsymbol{x}^2 - dw^2, \qquad \boldsymbol{x}^2 - w^2 = -a^2, \tag{A.1.3}$$
where a^2 is an arbitrary positive constant.

For the last two cases, the coordinates will be rescaled as $x \to ax$ and $w \to aw$, in order for the curvature parameter K to assume the values $+1$ and -1 in the final result for positive and negative curvature, respectively. Then, the differential line elements and embedding conditions for the spherical and hyperbolic 3-spaces become:

$$d\ell^2 = a^2 \left(d\boldsymbol{x}^2 \pm dw^2\right), \qquad \boldsymbol{x}^2 \pm w^2 = \pm 1, \tag{A.1.4}$$

where the plus and minus variants refer to the spherical and hyperbolic 3-space, respectively. For flat Euclidean space the line element changes to:

$$d\ell^2 = a^2 d\boldsymbol{x}^2. \tag{A.1.5}$$

Subject to the introduced coordinate rescaling, the coordinates $\boldsymbol{x}$ and w become dimensionless and a carries the dimension of length. Computing the differential of the embedding condition gives:

$$\boldsymbol{x}^2 \pm w^2 = \pm 1 \quad \longrightarrow \quad w\, dw = \mp \boldsymbol{x} \cdot d\boldsymbol{x}. \tag{A.1.6}$$

From the above differential, dw^2 can be found as function of $\boldsymbol{x}$ and $d\boldsymbol{x}$ as follows:

$$dw^2 = \left(\frac{\boldsymbol{x} \cdot d\boldsymbol{x}}{\mp w}\right)^2 = \frac{(\boldsymbol{x} \cdot d\boldsymbol{x})^2}{w^2} = \frac{(\boldsymbol{x} \cdot d\boldsymbol{x})^2}{1 \mp \boldsymbol{x}^2}. \tag{A.1.7}$$

In the last equality, the embedding condition is used to obtain $w^2 = 1 \mp \boldsymbol{x}^2$. Substitution of dw^2 in the differential line element $d\ell^2$ as given by Eq. A.1.4 yields:

$$d\ell^2 = a^2 \left[d\boldsymbol{x}^2 \pm \frac{(\boldsymbol{x} \cdot d\boldsymbol{x})^2}{1 \mp \boldsymbol{x}^2}\right]. \tag{A.1.8}$$

The rescaled line element for the hyperbolic and spherical cases can be unified with the rescaled line element for flat space (Eq. A.1.5) by writing:

$$d\ell^2 = a^2 \left[d\boldsymbol{x}^2 + K\frac{(\boldsymbol{x} \cdot d\boldsymbol{x})^2}{1 - K\boldsymbol{x}^2}\right] \equiv a^2 \gamma_{ij}(\boldsymbol{x})\, dx^i dx^j. \tag{A.1.9}$$

The introduced quantity $\gamma_{ij}(\boldsymbol{x})$ is the 3-metric describing the 3-spaces and has the form:

$$\boxed{\gamma_{ij}(\boldsymbol{x}) = \delta_{ij} + K\frac{x_i x_j}{1 - K(x_k x^k)},} \tag{A.1.10}$$

where K takes on the values -1 0 and $+1$ for negatively curved, flat and positively curved space, respectively.

A.2 FRW Christoffel Symbols

Using the FRW metric $g_{\mu\nu}$, as given by Eq. 1.2.4, the Christoffel symbols $\Gamma^{\mu}_{\rho\sigma}$ can be computed. The Christoffel symbols are defined in terms of the metric tensor and its first derivatives as:

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

Furthermore, they are symmetric in interchanging the lower indices: $\Gamma^{\mu}_{\rho\sigma} = \Gamma^{\mu}_{\sigma\rho}$. All Christoffel symbols with two (or three) temporal indices vanish, that is:

$$\Gamma^{\mu}_{00} = \Gamma^{0}_{0\rho} = \Gamma^{0}_{\sigma 0} = 0. \tag{A.2.2}$$

This only applies when working in coordinate time t, in conformal time $d\tau \equiv dt/a$ this statement is not longer valid and the purely temporal Christoffel is non-zero. Before deriving the non-zero Christoffel symbols below, the final results are given beforehand, so that the reader may decide to ignore the derivations. The non-vanishing Christoffel symbols are given by:

$$\boxed{\Gamma^{0}_{ij} = H g_{ij}, \qquad \Gamma^{i}_{0j} = H\delta^{i}_{j}, \qquad \Gamma^{i}_{jk} = \frac{1}{2}\gamma^{il}\left(\partial_j\gamma_{kl} + \partial_k\gamma_{jl} - \partial_l\gamma_{jk}\right),} \tag{A.2.3}$$

or are related to these by the symmetry in the lower pair of indices.

Derivation of FRW Christoffel Symbols

Upper index zero.—To compute the Christoffel symbol with the upper index equal to zero $\Gamma^{0}_{\rho\sigma}$, first the definition of the Christoffel symbol in terms of the metric is written for $\mu = 0$:

$$\Gamma^{0}_{\rho\sigma} = \frac{1}{2}g^{0\lambda}(\partial_\rho g_{\sigma\lambda} + \partial_\sigma g_{\rho\lambda} - \partial_\lambda g_{\rho\sigma}). \tag{A.2.4}$$

Using the fact that $g^{0\lambda}$ is only non-zero in case $\lambda = 0$, then $g^{00} = -1$, the expression for $\Gamma^{0}_{\rho\sigma}$ becomes:

$$\Gamma^{0}_{\rho\sigma} = -\frac{1}{2}(\underbrace{\partial_\rho g_{\sigma 0} + \partial_\sigma g_{\rho 0}}_{0} - \partial_0 g_{\rho\sigma}). \tag{A.2.5}$$

The first two terms between parenthesis are zero since the only non-zero component of the metric occurs for $\sigma = \rho = 0$, in which case $g_{00} = -1$. Since this is a constant, the result of acting with derivatives ∂_ρ and ∂_σ will be zero. The last term vanishes unless ρ and σ represent spatial indices, since the time-time component g_{00} is time-independent. Using $g_{ij} = a^2\gamma_{ij}$ the time derivative and the final Christoffel symbol are found to be:

$$\partial_0 g_{ij} = 2a\dot{a}\gamma_{ij} = 2\frac{\dot{a}}{a}g_{ij} \quad \longrightarrow \quad \Gamma^{0}_{ij} = H g_{ij}, \tag{A.2.6}$$

where $H \equiv \dot{a}/a$.

Lower index zero.—Setting a lower index equal to zero gives the following expression for the Christoffel symbol as function of the metric:

$$\Gamma^{\mu}_{0\sigma} = \frac{1}{2}g^{\mu\lambda}(\partial_0 g_{\sigma\lambda} + \underbrace{\partial_\sigma g_{0\lambda} - \partial_\lambda g_{0\sigma}}_{0}). \tag{A.2.7}$$

For the same reason as in the previous case, now the last two terms are zero. The first term is zero for $\sigma = \lambda = 0$, so only the spatial part of the metric g_{ij} is relevant. The above expression therefore reduces to:

$$\Gamma^{\mu}_{0i} = \frac{1}{2}g^{\mu j}\partial_0 g_{ij} = \frac{\dot{a}}{a}g^{\mu j}g_{ij}. \tag{A.2.8}$$

Note that the product $g^{\mu j} g_{ij}$ is not possible for $\mu = 0$, so only spatial indices $\mu = k$ are relevant and the final expression for the Christoffel symbol Γ^k_{0i} becomes:

$$\Gamma^k_{0i} = \frac{\dot{a}}{a} g^{kj} g_{ji} = H \delta^k_i, \tag{A.2.9}$$

since $g_{ij} = g_{ji}$ by symmetry and $g^{kj} g_{ji} = \delta^k_i$.

All indices spatial.—For all indices spatial, that is a Christoffel symbol of the form Γ^i_{jk}, the starting point is:

$$\Gamma^i_{jk} = \frac{1}{2} g^{il} \left(\partial_j g_{kl} + \partial_k g_{jl} - \partial_l g_{jk} \right). \tag{A.2.10}$$

Using $g^{il} = \gamma^{il}/a^2$ and $\partial_j g_{kl} = a^2 \partial_j \gamma_{kl}$ the final result is obtained as:

$$\Gamma^i_{jk} = \frac{1}{2} \gamma^{il} \left(\partial_j \gamma_{kl} + \partial_k \gamma_{jl} - \partial_l \gamma_{jk} \right). \tag{A.2.11}$$

Now all Christoffel symbols as stated in Eq. A.2.3 are derived.

A.3 Kinematics in FRW Universe

In the absence of any non-gravitational forces acting on a particle, the particle said to be *freely* falling and moves along a geodesic. Geodesics are the curved-space generalizations of straight lines in flat Euclidean space. Mathematically, geodesics are described by the geodesic equation. Assuming a particle follows the curve $x^\mu(\eta)$ parametrized by proper time η, the geodesic equation is:

$$\boxed{\frac{d^2 x^\mu}{d\eta^2} + \Gamma^\mu_{\rho\sigma} \frac{dx^\rho}{d\eta} \frac{dx^\sigma}{d\eta} = 0.} \tag{A.3.1}$$

Defining 4-velocity U^μ as:

$$U^\mu \equiv \frac{dx^\mu}{d\eta}, \tag{A.3.2}$$

and substituting this definition in the geodesic equation yields:

$$\frac{dU^\mu}{d\eta} + \Gamma^\mu_{\rho\sigma} U^\rho U^\sigma = 0. \tag{A.3.3}$$

p$^\mu Theremainingdifferential$,dU$^\mu/d\eta$, can be rewritten using the chain rule and the geodesic equation becomes:

$$\frac{dU^\mu}{d\eta} = \frac{dU^\mu}{dx^\rho} \frac{dx^\rho}{d\eta} = U^\rho \frac{dU^\mu}{dx^\rho} \quad \longrightarrow \quad U^\rho \left(\frac{dU^\mu}{dx^\rho} + \Gamma^\mu_{\rho\sigma} U^\sigma \right) = 0. \tag{A.3.4}$$

For massive particles, the 4-velocity is related to the 4-momentum via the relation $p^\mu = mU^\mu$ and using 4-momentum the geodesic equation can be written in the following final form:

$$\boxed{p^\sigma \frac{dp^\mu}{dx^\rho} = -\Gamma^\mu_{\rho\sigma} p^\rho p^\sigma.} \tag{A.3.5}$$

Notice that for massless particles $m = 0$ and the utilized relationship between 4-momentum and 4-velocity breaks down. However, the final result for massless particles, such as photons, is identical to Eq. A.3.5.

By the homogeneity of the FRW background, the 4-momentum p^μ is not allowed to vary with space and hence the spatial derivative of p^μ must vanish: $\partial_i p^\mu = 0$. Therefore, the left-hand side of Eq. A.3.5 vanishes unless the index ρ equals zero:

$$\begin{aligned} p^0 \frac{dp^0}{dt} &= -\Big(\Gamma^\mu_{ij} p^i p^j + \underbrace{\Gamma^\mu_{0j} p^0 p^j + \Gamma^\mu_{i0} p^j p^0}_{2\Gamma^\mu_{0j} p^i p^j} + \underbrace{\Gamma^\mu_{00} p^0 p^0}_{0}\Big) \\ &= -\Big(\Gamma^\mu_{ij} p^i p^\rho + 2\Gamma^\mu_{0j} p^0 p^j\Big), \end{aligned} \tag{A.3.6}$$

where the right hand side is expanded into spatial, temporal and mixed terms. The two mixed terms are collected on account of the symmetry $\Gamma^\mu_{i0} = \Gamma^\mu_{0j}$ and the purely temporal term is zero since $\Gamma^\mu_{00} = 0$. Finally, factoring out p^j gives:

$$\boxed{p^0 \frac{dp^\mu}{dt} = -\big(\Gamma^\mu_{ij} p^i + 2\Gamma^\mu_{0j} p^0\big) p^j.} \tag{A.3.7}$$

The first notable result to be extracted from Eq. A.3.7, is that particles at rest in the comoving frame will stay at rest in that frame. For these particles $p^j = 0$ and the right-hand side of Eq. A.3.7 vanishes, resulting in:

$$p^j = 0 \quad \longrightarrow \quad \frac{dp^j}{dt} = 0. \tag{A.3.8}$$

Next, we relax the assumption that the particles are at rest in the comoving frame and the $\mu = 0$ component of Eq. A.3.7 is considered. Using $p^0 = E$ and the fact that the second term on the left hand side vanishes since Γ^0_{0j} yields:

$$E \frac{dE}{dt} = -\Gamma^0_{ij} p^i p^j = -\frac{\dot{a}}{a} g_{ij} p^i p^j, \tag{A.3.9}$$

where in the second equality follows from Eq. A.2.3. Defining the amplitude of physical 3-momentum p^2 as:

$$p^2 \equiv g_{ij} p^i p^j = a^2(t) \gamma_{ij} p^i p^j, \tag{A.3.10}$$

and using the energy momentum relation $E^2 = m^2 + p^2$ to find $E\, dE = p\, dp$ gives:

$$\boxed{\frac{\dot{p}}{p} = -\frac{\dot{a}}{a} \quad \longrightarrow \quad p \propto \frac{1}{a}} \tag{A.3.11}$$

From the inverse relation between p and a, it follows that the physical 3-momentum of both massive and massless particles decays with the expansion of the universe.

For massless particles, e.g. photons, $E = p$ and it follows that both the energy and the momentum decay with the expansion of the universe:

$$E = p \propto \frac{1}{a}. \tag{A.3.12}$$

For massive particles the spatial components of 4-momentum are related to the 4-velocity via $p^i = mU^i$, which may be rewritten as:

$$p^i = mU^i = m\frac{dx^i}{d\eta} = m\frac{dt}{d\eta}\frac{dx^i}{dt} = mv^i\frac{dt}{d\eta}, \tag{A.3.13}$$

where $v^i = dx^i/dt$ is the *comoving* peculiar velocity and $dt/d\eta$ equals the Lorentz factor:

$$\frac{dt}{d\eta} = \frac{1}{\sqrt{1-v^2}} \equiv \gamma. \tag{A.3.14}$$

In expression for the Lorentz factor, v^2 is the physical peculiar velocity $v^2 = g_{ij}v^iv^j$. Using $p^2 = g_{ij}p^ip^j$, the expression for the physical momentum becomes:

$$p = \frac{mv}{\sqrt{1-v^2}} \propto \frac{1}{a}, \tag{A.3.15}$$

for massive particles. Since the scale factor grows as the universe expands, the momentum of freely falling particles eventually decays to zero and they will converge in to the Hubble flow.

Appendix B

Dynamics of the Universe

B.1 Energy-Momentum Tensor and the Cosmological Principle

Here we will derive, on account of the cosmological principle (CP), that the energy-momentum of any constituent in the universe should have the perfect-fluid form at background level. The energy momentum is defined in terms of the 3-scalar T_{00}, 3-vector T_{0i} and the 3-tensor T_{ij} as:

$$T_{\mu\nu} = \begin{bmatrix} T_{00} & T_{0j} \\ T_{i0} & T_{ij} \end{bmatrix}, \tag{B.1.1}$$

where T_{ij} and T_{0j} satisfy the symmetry constraints $T_{0j} = T_{i0}$ and $T_{ij} = T_{ji}$. Now, the components of the energy-momentum will be constrained in such a way that they satisfy homogeneity and isotropy. Recall that the time-time component T_{00} represents the energy density ρ. By homogeneity, ρ may be a function of time, but cannot posses spatial dependence:

$$T_{00} = \rho(t). \tag{B.1.2}$$

The 3-vectors T_{i0} and T_{0j} are equivalent to each other and describe the rate of flow of energy and momentum in each spatial direction. On account of isotropy, the average values of these components vanish:

$$T_{i0} = T_{0j} = 0. \tag{B.1.3}$$

Finally, the shear stress tensor T_{ij} is restricted to be proportional to the metric:[1]

$$T_{ij} \propto g_{ij} \propto \delta_{ij}(\boldsymbol{x} = 0), \tag{B.1.4}$$

since non-vanishing shear stress (off-diagonal) components would break isotropy. The diagonal terms of T_{ij} represent the pressure of the fluid P:

$$T_{ij} = P(t)\, g_{ij}, \tag{B.1.5}$$

where the pressure P is restricted to be a function of time only by isotropy. Using the results above, the energy-momentum tensor becomes:

$$T_{\mu\nu} = \begin{bmatrix} \rho & 0 \\ 0 & P g_{ij} \end{bmatrix}. \tag{B.1.6}$$

[1] From the FRW 3-metric (Eq. A.1.10), it follows that $g_{ij} \propto \delta_{ij}$ at $\boldsymbol{x} = 0$.

B.2 Energy-Momentum Conservation and Continuity Equation

To obtain the continuity equation (Eq. 1.3.6) for the energy density ρ, the $\nu = 0$ component of the conservation equation $\nabla_\mu T^{\mu\nu} = 0$ is considered:

$$\partial_\mu T^{\mu 0} + \Gamma^\mu_{\mu\alpha} T^{\alpha 0} + \Gamma^0_{\mu\alpha} T^{\mu\alpha} = 0. \tag{B.2.1}$$

The first term vanishes unless $\mu = 0$ and therefore it is equal to $\dot{\rho}$. Following the same lines of reasoning, the second term becomes $\Gamma^\mu_{\mu 0}\rho$. Finally, the Christoffel symbol in the last term is only non-vanishing for spatial components and hence the term simplifies to $\Gamma^0_{ij}T^{ij}$ with $T^{ij} = Pg^{ij}$. Taking these results together, the previous equation changes to:

$$\dot{\rho} + \rho\Gamma^\mu_{\mu 0} + \Gamma^0_{ij} P g^{ij} = 0. \tag{B.2.2}$$

Notice that in the second term only the spatial components $\mu = i$ are relevant since the Christoffel symbol is $\Gamma^i_{i0} = H\delta^i_i = 3H$. Similarly, for the third term $\Gamma^0_{ij} = Hg_{ij}$. Now, writing the Christoffel symbols out explicitly yields:

$$\dot{\rho} + 3H\rho + HPg_{ii}g^{ii} = 0. \tag{B.2.3}$$

Finally, using $g_{ii}g^{ii} = \delta^i_i = 3$, the *continuity equation* is obtained for the energy density:

$$\dot{\rho} + 3H(\rho + P) = 0. \tag{B.2.4}$$

B.3 Einstein Tensor FRW Universe

Here, we will explicitly derive the Ricci tensor $R_{\mu\nu}$ and scalar R for the FRW universe, which can then be used to determine the FRW Einstein tensor $G_{\mu\nu}$.

Ricci Tensor: Time-Time Component R_{00}

Setting both indices equal to zero, i.e. $\mu = \nu = 0$, yields the following equation for the time-time component R_{00} in terms of Christoffel symbols:

$$R_{00} = \partial_\lambda \Gamma^\lambda_{00} - \partial_0 \Gamma^\lambda_{\lambda 0} + \Gamma^\lambda_{\lambda\rho}\Gamma^\rho_{00} - \Gamma^\rho_{0\lambda}\Gamma^\lambda_{0\rho} = -\underbrace{\partial_0 \Gamma^\lambda_{\lambda 0}}_{(*)} - \underbrace{\Gamma^\rho_{0\lambda}\Gamma^\lambda_{0\rho}}_{(**)}, \tag{B.3.1}$$

where the first and third term vanish since $\Gamma^\lambda_{00} = \Gamma^\rho_{00} = 0$. The $(*)$-term is computed as:

$$(*) = \partial_0 \Gamma^\lambda_{\lambda 0} = \partial_0 \Gamma^i_{i0} = \partial_0 \left(3\frac{\dot{a}}{a}\right) = 3\left(\frac{a\ddot{a} - \dot{a}^2}{a^2}\right), \tag{B.3.2}$$

where only the spatial components $\lambda = i$ are considered since the expression vanishes for $\lambda = 0$. Furthermore, the Christoffel symbol is $\Gamma^i_{i0} = (\dot{a}/a)\delta^i_i = 3(\dot{a}/a)$ because i is summed over and hence $\delta^i_i = 3$. The $(**)$-term is evaluated as follows:

$$(**) = \Gamma^\rho_{0\lambda}\Gamma^\lambda_{0\rho} = \Gamma^i_{0j}\Gamma^j_{0i} = \Gamma^i_{0j}\Gamma = \left(\frac{\dot{a}}{a}\right)^2 \delta^i_j \delta^j_i = 3\left(\frac{\dot{a}}{a}\right)^2, \tag{B.3.3}$$

since the expression vanishes for $\rho = \lambda = 0$ and $\delta^i_j\delta^j_i = \delta^i_i = 3$. Finally, the time-time component of the Ricci scalar R_{00} is:

$$R_{00} = -(*) - (**) = -3\frac{\ddot{a}}{a}. \tag{B.3.4}$$

Ricci Tensor: Spatial Components R_{ij}

Computing the spatial part of the Ricci tensor R_{ij} is reasonably more involved compared with the time-time component R_{00}. In order to simplify matters, the following approach will be taken. First, the spatial part of the Ricci tensor will be evaluated at $\boldsymbol{x} = 0$, that is $R_{ij}(\boldsymbol{x} = 0)$. Then, after performing the computations the results will be generalized to arbitrary $\boldsymbol{x}$. Setting $\boldsymbol{x} = 0$ causes many terms to vanish and hence the derivation will become less computationally demanding.

It should be noted that this approach must be taken with caution. Naively, one could use Eq. A.1.10 for the spatial metric γ_{ij} and set $\boldsymbol{x} = 0$ to obtain:

$$\gamma_{ij}(\boldsymbol{x}) = \delta_{ij} + K\frac{x_i x_j}{1 - K(x_k x^k)} \quad \xrightarrow{\boldsymbol{x}\,=\,0} \quad \gamma_{ij} = \delta_{ij}, \tag{B.3.5}$$

and proceed with this expression for the spatial metric. However, the Ricci tensor R_{ij} contains second order derivatives with respect to the spatial metric. Although γ_{ij} reduces to δ_{ij} for $\boldsymbol{x} = 0$, these second order derivatives may get a contribution from the second term of γ_{ij} at $\boldsymbol{x} = 0$. Therefore the considered metric will be:

$$\gamma_{ij}(\boldsymbol{x} = 0) = \delta_{ij} + K x_i x_j, \tag{B.3.6}$$

since the denominator of the second term can be safely ignored ($x_k x^k = 0$). With the suitable form of the metric at hand, the R_{ij} can be computed according to:

$$R_{ij}(\boldsymbol{x} = 0) = \underbrace{\partial_\lambda \Gamma^\lambda_{ij}}_{(1)} - \underbrace{\partial_j \Gamma^\lambda_{i\lambda}}_{(2)} + \underbrace{\Gamma^\lambda_{\lambda\rho}\Gamma^\rho_{ij}}_{(3)} - \underbrace{\Gamma^\rho_{i\lambda}\Gamma^\lambda_{j\rho}}_{(4)}. \tag{B.3.7}$$

Below, the indicated terms (1) up to and including (4) will be evaluated individually.

First Term.—Expanding the first term $\partial_\lambda \Gamma^\lambda_{ij}$ into $\lambda = 0$ and $\lambda = i$ yields:

$$\partial_\lambda \Gamma^\lambda_{ij} = \partial_0 \Gamma^0_{ij} + \partial_k \Gamma^k_{ij}. \tag{B.3.8}$$

Computing the temporal term in this expansion gives:

$$\partial_0 \Gamma^0_{ij} = \partial_0 \left(\frac{\dot{a}}{a} g_{ij}\right) = \partial_0(\dot{a} a \gamma_{ij}) = (a\ddot{a} + \dot{a}^2)\delta_{ij}, \tag{B.3.9}$$

where, after computing the derivative, the metric γ_{ij} is set to δ_{ij} in the last equality. The spatial term $\partial_k \Gamma^k_{ij}$ requires the explicit computation of the spatial Christoffel symbol Γ^k_{ij} for $\gamma_{ij}(\boldsymbol{x} = 0)$:

$$\Gamma^k_{ij} = \frac{1}{2}\gamma^{kl}(\partial_i \gamma_{jl} + \partial_j \gamma_{il} - \partial_l \gamma_{ij}). \tag{B.3.10}$$

Evaluating the partial derivative $\partial_i \gamma_{jl}$ gives the following result:

$$\partial_i \gamma_{jl} = \partial_i(\delta_{jl} + K x_j x_l) = K(x_l \delta_{ij} + x_j \delta_{il}). \tag{B.3.11}$$

In a similar way, the two remaining derivatives can be found and Γ^k_{ij} can be written becomes:

$$\begin{aligned}\Gamma^k_{ij} &= \frac{1}{2}\gamma^{kl}\left(K(x_l\delta_{ij} + x_j\delta_{il}) + K(x_l\delta_{ij} + x_i\delta_{jl}) - K(x_j\delta_{li} + x_i\delta_{lj})\right) \\ &= \gamma^{kl}(K x_l \delta_{ij}) = K x^k \delta_{ij},\end{aligned} \tag{B.3.12}$$

since $\gamma^{kl}x_l = x^k$. Consequently, the partial derivative $\partial_k \Gamma^k_{ij}$ is:

$$\partial_k \Gamma^k_{ij} = \partial_k (K x^k \delta_{ij}) = 3K\delta_{ij}. \tag{B.3.13}$$

Finally, the first term becomes:

$$(1) = \partial_\lambda \Gamma^\lambda_{ij} = (a\ddot{a} + \dot{a}^2 + 3K)\delta_{ij}. \tag{B.3.14}$$

Second Term.—Expanding the second term into a temporal and spatial term with respect to λ gives:

$$\partial_j \Gamma^\lambda_{i\lambda} = \partial_j \Gamma^0_{i0} + \partial_j \Gamma^k_{ik} = \partial_j (K x^k \delta_{ik}) = K\delta_{ij}, \tag{B.3.15}$$

since $\Gamma^0_{i0} = 0$ and $x^k \delta_{ik} = x_i$. Hence, the second term yields:

$$(2) = \partial_j \Gamma^\lambda_{i\lambda} = K\delta_{ij}. \tag{B.3.16}$$

Third Term.—Expansion of the third term $\Gamma^\lambda_{\lambda\rho}\Gamma^\rho_{ij}$ yields:

$$\Gamma^\lambda_{\lambda\rho}\Gamma^\rho_{ij} = \Gamma^0_{0\rho}\Gamma^\rho_{ij} + \Gamma^l_{l\rho}\Gamma^\rho_{ij} = \Gamma^l_{l\rho}\Gamma^\rho_{ij} = \underbrace{\Gamma^l_{l0}\Gamma^0_{ij}}_{(*)} + \underbrace{\Gamma^l_{lm}\Gamma^m_{ij}}_{(**)}. \tag{B.3.17}$$

The $(*)$ term gives:

$$(*) = \Gamma^l_{l0}\Gamma^0_{ij} = \left(\frac{\dot{a}}{a}\delta^l_l\right)(\dot{a}a\gamma_{ij}) = 3\dot{a}^2\gamma_{ij} = 3\dot{a}^2\delta_{ij}, \tag{B.3.18}$$

where in the last equality use is made of the fact that $\gamma_{ij} = \delta_{ij}$ at $\boldsymbol{x} = 0$. Computing the $(**)$ term gives:

$$(**) = \Gamma^l_{lm}\Gamma^m_{ij} = (K x^l \delta_{lm})(K x^m \delta_{ij}) = K^2 x_m x^m \delta_{ij} = 0, \tag{B.3.19}$$

since $x_m x^m = 0$ at $\boldsymbol{x} = 0$. Hence, the third term in the expansion of R_{ij} becomes:

$$(3) = (*) + (**) = 3\dot{a}^2\delta_{ij}. \tag{B.3.20}$$

Fourth Term.—Finally, the fourth term can be expanded into:

$$\Gamma^\rho_{i\lambda}\Gamma^\lambda_{j\rho} = \Gamma^\rho_{i0}\Gamma^0_{j\rho} + \Gamma^\rho_{im}\Gamma^m_{j\rho} = \Gamma^0_{i0}\Gamma^0_{j0} + \underbrace{\Gamma^m_{i0}\Gamma^0_{jm}}_{(*)} + \underbrace{\Gamma^0_{im}\Gamma^m_{j0}}_{(**)} + \underbrace{\Gamma^l_{im}\Gamma^m_{jl}}_{(***)}, \tag{B.3.21}$$

where the first term in the final expansion vanishes and the labelled terms will be computed separately. The $(*)$ and $(**)$ terms are straightforward by now:

$$(*) = \Gamma^m_{i0}\Gamma^0_{jm} = \left(\frac{\dot{a}}{a}\delta^m_i\right)(\dot{a}a\delta_{jm}) = \dot{a}^2\delta_{ij} \tag{B.3.22}$$

$$(**) = \Gamma^0_{im}\Gamma^m_{j0} = (\dot{a}a\gamma_{im})\left(\frac{\dot{a}}{a}\delta^m_j\right) = \dot{a}^2\delta_{ij}, \tag{B.3.23}$$

since $\gamma_{ij} = \delta_{ij}$ at $\boldsymbol{x} = 0$. Computing the $(***)$ term gives:

$$\begin{aligned}(***) &= \Gamma^l_{im}\Gamma^m_{jl} = (Kx^l\delta_{im})(Kx^m\delta_{jl}) = K^2x^l\delta_{jl}x^m\delta_{im} \\ &= K^2x_jx_m = K^2x^mx_m\delta^{mj} = 0,\end{aligned} \tag{B.3.24}$$

since, again, $x^mx_m = 0$ at $\boldsymbol{x} = 0$. Consequently, the fourth term is:

$$(4) = (*) + (**) + (***) = 2\dot{a}^2\delta_{ij} \tag{B.3.25}$$

Final Result.—Now that the terms (1) up to and including (4) are computed, the spatial part of the Ricci tensor R_{ij} at $\boldsymbol{x} = 0$ can be obtained by combining these four terms in the appropriate way:

$$R_{ij}(\boldsymbol{x} = 0) = (1) - (2) + (3) - (4) = (a\ddot{a} + 2\dot{a}^2 + 2K)\delta_{ij} \tag{B.3.26}$$

To generalize this result to arbitrary $\boldsymbol{x}$, note that $\gamma_{ij}(\boldsymbol{x} = 0) = \delta_{ij}$ and hence R_{ij} for any $\boldsymbol{x}$ is:

$$R_{ij}(\boldsymbol{x}) = (a\ddot{a} + 2\dot{a}^2 + 2K)\gamma_{ij} = \left[\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{K}{a^2}\right]g_{ij}. \tag{B.3.27}$$

Ricci Scalar R

The Ricci scalar is obtained via the definition $R = g^{\mu\nu}R_{\mu\nu}$, expanding this product in temporal, spatial and mixed terms gives:

$$R = g^{00}R_{00} + g^{i0}R_{i0} + g^{0j}R_{0j} + g^{ij}R_{ij} = g^{00}R_{00} + g^{ij}R_{ij}, \tag{B.3.28}$$

since the off-diagonal terms of the Ricci tensor vanish. Computing the remaining two terms yields:

$$\begin{aligned}R &= -R_{00} + \frac{1}{a^2}\gamma^{ij}R_{ij} = 3\frac{\ddot{a}}{a} + \frac{1}{a^2}(a\ddot{a} + 2\dot{a}^2 + 2K)\gamma_{ij}\gamma^{ij} \\ &= 3\frac{\ddot{a}}{a} + \frac{3}{a^2}(a\ddot{a} + 2\dot{a}^2 + 2K) = 6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2}\right],\end{aligned} \tag{B.3.29}$$

because $\gamma_{ij}\gamma^{ij} = \delta^i_i = 3$.

B.4 Friedmann Equations

Here, we will explicitly derive the first and second Friedmann equation (Eqs. 1.3.21 and 1.3.22).

Derivation of the Friedmann Equations

First Equation.—Considering the $\mu = \nu = 0$ equation of EFE's for an FRW universe (Eq. 1.3.20) yields:

$$G_{00} + \Lambda g_{00} = R_{00} - 3\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2} + \frac{\Lambda}{3}\right] g_{00} = \frac{1}{M_{\rm pl}^2} T_{00}. \tag{B.4.1}$$

Since $R_{00} = -3\ddot{a}/a$, $g_{00} = -1$ and $T_{00} = \rho$, the above equation becomes:

$$\begin{aligned} \frac{\rho}{M_{\rm pl}^2} &= -3\frac{\ddot{a}}{a} + 3\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2} + \frac{\Lambda}{3}\right] \\ \frac{\rho}{M_{\rm pl}^2} &= 3\left[\left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2} + \frac{\Lambda}{3}\right]. \end{aligned} \tag{B.4.2}$$

Rearranging immediately gives the first Friedmann equation (Eq. 1.3.21):

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2} = \frac{1}{3M_{\rm pl}^2}\left(\rho + \rho_\Lambda\right), \tag{B.4.3}$$

since the energy density due to the cosmological constant ρ_Λ is given by $\rho_\Lambda = \Lambda M_{\rm pl}^2$.

Second Equation.—As already mentioned, the second Friedmann equation is obtained by considering the spatial components $\mu = i$ and $\nu = j$ of Eq. 1.3.20:

$$G_{ij} + \Lambda g_{ij} = R_{ij} - \frac{1}{2} R g_{ij} + \Lambda g_{ij} = \frac{T_{ij}}{M_{\rm pl}^2}. \tag{B.4.4}$$

Substituting the obtained expressions for the spatial Ricci tensor R_{ij} (Eq. 1.3.16), the Ricci scalar R (Eq. 1.3.18) and the spatial part of the FRW energy-momentum tensor T_{ij} (Eq. 1.3.2) yields:

$$\left[\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{K}{a^2}\right] g_{ij} - 3\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2}\right] g_{ij} + \Lambda g_{ij} = \frac{P}{M_{\rm pl}^2} g_{ij}. \tag{B.4.5}$$

Since all terms are proportional to the spatial metric g_{ij}, the prefactors must satisfy the following equation:

$$\begin{aligned} \frac{P}{M_{\rm pl}^2} &= \left[\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{K}{a^2}\right] - 3\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2}\right] + \Lambda \\ \frac{P}{M_{\rm pl}^2} &= -2\frac{\ddot{a}}{a} - \left(\frac{\dot{a}}{a}\right)^2 - \frac{K}{a^2} + \Lambda. \end{aligned} \tag{B.4.6}$$

The above equation can not be simplified further in a convenient way. In order to proceed, the first Friedmann equation (Eq. 1.3.21) is added to the preceding equation, yielding:

$$-2\frac{\ddot{a}}{a} + \frac{2}{3}\Lambda = \frac{1}{3M_{\rm pl}^2}(3P + \rho). \tag{B.4.7}$$

Finally using the energy density $\rho_\Lambda = \Lambda M_{\rm pl}^2$ reproduces the second Friedmann equation (Eq. 1.3.22):

$$\frac{\ddot{a}}{a} = -\frac{1}{6M_{\rm pl}^2}(3P + \rho - 2\rho_\Lambda). \tag{B.4.8}$$

Appendix C

Big Bang Puzzles and Classical Dynamics of Inflation

C.1 Horizon Problem: Quantitative Analysis

Here, we will discuss the horizon problem in a quantitative way and show that CMB patches separated by more than two degrees should be causally disconnected. To show this, it will turn out to be convenient to first define a number of quantities using Fig. 2.3 (right). The comoving distances between the singularity and recombination and between recombination and now are defined as:

$$d_{\rm hor} \equiv \tau_{\rm rec} - \tau_i, \qquad d_{\rm A} \equiv \tau_0 - \tau_{\rm rec}, \tag{C.1.1}$$

respectively (recall that $c \equiv 1$). Since $d_{\rm hor} \ll d_{\rm A}$, the arclength on the green circle between the two dotted radial lines in Fig. 2.3 (right) may be approximated by $d_{\rm hor}$. The angle θ_c is then equal to:

$$\theta_c = \frac{d_{\rm hor}}{d_{\rm A}}. \tag{C.1.2}$$

In order to find θ_c and hence θ, both $d_{\rm hor}$ and $d_{\rm A}$ must be evaluated. This can be done by using the definition of the particle horizon. Using Eq. 2.1.5, the difference between $\tau_2 - \tau_1$ with $\tau_2 > \tau_1$ can be written as the following integral:

$$\boxed{\mathcal{I}(t_1, t_2) \equiv \tau_2 - \tau_1 = \int_{t_1}^{t_2} \frac{dt}{a(t)},} \tag{C.1.3}$$

such that $d_{\rm hor} = \mathcal{I}(t_i, t_{\rm rec})$ and $d_{\rm A} = \mathcal{I}(t_{\rm rec}, t_0)$. To simplify matters, the integral $\mathcal{I}$ will be rewritten in terms the redshift z. The definition of redshift is:

$$z + 1 = \frac{1}{a}, \tag{C.1.4}$$

where $a < a_0 \equiv 1$. This definition can be used to relate the differential dt to dz as follows:

$$dz = -\frac{da}{a^2} = -\frac{\dot{a}}{a^2}dt = -\frac{H}{a}dt \quad \longrightarrow \quad dt = -\frac{a}{H}dz, \tag{C.1.5}$$

and the integral $\mathcal{I}(t_1, t_2)$ can be changed to:

$$\begin{aligned}\mathcal{I}(z_1, z_2) &= \int_{t_1}^{t_2} \frac{dt}{a(t)} = \int_{z_2}^{z_1} \frac{dz}{H(z)} \\ &= \int_{z_2}^{z_1} \frac{dz}{H_0\sqrt{\Omega_\Lambda + \Omega_m(1+z)^3 + \Omega_r(1+z)^4}}.\end{aligned} \tag{C.1.6}$$

In the last equality H is expressed in terms of z and density parameters Ω_Λ, Ω_m and Ω_r.

Using the above result and the fact that recombination occurs at a redshift of approximately $z_{\rm rec} = 1090$, $d_{\rm hor}$ and $d_{\rm A}$ are given by:

$$d_{\rm hor} = \mathcal{I}(\infty, z_{\rm rec}) = \int_{z_{\rm rec}}^{\infty} \frac{dz}{H(z)}, \qquad d_{\rm A} = \mathcal{I}(z_{\rm rec}, 0) = \int_0^{z_{\rm rec}} \frac{dz}{H(z)}, \tag{C.1.7}$$

and numerical integration yields the following result for the ratio between them:

$$\theta_c = \frac{d_{\rm hor}}{d_{\rm A}} = 0.020 \text{ rad} = 1.18^\circ \tag{C.1.8}$$

Hence, θ is indeed approximately 2°. This result verifies the statement that CMB patches separated by more than two degrees should be causally disconnected according to conventional Big Bang theory.

C.2 Equivalent Definitions of Inflation

In the main text, inflation is defined as period in the early universe characterized by a shrinking Hubble sphere $(aH)^{-1}$ (Eq. 2.4.2), since this definition of inflation relates most easily to the horizon and flatness problems. However, the shrinking Hubble sphere is not only possible definition of inflation. In literature, other definitions are also used to describe inflation. Here, the equivalence between the shrinking Hubble sphere and other commonly used definitions of inflation will be shown.

▷ *Accelerated Expansion.*—Perhaps the best known definition of inflation is a period accelerated expansion, characterized by the condition that second derivative of the scale factor is greater than zero:

$$\ddot{a} > 0. \tag{C.2.1}$$

To see that this definition of inflation is equivalent to the shrinking Hubble sphere, explicitly differentiate $(aH)^{-1}$ with respect to time:

$$\frac{d}{dt}(aH)^{-1} = \frac{d}{dt}(\dot{a})^{-1} = -\frac{\ddot{a}}{\dot{a}^2} < 0. \tag{C.2.2}$$

Since $\dot{a}^2$ is always positive, it follows directly that a shrinking Hubble sphere is equivalent to accelerated expansion $\ddot{a} > 0$.

▷ *Almost Constant Hubble Parameter.*—The time derivative of the Hubble sphere may also be evaluated by keeping H as a time dependent variable instead of rewriting aH as $\dot{a}$. The time derivative then becomes:

$$\frac{d}{dt}(aH)^{-1} = -\frac{\dot{a}H + a\dot{H}}{(aH)^2} = -\frac{1}{a}\left(1 + \frac{\dot{H}}{H^2}\right) \equiv -\frac{1}{a}(1-\varepsilon) < 0. \tag{C.2.3}$$

Here, the first Hubble flow parameter ε is defined as:

$$\boxed{\varepsilon \equiv -\frac{\dot{H}}{H^2} < 1,} \tag{C.2.4}$$

which must be smaller than one in order to satisfy $-(1-\varepsilon)/a < 0$. That is, the time variation of the Hubble parameter $\dot{H}$ is much smaller H and hence the Hubble constant is almost constant during inflation.

▷ *Quasi De Sitter space-time.*—Since $\dot{H} \simeq 0$ during inflation, one may find an approximate form for the scale factor by setting $H \neq H(t)$ such that the differential equation for a in terms of t becomes:

$$H \equiv \frac{\dot{a}}{a} \quad \longrightarrow \quad \frac{da}{dt} = Ha. \tag{C.2.5}$$

The solution to this differential equation is:

$$a(t) = e^{Ht}, \tag{C.2.6}$$

where the integration constant is set to one. Note that the above solution is only valid in the approximation that H is constant over time, or equivalently for $\varepsilon = 0$. Using this form of the scale factor, the space-time line element ds^2 becomes:

$$ds^2 = -dt^2 + a^2\, d\boldsymbol{x}^2 = -dt^2 + e^{2Ht}\, d\boldsymbol{x}^2, \tag{C.2.7}$$

which coincides with the De Sitter space-time. However, inflation has to end for the universe to be able transit to the successive eras such as radiation and matter dominated eras. Therefore, the Hubble constant cannot be completely time independent and hence ε cannot vanish entirely. Nevertheless, for small and finite ε, the De Sitter space-time remains a reasonable approximation to the inflationary background. For that reason, inflation is often described by a quasi De Sitter space-time.

▷ *Negative Pressure and SEC-Violation.*—Inflation requires a fluid with energy density ρ and pressure P that is described by an equation of state:

$$w = P/\rho < -1/3. \tag{C.2.8}$$

That is, the fluid driving inflation does not respect the strong energy condition $w > -1/3$ (SEC). Furthermore, for the fluid to satisfy $w < -1/3$, it must be characterized by a negative pressure, since ρ is positive.

To see why inflation requires $w = P/\rho < -1/3$, consider the second Friedmann equation (Eq. 1.3.22) without the explicit dark energy density ($\rho_\Lambda \equiv 0$):

$$\frac{\ddot{a}}{a} = \dot{H} + H^2 = -\frac{1}{6M_{\rm pl}^2}(\rho + 3P) = -\frac{\rho}{6M_{\rm pl}^2}(1+3w). \tag{C.2.9}$$

Combining the above equation with the first Friedmann equation $\rho = 3M_{\rm pl}^2 H^2$ (Eq. 1.3.21, again with $\rho_\Lambda \equiv 0$) gives:

$$\dot{H} + H^2 = -\frac{H^2}{2}(1+3w) \quad \longrightarrow \quad \varepsilon = -\frac{\dot{H}}{H^2} = \frac{3}{2}(1+w). \tag{C.2.10}$$

Since during inflation $\varepsilon < 1$ it follows that the fluid driving inflation satisfies $w < -1/3$, which violates the SEC. Note that it is unsurprising that inflation violates the SEC, since the SEC gives the condition for w to get a growing Hubble sphere: during inflation the opposite occurs and hence the violation is a logical consequence.

▷ *Constant Energy Density.*—Combining the continuity equation in the form of Eq. 1.3.8 with the result $2\varepsilon = 3(1+w)$ from the previous equation yields:

$$\left|\frac{d\ln\rho}{d\ln a}\right| = 2\varepsilon. \tag{C.2.11}$$

Hence, for small ε the energy density remains nearly constant during inflation.

C.3 Klein-Gordon Equation in FRW Space-Time

The Klein-Gordon (KG) equation for the scalar field (inflaton) is derived here via the principle of least action. That is, the action S_ϕ (corresponding to the Lagrangian $\mathcal{L}_\phi$ as given by Eq. 2.4.16) will be varied with respect to ϕ and the resulting expression is set equal to zero. The variation δS_ϕ is given by:

$$\delta S_\phi = \int d^4x\sqrt{-g}\Big[\underbrace{-g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi}_{(1)}\ \underbrace{-V_\phi(\phi)\delta\phi}_{(2)}\Big], \tag{C.3.1}$$

which contains two terms, labelled (1) and (2). The separate actions corresponding to these terms are:

$$\delta S_\phi^{(1)} = -\int d^4x\sqrt{-g}\ g^{\mu\nu}\partial_\mu\phi\partial_\nu\delta\phi, \qquad \delta S_\phi^{(2)} = -\int d^4x\sqrt{-g}\ V_\phi(\phi)\delta\phi. \tag{C.3.2}$$

Integrating the first contribution $\delta S_\phi^{(1)}$ by parts over the $\partial_\nu\delta\phi$ term yields:

$$\delta S_\phi^{(1)} = \underbrace{-\sqrt{-g}g^{\mu\nu}\partial_\mu\phi\ \delta\phi}_{0,\ \delta\phi\to 0}\Big|_{\partial\mathcal{M}} + \int d^4x\partial_\nu\left(\sqrt{-g}g^{\mu\nu}\partial_\mu\phi\right)\delta\phi. \tag{C.3.3}$$

The boundary term vanishes since it is assumed by the boundary conditions of ϕ that $\delta\phi$ vanishes on $\partial\mathcal{M}$. Combining the new form of $\delta S_\phi^{(1)}$ with $\delta S_\phi^{(2)}$ gives the following expression for δS_ϕ:

$$\delta S_\phi = \int d^4x\sqrt{-g}\left[\frac{1}{\sqrt{-g}}\partial_\nu\left(\sqrt{-g}g^{\mu\nu}\partial_\mu\phi\right) - V_\phi(\phi)\right]\delta\phi \equiv 0, \tag{C.3.4}$$

which is set equal to zero by the principle of least action. Since the above equation should hold for any $\delta\phi$ the equation of motion for ϕ is:

$$\frac{1}{\sqrt{-g}}\partial_\nu\left(\sqrt{-g}g^{\mu\nu}\partial_\mu\phi\right) = V_\phi(\phi). \tag{C.3.5}$$

The left hand side is by definition equal to $\Box\phi$, this confirms the Klein-Gordon equation as given by Eq. 2.5.2.

Appendix D

Cosmological Perturbations

D.1 The Central Limit Theorem

On account of the central limit theorem (CLT), we expect cosmological perturbations to be Gaussian to very high degree. We will illustrate the CLT here by considering a one-dimensional uncorrelated random walk. The walk consists of N independents steps of length x_n $(n = 1, \ldots, N)$ and the steps x_n are all drawn from a common probability density function $\rho(x_n)$ with finite variance, i.e. $\sigma^2 < \infty$. For simplicity, we assume the average displacement in each step is zero so that $\langle x_{(n)} \rangle = 0$. To generalize the result to be derived below to the case of an arbitrary finite average, simply add a drift term $N\langle x_{(n)} \rangle$.

The aim of this illustrative example will be to compute the PDF for the total distance covered by the one dimensional random walk after N steps. That is, we want calculate the PDF for the statistical variable:

$$X_N \equiv \sum_{n=1}^{N} x_n. \tag{D.1.1}$$

However, for convenience we consider the rescaled variable $U_N \equiv X_N/\sqrt{N}$. In both cases, the PDF to be computed is that of the sum of the independent random variables. The PDF of the sum of two independent random variables X and Y is given by the convolution of their seperate PDF's $\rho_X(x)$ and $\rho_Y(y)$. The convolution is defined as:

$$\rho_{X+Y} \equiv (\rho_Y * \rho_X)(x) = \int_{-\infty}^{+\infty} \rho_Y(y)\rho_X(x-y)dy. \tag{D.1.2}$$

Generalizing to the sum of several random variables $X_1, \ldots X_N$ the convolution gives:

$$\rho_{X_1+\cdots+X_N}(x) = (\rho_{X_1} * \cdots * \rho_{X_N})(x). \tag{D.1.3}$$

A more practical expression for the PDF of some variable Y that is a function F of the identically distributed independent variables $X_1, \ldots, X_N$, that is:

$$Y = F(X_1, \ldots X_N). \tag{D.1.4}$$

The PDF for the random variable Y, denoted as $\rho_Y(y)$, can be computed via the following equation involving the one-dimensional Dirac delta function:

$$\rho_Y(y) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \rho_{X_1}(x_1) \cdots \rho_{X_N}(x_N)\ \delta^{(1)}(F-y)\ dx_1 \cdots dx_N. \tag{D.1.5}$$

Evaluating this expression for the case of the N-step random walk in the limit of a large number of steps shows that ρ_U is a Gaussian:

$$\rho_U \equiv \lim_{N\to\infty} \rho_{U_N} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{U^2}{2\sigma^2}\right], \tag{D.1.6}$$

where we denote the variance as $\sigma^2 = \langle x^2 \rangle$. This is an illustrative example in which the combination of a large number of random processes (the individual steps) give rise to Gaussianity, since the total distance covered by the walk is Gaussian-distributed. The derivation of the above result is shown in the box below.

Derivation: Random Walk and CLT

Here, we will explicitly evaluate Eq. D.1.5 for the N-step one dimensional random walk (based on [41]). Substituting the relevant expressions into Eq. D.1.5, the PDF for the rescaled distance $U_N \equiv X_N/\sqrt{N}$ covered by a random walk of N steps is given by:

$$\begin{aligned}\rho_{U_N} &= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \left[\prod_{n=1}^{N} dx_n\, \rho(x_n)\right] \delta^{(1)}\Big(\sum_{n=1}^{N} x_n/\sqrt{N} - U_N\Big) \\ &= \frac{1}{2\pi} \int dk\, e^{-ikU_N} \left[\prod_{n=1}^{N} \int_{-\infty}^{\infty} dx_n\, \rho(x_n)\, e^{ikx_n/\sqrt{N}}\right]. \end{aligned} \tag{D.1.7}$$

In the last line, we Fourier transformed the Dirac delta, the factor of $(2\pi)^{-1}$ is included as this is a one-dimensional Dirac delta function (and therefore denoted using a superscript (1) as $\delta^{(1)}$). As the PDF is the same for each step, we can write the product as:

$$\prod_{n=1}^{N} \int_{-\infty}^{\infty} dx_n\, \rho(x_n)\, e^{ikx_n/\sqrt{N}} = \left[\int dx\, \rho(x)\, e^{ikx/\sqrt{N}}\right]^N. \tag{D.1.8}$$

Expanding the complex exponential as a Taylor series yields:

$$e^{ikx/\sqrt{N}} = 1 + \frac{ik}{\sqrt{N}} x + \frac{1}{2}\left(\frac{ik}{\sqrt{N}}\right)^2 x^2 + \mathcal{O}(k^3/N^{3/2}). \tag{D.1.9}$$

The product then becomes:

$$\begin{aligned}\prod_{n=1}^{N} \int_{-\infty}^{\infty} dx_n\, \rho(x_n)\, e^{ikx_n/\sqrt{N}} &= \left[1 + \frac{ik}{\sqrt{N}}\langle x\rangle + \frac{1}{2}\left(\frac{ik}{\sqrt{N}}\right)^2 \langle x^2\rangle + \mathcal{O}(k^3/N^{3/2})\right]^N \\ &= \left[1 - \frac{k^2}{2N}\langle x^2\rangle + \mathcal{O}(k^3/N^{3/2})\right]^N. \end{aligned} \tag{D.1.10}$$

By assumption, we took the average to vanish, so that the second term in the expansion vanishes. The third term is finite since we assumed the variance to be finite. Now, consider the limit of a large number of steps, corresponding to $N \to \infty$. In that case higher order terms than N^{-1} can be neglected and the limit can be written as:

$$\lim_{N\to\infty} \prod_{n=1}^{N} \int_{-\infty}^{\infty} dx_n\, \rho(x_n)\, e^{ikx_n/\sqrt{N}} = \exp\left[-\frac{k^2}{2}\langle x^2\rangle\right]. \tag{D.1.11}$$

Substituting this result into the equation for ρ_U and evaluating the integral shows that the PDF is a Gaussian in the limit of a large number of steps:

$$\rho_U \equiv \lim_{N\to\infty} \rho_{U_N} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} dk\ \exp\left[-ikU - \frac{k^2}{2}\langle x^2\rangle\right] = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{U^2}{2\sigma^2}\right], \quad \text{(D.1.12)}$$

where we defined $U \equiv \lim_{N\to\infty} U_N$ and denote the variance as $\sigma^2 = \langle x^2\rangle$.

D.2 Vector Perturbations

In the main text, and in particular in chapters 3 and 4, it is mentioned that vector perturbations constitute a subdominant contribution to the perturbed dynamics of the early universe and can typically be neglected. Here, we will make this statements more precise, following [15]. For purely vectorial perturbations, the perturbed line element reads:

$$ds^2 = a\Big[-d\tau^2 + 2\hat{B}_i\, dx^i d\tau + (\delta_{ij} + 2\partial_{(i}\hat{E}_{j)})dx^i dx^j\Big]. \quad \text{(D.2.1)}$$

Under the vector gauge transformation:

$$\boldsymbol{x} \to \boldsymbol{x} + \boldsymbol{\xi}, \quad \text{(D.2.2)}$$

where $\partial^i \xi_i = 0$, the perturbations transform as:

$$\hat{B}_i \to \hat{B}_i + \xi_i', \qquad \hat{E}_i \to \hat{E}_i - \xi_i. \quad \text{(D.2.3)}$$

Hence, the combination $\hat{\sigma}_i \equiv \hat{E}_i' + \hat{B}_i$ is manifestly gauge-invariant by construction and is called the vector-shear perturbation.

For purely vectorial perturbations, there are two perturbed field equations. Assuming a fluid with anisotropic stress contribution $\delta\Sigma_{ij}$ (see section 4.8.1) and velocity v^i, the vectorial contributions to the anistropic stress and velocity are given by $\delta\Sigma^{\rm V}_{ij} = \partial_{(i}\hat{\Sigma}_{j)}$ and $v_i^{\rm V} = \hat{v}_i$, respectively. Explicitly, in conformal time the perturbed field equations read [15]:

$$\hat{v}_i' + 3\mathcal{H}\hat{v}_i = -a\partial^2\delta\hat{\Sigma}_i, \quad \text{(D.2.4)}$$

$$\hat{v}_i = -(M_{\rm pl}^2/2)\partial^2\hat{\sigma}_i. \quad \text{(D.2.5)}$$

In the absence of anisotropic stress, which is the case in the single field scenario, we can set $\delta\hat{\Sigma}_i \equiv 0$ and the solution to $\hat{v}_i$ is a decaying mode:

$$\hat{v}_i = \frac{3\hat{\mathcal{C}}_i}{a}, \quad \text{(D.2.6)}$$

where $\hat{\mathcal{C}}_i$ is a generic time-independent divergenceless 3-vector. Hence, at late times, corresponding to large a, the vectorial velocity perturbation $\hat{v}_i$ vanishes. On account of the second field equation, we also conclude that the gauge invariant vector-shear perturbation $\hat{\sigma}_i$ vanishes as $\hat{v}_i$ tends to zero. Therefore, we conclude indeed that vector perturbations play a subdominant role and they can safely be ignored.

D.3 Independent Evolution of SVT Components

The independent evolution of Scalar-Vector-Tensor (SVT) components is most easily shown in Fourier space as the different $k-$modes evolve independently at linear order (this statement will be proved below as well). A second important reason to work in Fourier space is the fact that a particular $k-$mode of a perturbation can be compared directly to the comoving Hubble radius and hence one can easily discriminate different stages in its evolution, e.g. sub-horizon and super-horizon stages (recall Fig. 3.2). First, we will prove that the SVT components evolve independently and subsequently the decomposition will be performed for a 3-scalar, 3-vector and 3-tensor in both spatial $\boldsymbol{x}-$space and momentum $\boldsymbol{k}-$space via the Fourier transform.

The proofs presented below are based on spatial and rotational invariance of the background, and are therefore quite generic. The proofs below are originally from Seljak and Hirata and first presented in TASI lecture notes on inflation by Baumann [15]. The approach of this subsection is as follows: first translational invariance will be used to prove different $k-$modes evolve independently at linear order. Then, we exploit rotational invariance to show that there is no coupling between scalar, vector and tensor perturbations at linear order.

Independent Fourier Modes

Consider the time-evolution from initial time t_1 to final time t_2 of N generic fields $\mathcal{Q}_{\mathcal{I}}$ and perturbations in these fields $\mathcal{Q}_{\mathcal{I}}(\boldsymbol{x}, t)$, where $\mathcal{I} = 1, \ldots, N$. For notational convenience, we will denote a Fourier mode here as $\delta\mathcal{Q}(\boldsymbol{k}, t)$, instead of the usual notation $\delta\mathcal{Q}_{\boldsymbol{k}}(t)$. The time-evolution $t_1 \to t_2$ of these perturbations can be written in terms of a transfer matrix $T_{\mathcal{I}\mathcal{J}}(t_2, t_1, \boldsymbol{k'}, \boldsymbol{k})$, which in principle can be computed from the linearly perturbed Einstein equations. The perturbations at time t_2 can then be related to those at initial time t_1 via the expression [15]:

$$\delta\mathcal{Q}_{\mathcal{I}}(t_2, \boldsymbol{k}) = \sum_{\mathcal{J}=1}^{N} \int d^3\boldsymbol{k'}\, T_{\mathcal{I}\mathcal{J}}(t_2, t_1, \boldsymbol{k'}, \boldsymbol{k})\, \delta\mathcal{Q}_{\mathcal{J}}(t_1, \boldsymbol{k'}). \tag{D.3.1}$$

Note that within the transfer matrix we have allowed for the possible coupling between $\boldsymbol{k}$ and $\boldsymbol{k'}$.

The proof proceeds using the principle of contradiction: based on translational invariance, we will show that the mixing of different modes as introduced in the transfer matrix is in fact forbidden. Consider the spatial translation $\boldsymbol{x} \to \boldsymbol{x'}$ expressed as:

$$\boldsymbol{x'} = \boldsymbol{x} + \boldsymbol{\xi}(\boldsymbol{x}, t). \tag{D.3.2}$$

Under such a transformation, the perturbations transform in Fourier space as:

$$\delta\mathcal{Q}'_{\mathcal{I}}(t, \boldsymbol{k}) = e^{-i\boldsymbol{k}\cdot\boldsymbol{\xi}}\, \delta\mathcal{Q}_{\mathcal{I}}(t, \boldsymbol{k}). \tag{D.3.3}$$

In the primed coordinate system, the evolution of the perturbations from t_1 to t_2 and the transfer matrix can be written as:

$$\delta\mathcal{Q}'_{\mathcal{I}}(t_2, \boldsymbol{k}) = \sum_{\mathcal{J}=1}^{N} \int d^3\boldsymbol{k'}\, e^{-i\boldsymbol{k}\cdot\boldsymbol{\xi}}\, T_{\mathcal{I}\mathcal{J}}(t_2, t_1, \boldsymbol{k'}, \boldsymbol{k})\, e^{i\boldsymbol{k'}\cdot\boldsymbol{\xi}}\, \delta\mathcal{Q}'_{\mathcal{J}}(t_1, \boldsymbol{k'}), \tag{D.3.4}$$

$$T'_{\mathcal{I}\mathcal{J}}(t_2, t_1, \boldsymbol{k'}, \boldsymbol{k}) \equiv e^{i(\boldsymbol{k'}-\boldsymbol{k})\cdot\boldsymbol{\xi}}\, T_{\mathcal{I}\mathcal{J}}(t_2, t_1, \boldsymbol{k'}, \boldsymbol{k}), \tag{D.3.5}$$

which should hold for any spatial shift $\boldsymbol{\xi}(\boldsymbol{x}, t)$.

Now, based on translational invariance of the equations of motion, we conclude that the transfer matrix should be the same in both coordinate systems:

$$T_{\mathcal{I}\mathcal{J}} = T'_{\mathcal{I}\mathcal{J}} \quad \longrightarrow \quad \boldsymbol{k} = \boldsymbol{k'} \;\; \vee \;\; T_{\mathcal{I}\mathcal{J}}(t_2, t_1, \boldsymbol{k'}, \boldsymbol{k}) = 0. \tag{D.3.6}$$

Therefore, disregarding the second solution, we find that $\boldsymbol{k} = \boldsymbol{k'}$ and hence there is no coupling between different Fourier modes at linear order. In other words, different $k-$modes evolve independently. □

Independent Evolution of SVT Components

We proceed by proving that, on account of rotational invariance of the background, the SVT components evolve independently at linear order. However, before we prove this statement, we introduce a preliminary concept first: helicity.

Consider a mode with wave-vector $\boldsymbol{k}$. Now we introduce a rotation of angle θ along $\boldsymbol{k}$. Under such a rotation, the perturbation transforms as:

$$\delta\mathcal{Q}_{\boldsymbol{k}} \to e^{im\theta}\, \delta\mathcal{Q}_{\boldsymbol{k}}, \tag{D.3.7}$$

where m is the helicity. The helicity m takes on the values $0, \pm 1, \pm 2$ for scalar, vector and tensor components, respectively. Notice that for scalar perturbations, there is no effect of the rotation on the perturbation, as it should be by definition of a scalar quantity. On account of the rotational invariance of the background, we may set $\boldsymbol{k} = (0, 0, k)$ without loss of generality, so that the spatial dependence of the perturbation contained in the complex exponential of the Fourier transform becomes $e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = e^{ikx^3}$. Rotating a spatial vector $\boldsymbol{x}$ by an angle θ around x^3 (in the Cartesian basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$) can be represented using a linear transformation with matrix $\Lambda(\theta)$ as follows:

$$\boldsymbol{x} \to \boldsymbol{x'} = \Lambda(\theta)\, \boldsymbol{x}, \qquad \Lambda(\theta) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{D.3.8}$$

In tensor notation, the above rotation can be written as $x'^i = \Lambda^i_j(\theta)\, x^j$, where $\Lambda^i_j(\theta)$ may be regarded as the spatial part ($\mu = i, \nu = j$) of the Lorentz tensor Λ^μ_ν describing rotations around the 3-axis, see e.g. chapter 1 of [82]. For later convenience, we will explicitly show how the unit vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ transform when acted on with the rotation matrix $\Lambda(\theta)$:

$$\begin{aligned} \boldsymbol{e}'_1 &= \boldsymbol{e}_1 \cos\theta - \boldsymbol{e}_2 \sin\theta, \\ \boldsymbol{e}'_2 &= \boldsymbol{e}_2 \cos\theta + \boldsymbol{e}_1 \sin\theta. \end{aligned} \tag{D.3.9}$$

Furthermore, it is convenient to move from the Cartesian basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ to the so-called helicity basis:

$$\boldsymbol{e}_\pm = \frac{\boldsymbol{e}_1 \pm i\boldsymbol{e}_2}{\sqrt{2}}, \quad \boldsymbol{e}_3. \tag{D.3.10}$$

In this basis, a rotation around the 3-axis by an angle θ will transform the unit vectors as:

$$\boldsymbol{e}'_\pm = e^{\pm i\theta}\, \boldsymbol{e}_\pm, \quad \boldsymbol{e}'_3 = \boldsymbol{e}_3. \tag{D.3.11}$$

To derive the first transformation equation, we used the explicit transformations of $\boldsymbol{e}_{1,2}$ under a rotation around the 3-axis in the Cartesian basis and Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$. The components of any contravariant tensor transform as:

$$T'_{i_1,\ldots,i_n} = e^{i(n_+ - n_-)\theta} T_{i_1,\ldots,i_n} \equiv e^{im\theta} T_{i_1,\ldots,i_n}, \tag{D.3.12}$$

in which the indices i_n can be $\pm$ or 3 in the helicity basis $\{\boldsymbol{e}_\pm, \boldsymbol{e}_3\}$ and the helicity is defined as $m \equiv n_+ - n_-$. The variables $n_\pm$ count the number of plus and minus indices present in the considered tensor.

In the helicity basis, a scalar α has no indices and therefore has helicity 0. A vector β_i has components $\beta_\pm$ and β_3, corresponding to helicities ± 1 and 0, respectively. Equivalently, the vector β is decomposed into scalar $(\beta_3, m = 0)$ and vector $(\beta_\pm, m = \pm 1)$ components. For a generic tensor γ_{ij}, the components become $\gamma_{\pm\pm}, \gamma_{\pm 3}, \gamma_{3\pm}, \gamma_{\pm\mp}$ and γ_{33}. It follows that the tensor can be decomposed into scalar $(\gamma_{33}, \gamma_{\pm\mp})$, vector $(\gamma_{\pm 3}, \gamma_{3\pm})$ and tensor $(\gamma_{\pm\pm})$ components in the helicity basis.

Helicity: Mathematical Background

The above discussion of helicity is considered sufficient for its purpose within cosmological perturbation theory. Nevertheless, the discussion lacks full mathematical rigor. Here, we discuss helicity from a more mathematically oriented viewpoint (see [44] for more details). The background, described by the unperturbed flat FRW metric, is invariant under spatial translations as well as rotations. However, going to Fourier space and picking a specific k-mode *breaks* the rotational symmetry. Invariance remains for rotations around any axis parallel to the $\boldsymbol{k}-$axis, corresponding to the symmetry group SO(2).

Any 3-tensor (e.g. β_i and γ_{ij}) can be decomposed into irreducible representations of the symmetry group SO(2). The tensors resulting from this decomposition are eigentensors (analogous to eigenvectors/eigenstates in quantum mechanics) of the helicity operator:

$$\hat{L}_\theta \equiv -i\frac{\partial}{\partial\theta}, \tag{D.3.13}$$

where θ is the angle subtended in the plane perpendicular to the wave-vector $\boldsymbol{k}$ when going from the initial to the final basis. The eigenvalues of $\hat{L}_\theta$ are equal to the helicity of the considered eigentensor.

Let us illustrate this with a generic vector $\boldsymbol{\beta}$, defined in the helicity basis $\{\boldsymbol{e}_\pm, \boldsymbol{e}_3\}$:

$$\boldsymbol{\beta} = \beta_\pm \boldsymbol{e}_\pm + \beta_3 \boldsymbol{e}_3. \tag{D.3.14}$$

Now we rotate the basis by an angle θ in the plane perpendicular to $\boldsymbol{e}_3$, to obtain a new basis $\{\boldsymbol{e}'_\pm, \boldsymbol{e}'_3\}$. The transformed vector $\boldsymbol{\beta}'$ can be related to the unprimed basis by using Eqs. D.3.11 as follows:

$$\boldsymbol{\beta}' = e^{\pm i\theta}\beta_\pm\, \boldsymbol{e}_\pm + \beta_3 \boldsymbol{e}_3 \equiv \boldsymbol{\beta}^\theta_\pm + \boldsymbol{\beta}_3, \tag{D.3.15}$$

where the upper index θ refers to the complex exponential that should be included. Acting with the helicity operator $\hat{L}_\theta$ on the two components of $\boldsymbol{\beta}'$ as defined in the last equation will yield the following:

$$\hat{L}_\theta \boldsymbol{\beta}^\theta_\pm = \pm\boldsymbol{\beta}^\theta_\pm, \quad \hat{L}_\theta \boldsymbol{\beta}_3 = 0. \tag{D.3.16}$$

Therefore, we conclude that the vector $\boldsymbol{\beta}$ can be decomposed into components with helicity $m = 0, \pm 1$ as already claimed above.

Now that the concept of helicity is introduced, we will prove that perturbations of different helicities evolve in an independent way. In other words, that one can study scalar perturbations while ignoring the presence of tensor perturbations entirely. Again, consider N perturbations $\delta\mathcal{Q}_{\mathcal{I}}$ with helicities m_{I}, defined on the helicity basis $\{\boldsymbol{e}_{\pm}, \boldsymbol{e}_3\}$. Their evolution from initial t_1 to final time t_2 can be written as [15]:

$$\delta\mathcal{Q}_{\mathcal{I}}(t_2, \boldsymbol{k}) = \sum_{\mathcal{J}=1}^{N} T_{\mathcal{I}\mathcal{J}}(t_2, t_1, \boldsymbol{k})\, \delta\mathcal{Q}_{\mathcal{J}}(t_1, \boldsymbol{k}), \tag{D.3.17}$$

where the transfer matrix $T_{\mathcal{I}\mathcal{J}}(t_2, t_1, \boldsymbol{k})$ again follows from the Einstein equations. Under $\theta-$rotations along the $\boldsymbol{k}-$axis, the perturbations transform as follows:

$$\delta\mathcal{Q}'_{\mathcal{I}}(t, \boldsymbol{k}) = e^{im_{\mathcal{I}}\theta}\delta\mathcal{Q}_{\mathcal{I}}(t, \boldsymbol{k}). \tag{D.3.18}$$

Henceforth, after the rotation the evolution equations for the perturbations become:

$$\delta\mathcal{Q}'_{\mathcal{I}}(t_2, \boldsymbol{k}) = \sum_{\mathcal{J}=1}^{N} e^{im_{\mathcal{I}}\theta}\, T_{\mathcal{I}\mathcal{J}}(t_2, t_1, \boldsymbol{k})\, e^{-im_{\mathcal{J}}\theta}\, \delta\mathcal{Q}'_{\mathcal{I}}(t, \boldsymbol{k}), \tag{D.3.19}$$

for any angle θ. The transfer matrix in both frames are related via:

$$T'_{\mathcal{I}\mathcal{J}}(t_2, t_1, \boldsymbol{k}) = e^{i(m_{\mathcal{I}}-m_{\mathcal{J}})\theta}\, T_{\mathcal{I}\mathcal{J}}(t_2, t_1, \boldsymbol{k}). \tag{D.3.20}$$

By rotational invariance of the equations of motion, $T_{\mathcal{I}\mathcal{J}} = T'_{\mathcal{I}\mathcal{J}}$ which is satisfied for:

$$m_{\mathcal{I}} = m_{\mathcal{J}} \quad \vee \quad T_{\mathcal{I}\mathcal{J}}(t_2, t_1, \boldsymbol{k}) = 0. \tag{D.3.21}$$

Disregarding the second solution, we find that $\delta\mathcal{Q}_{\mathcal{I}}$ and $\delta\mathcal{Q}_{\mathcal{J}}$ have the same helicity and hence in the linearly perturbed equations of motion modes of different helicity (i.e. scalar, vector and tensor components) do not couple and thus evolve independently. □

D.4 Specific Gauges

Here, we will introduce a few commonly used gauges for scalar perturbations (in addition to the Newtonian gauge) and we show how to choose the scalar gauge transformation to move to these gauges. Gauge choices can be motivated by considerations concerned with the perturbed metric or the perturbed energy-momentum tensor for the matter. Gauges based on both types of motivations will be introduced below.[1]

D.4.1 Uniform Curvature or Spatially Flat Gauge

This gauge is chosen such that the spatial 3-metric γ_{ij} remains unperturbed, requiring $\tilde{\Psi} = \tilde{E} = 0$. To obtain a vanishing scalar shear and spatial curvature, the scalar shifts ξ and ξ^0 should be chosen as:

$$\xi = E, \qquad \xi^0 = -\Psi/\mathcal{H}. \tag{D.4.1}$$

[1] In this section, we will consistently adopt the convention that barred quantities correspond to the zeroth order (background) contribution to a variable and the δ-notation is used for first order perturbations.

The remaining potentials $\tilde{\Phi}$ and $\tilde{B}$, which we denote as Φ_{F} and Ψ_{F} to denote that they are defined in the flat gauge, can be written in the following gauge-invariant way:

$$\Phi_{\mathrm{F}} \equiv \Phi + \Psi + (\Psi/\mathcal{H})', \tag{D.4.2}$$

$$\Psi_{\mathrm{F}} \equiv B - E' - \Psi/\mathcal{H}. \tag{D.4.3}$$

The gauge-invariance of the above potentials only applies when the spatial and temporal shifts (ξ, ξ^0) are chosen as above. In the flat gauge, the scalar field is defined as follows on account of Eq. 4.3.39:

$$\delta\phi_{\mathrm{F}} \equiv \delta\phi + \bar{\phi}\frac{\Psi}{\mathcal{H}}. \tag{D.4.4}$$

The variable $\delta\phi_{\mathrm{F}}$ is gauge-invariant by construction and often referred to as the Mukhanov-Sasaki variable [68, 75], denoted as Q or v.

D.4.2 Comoving Orthogonal Gauge

The comoving orthogonal gauge is defined by a vanishing 3-velocity: $\tilde{v}^i \equiv 0$. Orthogonality of constant time hypersurfaces to 4-velocity U^μ requires the additional constraint $\tilde{B} = 0$, so that momentum vanishes as well in this gauge [66]. The shifts required to move to this gauge are given by:

$$\xi^0 = -(v + B), \qquad \xi = -\int v \, d\tau + \mathcal{C}(\boldsymbol{x}), \tag{D.4.5}$$

where v is the velocity potential, related to the velocity perturbation as $v^i = \partial^i v$.[2] The integration constant $\mathcal{C}$ represents the residual gauge-freedom in shifts of the spatial coordinates. The remaining gauge-invariant scalar perturbations are given by:

$$\Phi_{\mathrm{C}} \equiv \Phi + \mathcal{H}(B + v) + (B + v)', \tag{D.4.6}$$

$$\Psi_{\mathrm{C}} \equiv \Psi - \mathcal{H}(B + v) \equiv \mathcal{R}, \tag{D.4.7}$$

$$E_{\mathrm{C}} \equiv E + \int v \, d\tau - \mathcal{C}. \tag{D.4.8}$$

Here, we defined the comoving curvature perturbation $\mathcal{R}$ in terms of the shift and velocity potentials, B and v. It is possible to write $\mathcal{R}$ manifestly in terms of metric perturbations instead of the velocity potential v by using the perturbed EFE's, to be derived in the upcoming sections. The resulting expression is:

$$\mathcal{R} = \Psi - \frac{\mathcal{H}(\Psi' + \mathcal{H}\Phi)}{\mathcal{H}' - \mathcal{H}^2}. \tag{D.4.9}$$

Since the comoving curvature perturbation is of such fundamental importance in this work, we will provide a geometric interpretation of $\mathcal{R}$ in the next section.

Derivation: Comoving Curvature Perturbation in Metric Peturbations

Here, we will derive the comoving curvature perturbation $\mathcal{R}$ solely in terms of metric

[2] When studying scalar perturbations, the divergence-free vector contribution $\hat{v}^i$ to the velocity need not be considered.

perturbations. One of the results of perturbing the EFE's for a perfect fluid will be the following equation relating the linear combination $v + B$ to the metric potentials Φ and Ψ [66]:

$$\Psi' + \mathcal{H}\Phi = -\frac{a^2}{2M_{\rm pl}^2}(\bar{\rho} + \bar{P})(v + B). \tag{D.4.10}$$

From the background evolution equations, we find that the combination $\bar{\rho} + \bar{P}$ equals:

$$\bar{\rho} + \bar{P} = \frac{2M_{\rm pl}^2}{a^2}(\mathcal{H}^2 - \mathcal{H}'), \tag{D.4.11}$$

which can be used to find:

$$v + B = \frac{\Psi' + \mathcal{H}\Phi}{\mathcal{H}' - \mathcal{H}^2}. \tag{D.4.12}$$

Solely in terms of metric perturbations, the comoving curvature perturbations thus becomes:

$$\mathcal{R} = \Psi - \frac{\mathcal{H}(\Psi' + \mathcal{H}\Phi)}{\mathcal{H}' - \mathcal{H}^2}. \tag{D.4.13}$$

Assuming the universe is filled by a scalar field, the comoving curvature perturbation can also be related to perturbations in the scalar field. In terms of the potentials B and v, the comoving curvature perturbation is defined as:

$$\mathcal{R} = \Psi - \mathcal{H}(B + v). \tag{D.4.14}$$

This form of $\mathcal{R}$ is gauge-invariant by construction. By comparing the mixed components of the perturbed energy-momentum tensor of a perfect fluid and a scalar field, a relation between $B + v$ and the scalar field perturbations can be found. For a perfect fluid and scalar field, the components δT_i^0 are defined as:

$$\delta T_i^0 = (\bar{\rho} + \bar{P})\partial_i(v + B), \qquad \delta T_i^0 = -\frac{1}{a^2}\bar{\phi}'\partial_i\delta\phi, \tag{D.4.15}$$

respectively. These relation will be derived in section 4.8. Using the fact that for scalar fields $\bar{\rho} + \bar{P} = \bar{\phi}'^2/a^2$ and equating the above two expressions yields:

$$v + B = -\frac{\delta\phi}{\bar{\phi}'}. \tag{D.4.16}$$

The comoving curvature in a scalar field dominated universe can therefore be written as:

$$\mathcal{R} = \Psi + \mathcal{H}\frac{\delta\phi}{\bar{\phi}'}. \tag{D.4.17}$$

This direct relation between $\mathcal{R}$, Ψ and $\delta\phi$ will allow us to connect inflaton fluctuations to the comoving curvature perturbation.

D.4.3 Uniform Density Gauge

In this gauge, the spatial hypersurfaces are chosen such that the energy density on these hypersurfaces is homogeneous, i.e. the density perturbation is zero $\delta\rho = 0$. According to Eq.

4.3.39, the temporal shift ξ^0 should be chosen as:

$$\xi^0 = \frac{\delta\rho}{\bar{\rho}'}, \tag{D.4.18}$$

to obtain $\delta\tilde{\rho} = 0$ and hence a uniform energy density on the spatial slices. The spatial curvature perturbation Ψ then transforms as:

$$\Psi_{\mathrm{UD}} \equiv \Psi + \mathcal{H}\frac{\delta\rho}{\bar{\rho}'} \equiv \zeta. \tag{D.4.19}$$

Here we defined the curvature perturbation on slices of uniform energy density as ζ, following the sign convention of Riotto [74]. This quantity is gauge-invariant by construction and represents the curvature perturbation Ψ on slices of uniform energy density:

$$\zeta = \Psi\big|_{\delta\rho=0}. \tag{D.4.20}$$

Notice that the spatial shift ξ is unused in the definition of the uniform density gauge. To completely specify the gauge, one can choose one of the remaining scalar perturbations (Φ, B or E) to vanish by appropriate choice of ξ.

Appendix E

Evolution Outside the Horizon

E.1 Derivation of $\mathcal{R}'$ in Field Equations Approach

To derive Eq. 6.2.4 for the evolution of $\mathcal{R}$, we start from Eq. 6.2.3 and use the following background relations for the equation of state and the sound speed $c_s^2 = P'/\rho'$ (Eqs. 4.1.14):

$$\frac{\mathcal{H}'}{\mathcal{H}^2} = -\frac{1}{2}(1+3w), \qquad \frac{w'}{w+1} = 3\mathcal{H}(w - c_s^2), \tag{E.1.1}$$

Using those, we can rewrite the above expression for $\mathcal{R}'$ as:

$$\frac{3}{2}(1+w)\mathcal{H}^{-1}\mathcal{R}' = 3c_s^2(\mathcal{H}^{-1}\Psi' + \Phi) + \mathcal{H}^{-2}\Psi'' + 2\mathcal{H}^{-1}\Psi' + \mathcal{H}^{-1}\Phi' - 3w\Phi. \tag{E.1.2}$$

Now we are in the position to rewrite the above equation for the evolution of $\mathcal{R}$ by invoking the field equations in Newtonian gauge. We can rewrite Eq. 4.9.18 by invoking the first Friedmann equation as:

$$\mathcal{H}^{-2}\Psi'' + \mathcal{H}^{-1}(\Phi' + 2\Psi') + \left(1 + \frac{2\mathcal{H}'}{\mathcal{H}^2}\right)\Phi + \frac{1}{2}\partial^2(\Phi - \Psi) = \frac{3\delta P}{2\rho}, \tag{E.1.3}$$

where we recognize the prefactor of the third term as $-3w$ according to background relation for $\mathcal{H}'/\mathcal{H}^2$. In Fourier space, the above equation becomes:

$$\mathcal{H}^{-2}\Psi'' + \mathcal{H}^{-1}(\Phi' + 2\Psi') - 3w\Phi = \frac{3\delta P}{2\rho} + \frac{1}{2}\left(\frac{k}{\mathcal{H}}\right)^2(\Phi - \Psi). \tag{E.1.4}$$

Hence, by substitution of this result, we can rewrite Eq. E.1.2 to give:

$$\frac{3}{2}(1+w)\mathcal{H}^{-1}\mathcal{R}' = 3c_s^2(\mathcal{H}^{-1}\Psi' + \Phi) + \frac{3\delta P}{2\rho} + \frac{1}{2}\left(\frac{k}{\mathcal{H}}\right)^2(\Phi - \Psi). \tag{E.1.5}$$

Now invoking the first field equation in Newtonian gauge (Eq. 4.9.16) to obtain an expression for $\mathcal{H}^{-1}\Psi' + \Phi$ yields:

$$\frac{\Psi'}{\mathcal{H}} + \Phi = -\frac{\delta}{2} - \frac{1}{3}\left(\frac{k}{\mathcal{H}}\right)^2\Psi, \tag{E.1.6}$$

where δ denotes the density contrast: $\delta \equiv \delta\rho/\rho$. Inserting this expression gives:

$$\frac{3}{2}(1+w)\mathcal{H}^{-1}\mathcal{R}' = -c_s^2\left(\frac{k}{\mathcal{H}}\right)^2\Psi + \frac{1}{2}\left(\frac{k}{\mathcal{H}}\right)^2(\Phi - \Psi) + \frac{3}{2}\left(\frac{\delta P}{\rho} - c_s^2\delta\right). \tag{E.1.7}$$

Lastly, we use Eq. 4.6.13, valid in any gauge, to compute the ratio $\delta P/\rho$, which gives:

$$\frac{\delta P}{\rho} = c_s^2(\delta - 3(1+w)\mathcal{S}), \tag{E.1.8}$$

and hence we get:

$$\frac{3}{2}(1+w)\mathcal{H}^{-1}\mathcal{R}' = -c_s^2\left(\frac{k}{\mathcal{H}}\right)^2\Psi + \frac{1}{2}\left(\frac{k}{\mathcal{H}}\right)^2(\Phi - \Psi) - \frac{9}{2}(1+w)c_s^2\mathcal{S}. \tag{E.1.9}$$

Finally, we obtain the following equation for the time evolution of $\mathcal{R}$ in Fourier space:

$$\mathcal{R}' = -\frac{2\mathcal{H}}{3(1+w)}\left(\frac{k}{\mathcal{H}}\right)^2\left[c_s^2\Psi + \frac{1}{2}(\Psi - \Phi)\right] - 3\mathcal{H}c_s^2\mathcal{S}. \tag{E.1.10}$$

E.2 Derivation of $\mathcal{R}'$ in Energy-Momentum Approach

Here, we will discuss how Eq. 6.3.5, which is a core result in proving the constancy of $\mathcal{R}$ in the energy-momentum approach, is derived. In the super-horizon limit, the scalar component of the velocity perturbation $\partial_i v$ vanishes (because in Fourier space $\partial_i \to ik_i$, which vanishes as $k \to 0$). Therefore, we can set $T_0^i = 0$. Furthermore, we can simplify the (scalar) perturbed metric tensor (Eq. 4.3.4) by applying the super-horizon limit:

$$\begin{aligned} g_{\mu\nu} &= a^2\begin{pmatrix} -(1+2\Phi) & \partial_i B \\ \partial_i B & (1-2\Psi)\delta_{ij} + 2\partial_{\langle i}\partial_{j\rangle}E \end{pmatrix} \\ &\xrightarrow{k\to 0} a^2\begin{pmatrix} -(1+2\Phi) & 0 \\ 0 & (1-2\Psi)\delta_{ij} \end{pmatrix}. \end{aligned} \tag{E.2.1}$$

Notice that the super-horizon limit constrains the perturbed metric in any gauge to become equivalent to the metric in Newtonian gauge, since all terms involving gradients or derivatives vanish. The super-horizon form of the metric also affects the expressions for the perturbed Christoffel symbols.

In terms of Christoffel symbols and ordinary partial derivatives, the covariant derivative equation enforcing energy-momentum conservation can be written as:

$$\nabla_\mu T^\mu_\nu = \partial_\mu T^\mu_\nu + \Gamma^\mu_{\mu\lambda}T^\lambda_\nu - \Gamma^\lambda_{\mu\nu}T^\mu_\lambda = 0. \tag{E.2.2}$$

Perturbing this equation to first order to get an equation for the conservation of the perturbed part the energy-momentum tensor yields:

$$\nabla_\mu \delta T^\mu_\nu = \partial_\mu \delta T^\mu_\nu + \delta\Gamma^\mu_{\mu\lambda}T^\lambda_\nu + \Gamma^\mu_{\mu\lambda}\delta T^\lambda_\nu - \delta\Gamma^\lambda_{\mu\nu}T^\mu_\lambda - \Gamma^\lambda_{\mu\nu}\delta T^\mu_\lambda = 0. \tag{E.2.3}$$

To prove the super-horizon constancy of the comoving curvature perturbation $\mathcal{R}$, it suffices to consider the $\nu = 0$ component only, which gives:

$$\nabla_\mu \delta T^\mu_0 = \partial_\mu \delta T^\mu_0 + \delta\Gamma^\mu_{\mu\lambda}T^\lambda_0 + \Gamma^\mu_{\mu\lambda}\delta T^\lambda_0 - \delta\Gamma^\lambda_{\mu 0}T^\mu_\lambda - \Gamma^\lambda_{\mu 0}\delta T^\mu_\lambda = 0. \tag{E.2.4}$$

Evaluating of all terms yields the following equation for the time derivative of the energy density in terms of the energy density, pressure and gravitational potentials:

$$\delta\rho' = 3\Psi'(\rho + P) - 3\mathcal{H}(\delta\rho + \delta P), \tag{E.2.5}$$

which verifies the advocated result (Eq. 6.3.5). In the derivation box below we reflect briefly on the way the above equation is derived.

Partial Derivation: Conservation of Perturbed EM Tensor

To show explicitly how the above equation is derived, we show how to evaluate the term:

$$\nabla_\mu \delta T_0^\mu \supset \delta\Gamma^\mu_{\mu\lambda} T_0^\lambda + \Gamma^\mu_{\mu\lambda}\delta T_0^\lambda, \tag{E.2.6}$$

in the expression for $\nabla_\mu \delta T_0^\mu$. We first consider the first subterm:

$$\delta\Gamma^\mu_{\mu\lambda} T_0^\lambda = \delta\Gamma^\mu_{\mu 0} T_0^0 + \delta\Gamma^\mu_{\mu i} T_0^i = -\delta\Gamma^\mu_{\mu 0}\rho = -(\Gamma^0_{00} + \delta\Gamma^i_{i0})\rho = -\rho(\Phi' - 3\Psi'), \tag{E.2.7}$$

where we used that in the super-horizon limit $T_0^i = 0$. Furthermore, the perturbed Christoffel symbols are given by $\delta\Gamma^0_{00} = \Phi'$ and:

$$\delta\Gamma^i_{i0} = -\Psi'\delta^i_i + \frac{1}{2}\partial_{\langle i}\partial^{i\rangle} h' \quad \xrightarrow{k\to 0} \quad -3\Psi', \tag{E.2.8}$$

since in the super-horizon limit the derivative term proportional to $\partial_{\langle i}\partial^{i\rangle}$ vanishes. This gives the last equality. Similarly, the other terms in the expression for $\nabla_\mu \delta T_0^\mu$ can be evaluated as well, yielding:

$$\nabla_\mu \delta T^\mu_\nu = -\delta\rho' - \rho(\Phi' - 3\Psi') - 4\mathcal{H}\delta\rho + \rho\Phi' + 3\Psi' P + \mathcal{H}\delta\rho - 3\mathcal{H}\delta P. \tag{E.2.9}$$

Collecting terms and rewriting this equation yields the desired equation for $\delta\rho'$.

E.3 Lie Derivative of the Metric

Here we will derive the Lie derivative of the metric. That is, we will prove the relation:

$$\Delta(\delta g_{\mu\nu})(x) = -\bar{g}_{\mu\kappa}\partial_\nu\xi^\kappa - \bar{g}_{\kappa\nu}\partial_\mu\xi^\kappa - \partial_\kappa\bar{g}_{\mu\nu}\xi^\kappa \equiv \Delta_\xi\, g_{\mu\nu}, \tag{E.3.1}$$

introduced in the discussion on gauge transformations in section 6.4.2. We will start from Eq. 6.4.8:

$$\Delta(\delta g_{\mu\nu}) \equiv \tilde{g}_{\mu\nu}(x) - g_{\mu\nu}(x). \tag{E.3.2}$$

We can use the first order relation $x^\mu = \tilde{x}^\mu - \xi^\mu$ to write:

$$\Delta(\delta g_{\mu\nu}) \equiv \tilde{g}_{\mu\nu}(\tilde{x} - \xi) - g_{\mu\nu}(x) = \tilde{g}_{\mu\nu}(\tilde{x}) - \partial_\lambda\bar{g}_{\mu\nu}(x)\xi^\lambda(x) - g_{\mu\nu}(x). \tag{E.3.3}$$

At first order, it does not matter whether the partial derivative ∂_λ is taken with respect to x or $\tilde{x}$ (here we take x) and we can take the background metric since ξ^λ is already a first order perturbation. By the tensor transformation law for the metric, we can rewrite the first term in the expansion so that we get:

$$\begin{aligned}
\Delta(\delta g_{\mu\nu}) &= g_{\lambda\kappa}\frac{\partial x^\lambda}{\partial\tilde{x}^\mu}\frac{\partial x^\kappa}{\partial\tilde{x}^\nu} - g_{\mu\nu}(x) - \partial_\lambda\bar{g}_{\mu\nu}\xi^\lambda(x) \\
&= g_{\lambda\kappa}\frac{\partial(\tilde{x}^\lambda - \xi^\lambda)}{\partial\tilde{x}^\mu}\frac{\partial(\tilde{x}^\kappa - \xi^\kappa)}{\partial\tilde{x}^\nu} - g_{\mu\nu}(x) - \partial_\lambda\bar{g}_{\mu\nu}\xi^\lambda(x) \\
&= g_{\lambda\kappa}(\delta^\lambda_\mu - \partial_\mu\xi^\lambda)(\delta^\kappa_\nu - \partial_\nu\xi^\kappa) - g_{\mu\nu}(x) - \partial_\lambda\bar{g}_{\mu\nu}\xi^\lambda(x) \\
&= -\bar{g}_{\mu\kappa}\partial_\nu\xi^\kappa - \bar{g}_{\kappa\nu}\partial_\mu\xi^\kappa - \partial_\kappa\bar{g}_{\mu\nu}\xi^\kappa \equiv \Delta_\xi\, g_{\mu\nu}.
\end{aligned} \tag{E.3.4}$$

This is indeed the advocated result, the Lie derivative of the metric $\Delta_\xi\, g_{\mu\nu}$, used in the main text (Eq. 6.4.9)

Appendix F

Non-Gaussianity and Local Bispectrum

F.1 Derivation of Correlation Function $\langle \Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\Phi_{\mathsf{NL}}(\boldsymbol{k}_3)\rangle$

To derive the three-point correlation function for the local parametrization of NG, we will consider the first term in Eq. 7.4.3 – the remaining two permutations will be straightforward knowing the result of the first term. Expanding the first term yields:

$$\begin{aligned}\langle \Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\Phi_{\rm NL}(\boldsymbol{k}_3)\rangle &= f_{\rm NL}\left\langle \Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\int_{\boldsymbol{p}}\Phi(\boldsymbol{k}_3+\boldsymbol{p})\Phi^*(\boldsymbol{p}) - (2\pi)^3\delta^{(3)}(\boldsymbol{k}_3)\langle\Phi(\boldsymbol{x})^2\rangle\right\rangle \\ &= f_{\rm NL}\int_{\boldsymbol{p}}\langle\Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\Phi(\boldsymbol{k}_3+\boldsymbol{p})\Phi^*(\boldsymbol{p})\rangle \qquad (*) \\ &\quad - f_{\rm NL}(2\pi)^3\delta^{(3)}(\boldsymbol{k}_3)\langle\Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\rangle\langle\Phi(\boldsymbol{x})^2\rangle. \qquad (**) \end{aligned} \tag{F.1.1}$$

To evaluate the 4-point correlation function in the second line we use Wick's theorem, which states for a generic field $f(\boldsymbol{k})$ that [62]:

$$\langle f(\boldsymbol{k}_1)f(\boldsymbol{k}_2)f(\boldsymbol{k}_3)f(\boldsymbol{k}_4)\rangle = \sum \text{All possible 2-point contractions.} \tag{F.1.2}$$

Using this contraction identity, the 4-point correlation function can be written in the following way:

$$\begin{aligned}\int_{\boldsymbol{p}}\langle\Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\Phi(\boldsymbol{k}_3+\boldsymbol{p})\Phi^*(\boldsymbol{p})\rangle &= \int_{\boldsymbol{p}}\langle\Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\rangle\langle\Phi(\boldsymbol{k}_3+\boldsymbol{p})\Phi^*(\boldsymbol{p})\rangle \qquad \text{(A)} \\ &\quad + \int_{\boldsymbol{p}}\langle\Phi(\boldsymbol{k}_1)\Phi^*(\boldsymbol{p})\rangle\langle\Phi(\boldsymbol{k}_2+\boldsymbol{p})\Phi(\boldsymbol{k}_3)\rangle \qquad \text{(B)} \\ &\quad + \int_{\boldsymbol{p}}\langle\Phi(\boldsymbol{k}_2)\Phi^*(\boldsymbol{p})\rangle\langle\Phi(\boldsymbol{k}_3+\boldsymbol{p})\Phi^*(\boldsymbol{k}_1)\rangle \qquad \text{(C)}\end{aligned} \tag{F.1.3}$$

To proceed, we first combine the second term in Eq. F.1.1 with term (A) in the following way. Using the fact that $\Phi^*(\boldsymbol{p}) = \Phi(-\boldsymbol{p})$, we find:

$$\text{(A)} + (**) = f_{\rm NL}\langle\Phi(\boldsymbol{k}_1)\Phi(\boldsymbol{k}_2)\rangle \times (2\pi)^3\delta^{(3)}(\boldsymbol{k}_3)\left[\int_{\boldsymbol{p}}P_\Phi(\boldsymbol{p}) - \int_{\boldsymbol{p}}P_\Phi(\boldsymbol{p})\right] = 0. \tag{F.1.4}$$

Here we employed the following definitions:

$$\langle\Phi(\boldsymbol{k}_3+\boldsymbol{p})\Phi^*(\boldsymbol{p})\rangle = \langle\Phi(\boldsymbol{k}_3+\boldsymbol{p})\Phi(-\boldsymbol{p})\rangle = (2\pi)^3\delta^{(3)}(\boldsymbol{k}_3)\times P_\Phi(\boldsymbol{p}), \tag{F.1.5}$$

$$\langle\Phi(\boldsymbol{x})^2\rangle = \int_{\boldsymbol{p}}P_\Phi(\boldsymbol{p}). \tag{F.1.6}$$

Now, we expand the 2-point correlation functions in the second term (B):

$$\begin{aligned}(\mathrm{B}) &= \int_{\boldsymbol{p}} (2\pi)^2 \delta^{(3)}(\boldsymbol{k}_1 - \boldsymbol{p}) P_\Phi(k_1) \times (2\pi)^3 \delta^{(3)}(\boldsymbol{k}_2 + \boldsymbol{k}_3 + \boldsymbol{p}) P_\Phi(k_2) \\ &= (2\pi)^3 \delta^{(3)}(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) P_\Phi(k_1) P_\Phi(k_2), \end{aligned} \tag{F.1.7}$$

where the integral is evaluated using the first delta function. Lastly, we consider term (C) and expand the 2-point correlation functions in terms of $P(k_i)$ as follows:

$$\begin{aligned}(\mathrm{C}) &= \int_{\boldsymbol{p}} \langle \Phi(\boldsymbol{k}_2) \Phi^*(\boldsymbol{p}) \rangle \langle \Phi(\boldsymbol{k}_3 + \boldsymbol{p}) \Phi^*(\boldsymbol{k}_1) \rangle \\ &= \int_{\boldsymbol{p}} (2\pi)^3 \delta^{(3)}(\boldsymbol{k}_2 - \boldsymbol{p}) P_\Phi(k_2) \times (2\pi)^3 \delta^{(3)}(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{p}) P_\Phi(k_1) \\ &= (2\pi)^3 \delta^{(3)}(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) P_\Phi(k_1) \Phi(k_2). \end{aligned} \tag{F.1.8}$$

Therefore, we find the following expression for the correlation function:

$$\langle \Phi(\boldsymbol{k}_1) \Phi(\boldsymbol{k}_2) \Phi_{\mathrm{NG}}(\boldsymbol{k}_3) \rangle = 2 f_{\mathrm{NL}} \, (2\pi)^3 \delta^{(3)}(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3) \, P_\Phi(k_1) P_\Phi(k_2). \tag{F.1.9}$$

F.2 Fourier Transform of Correlation Function

To find the analog of the real space 3-point correlation function, we start from Eq. 7.5.14:

$$\langle \overline{\mathcal{R}}(\boldsymbol{x}_1) \mathcal{R}(\boldsymbol{x}_2) \mathcal{R}(\boldsymbol{x}_3) \rangle = \langle \overline{\mathcal{R}}(\boldsymbol{x}_1) \overline{\mathcal{R}}(\boldsymbol{x}_+) \rangle \Big[(\boldsymbol{x}_3 - \boldsymbol{x}_2) \cdot \boldsymbol{\nabla} \langle \mathcal{R}(\boldsymbol{x}_2) \mathcal{R}(\boldsymbol{x}_3) \rangle \Big] \Big|_0 . \tag{F.2.1}$$

The Fourier transform of the two terms can be written as:

$$(*) = \langle \overline{\mathcal{R}}(\boldsymbol{x}_1) \overline{\mathcal{R}}(\boldsymbol{x}_+) \rangle = \int_{\boldsymbol{k}_L} P(k_L) \, e^{i \boldsymbol{k}_L \cdot (\boldsymbol{x}_1 - \boldsymbol{x}_+)}, \tag{F.2.2}$$

$$(**) = (\boldsymbol{x}_3 - \boldsymbol{x}_2) \cdot \boldsymbol{\nabla} \langle \mathcal{R}(\boldsymbol{x}_2) \mathcal{R}(\boldsymbol{x}_3) \rangle = \int_{\boldsymbol{k}_S} P(k_S) \, (\boldsymbol{k}_S \cdot \partial_{\boldsymbol{k}_S}) \, e^{i \boldsymbol{k}_S \cdot \boldsymbol{x}_-}, \tag{F.2.3}$$

where $\boldsymbol{x}_- \equiv \boldsymbol{x}_3 - \boldsymbol{x}_2$. As the momenta serve as dummy integration variables, they can be chosen arbitrarely and for later convenience they are set to $\boldsymbol{k}_L$ and $\boldsymbol{k}_S$ for the terms $(*)$ and $(**)$, respectively. Performing intergration by parts once on $(**)$ gives:

$$(**) = - \int_{\boldsymbol{k}_S} \partial_{\boldsymbol{k}_S} \cdot \Big[\boldsymbol{k}_S P(k_S) \Big] \, e^{i \boldsymbol{k}_S \cdot \boldsymbol{x}_-} = - \int_{\boldsymbol{k}_S} P(k_S) \, \frac{d \ln(k_S^3 P(k_S))}{d \ln k_S} \, e^{i \boldsymbol{k}_S \cdot \boldsymbol{x}_-}. \tag{F.2.4}$$

In the second equality we used the fact that the derivative can be expressed in terms the momentum magnitude k_S as:

$$\partial_{\boldsymbol{k}_S} \cdot \Big[\boldsymbol{k}_S \cdot P(k_S) \Big] = P(k_S) \, \frac{d \ln(k_S^3 P(k_S)}{d \ln k_S}. \tag{F.2.5}$$

Based on momentum conservation, i.e. the momentum vectors $\boldsymbol{k}_i$ must form a closed triangle, we can always multiply by a conveniently chosen unity:

$$1 = \int_{\boldsymbol{k}_1} (2\pi)^3 \delta^{(3)}(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3), \tag{F.2.6}$$

and the product $(*)(**)$ can be written as:

$$\langle\overline{\mathcal{R}}(\boldsymbol{x}_1)\mathcal{R}(\boldsymbol{x}_2)\mathcal{R}(\boldsymbol{x}_3)\rangle = -\int_{\boldsymbol{k}_1}\int_{\boldsymbol{k}_L}\int_{\boldsymbol{k}_S} e^{i\boldsymbol{k}_L\cdot(\boldsymbol{x}_1-\boldsymbol{x}_+)+i\boldsymbol{k}_S\cdot\boldsymbol{x}_-} \times (2\pi)^2\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)P(k_L)P(k_S)\,\frac{d\ln(k_S^3P(k_S)}{d\ln k_S} \tag{F.2.7}$$

The above expression can be simplified as follows. First, note that on account of momentum conservation $\boldsymbol{k}_L = -\boldsymbol{k}_1$. The argument of the complex exponential can then be written as:

$$-\boldsymbol{k}_1\cdot\boldsymbol{x}_1 - \boldsymbol{k}_L\cdot\boldsymbol{x}_+ + \boldsymbol{k}_S\cdot\boldsymbol{x}_- = -(\boldsymbol{k}_1\cdot\boldsymbol{x}_1+\boldsymbol{k}_2\cdot\boldsymbol{x}_2+\boldsymbol{k}_3\cdot\boldsymbol{x}_3). \tag{F.2.8}$$

Changing the dummy integration variables as $\boldsymbol{k}_S \to \boldsymbol{k}_2$ and $\boldsymbol{k}_L \to \boldsymbol{k}_3$, the expression becomes:

$$\langle\overline{\mathcal{R}}(\boldsymbol{x}_1)\mathcal{R}(\boldsymbol{x}_2)\mathcal{R}(\boldsymbol{x}_3)\rangle = -\int_{\boldsymbol{k}_1}\int_{\boldsymbol{k}_2}\int_{\boldsymbol{k}_3} e^{-i(\boldsymbol{k}_1\cdot\boldsymbol{x}_1+\boldsymbol{k}_2\cdot\boldsymbol{x}_2+\boldsymbol{k}_3\cdot\boldsymbol{x}_3)} \times (2\pi)^3\delta^{(3)}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)P(k_1)P(k_2)\,\frac{d\ln(k_2^3P(k_2)}{d\ln k_2}, \tag{F.2.9}$$

where we used the fact that in the squeezed limit the $\boldsymbol{k}_S = \boldsymbol{k}_3 = -\boldsymbol{k}_2$ and hence $k_S = k_2$. From this result, the Fourier counterpart can be extracted directly to give Eq. 7.5.15.

Bibliography

[1] M. Abramowitz and I. A. Stegun: *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Vol. 55. Courier Corporation, 1964

[2] P. A. R. Ade et al.: "Planck 2013 Results. XXIV. Constraints on Primordial Non-Gaussianity". *Astron. Astrophys.* 571, A24 (2014). DOI: 10.1051/0004-6361/201321554. arXiv: 1303.5084 [astro-ph.CO]

[3] P. A. R. Ade et al.: "Planck 2015 results. XX. Constraints on inflation". *Astron. Astrophys.* 594, A20 (2016). DOI: 10.1051/0004-6361/201525898. arXiv: 1502.02114 [astro-ph.CO]

[4] P. Adshead, R. Easther, and E. A. Lim: "Cosmology With Many Light Scalar Fields: Stochastic Inflation and Loop Corrections". *Phys. Rev.* D79, 063504 (2009). DOI: 10.1103/PhysRevD.79.063504. arXiv: 0809.4008 [hep-th]

[5] P. Adshead, R. Easther, and E. A. Lim: "Cosmology With Many Light Scalar Fields: Stochastic Inflation and Loop Corrections". *Phys. Rev.* D79, 063504 (2009). DOI: 10.1103/PhysRevD.79.063504. arXiv: 0809.4008 [hep-th]

[6] M. Akhshik, H. Firouzjahi, and S. Jazayeri: "Cosmological Perturbations and the Weinberg Theorem". *JCAP* 1512.12, 027 (2015). DOI: 10.1088/1475-7516/2015/12/027. arXiv: 1508.03293 [hep-th]

[7] N. Arkani-Hamed and J. Maldacena: "Cosmological Collider Physics" (2015). arXiv: 1503.08043 [hep-th]

[8] R. L. Arnowitt, S. Deser, and C. W. Misner: "The Dynamics of general relativity". *Gen. Rel. Grav.* 40, 1997–2027 (2008). DOI: 10.1007/s10714-008-0661-1. arXiv: gr-qc/0405109 [gr-qc]

[9] F. Arroja and T. Tanaka: "A note on the role of the boundary terms for the non-Gaussianity in general k-inflation". *JCAP* 1105, 005 (2011). DOI: 10.1088/1475-7516/2011/05/005. arXiv: 1103.1102 [astro-ph.CO]

[10] V. Assassi, D. Baumann, and D. Green: "Symmetries and Loops in Inflation". *JHEP* 02, 151 (2013). DOI: 10.1007/JHEP02(2013)151. arXiv: 1210.7792 [hep-th]

[11] D. Babich, P. Creminelli, and M. Zaldarriaga: "The Shape of non-Gaussianities". *JCAP* 0408, 009 (2004). DOI: 10.1088/1475-7516/2004/08/009. arXiv: astro-ph/0405356 [astro-ph]

[12] J. M. Bardeen: "Gauge-invariant Cosmological Perturbations". *Physical Review D* 22.8, 1882 (1980)

[13] N. Bartolo, S. Matarrese, and A. Riotto: "Gauge-invariant temperature anisotropies and primordial non-Gaussianity". *Phys. Rev. Lett.* 93, 231301 (2004). DOI: 10.1103/PhysRevLett.93.231301. arXiv: astro-ph/0407505 [astro-ph]

[14] D. Baumann: "Cosmology, Mathematical Tripos, Part III, DAMPT Cambridge" (2010). URL: http://www.damtp.cam.ac.uk/user/db275/cosmology.pdf

[15] D. Baumann: "Inflation". *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.* 2011, 523–686. DOI: 10.1142/9789814327183_0010. arXiv: 0907.5424 [hep-th]

[16] D. Baumann: "Primordial Non-Gaussianity, DAMPT Cambridge" (2011). URL: http://www.damtp.cam.ac.uk/user/db275/TEACHING/INFLATION/NG.pdf

[17] D. Baumann and L. McAllister: *Inflation and String Theory.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2015. ISBN: 9781107089693, 9781316237182. DOI: 10.1017/CBO9781316105733. arXiv: 1404.2601 [hep-th]. URL: https://inspirehep.net/record/1289899/files/arXiv:1404.2601.pdf

[18] D. Baumann and L. McAllister: *Inflation and String Theory.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2015. ISBN: 9781107089693, 9781316237182. DOI: 10.1017/CBO9781316105733. arXiv: 1404.2601 [hep-th]. URL: http://www.cambridge.org/mw/academic/subjects/physics/theoretical-physics-and-mathematical-physics/inflation-and-string-theory?format=HB

[19] D. Baumann et al.: "Partially Massless Fields During Inflation". *JHEP* 04, 140 (2018). DOI: 10.1007/JHEP04(2018)140. arXiv: 1712.06624 [hep-th]

[20] P. Berglund and G. Ren: "Multi-Field Inflation from String Theory" (2009). arXiv: 0912.1397 [hep-th]

[21] E. Bertschinger: "Cosmological Perturbation Theory and Structure Formation". *Cosmology 2000: Proceedings, Conference, Lisbon, Portugal, 12-15 Jul 2000.* 2001, 1–25. arXiv: astro-ph/0101009 [astro-ph]

[22] D. Boer: "Symmetry in Physics, Lecture Notes". *University of Groningen* (2015)

[23] L. A. Boyle and P. J. Steinhardt: "Probing the Early Universe with Inflationary Gravitational Waves". *Phys. Rev.* D77, 063504 (2008). DOI: 10.1103/PhysRevD.77.063504. arXiv: astro-ph/0512014 [astro-ph]

[24] R. H. Brandenberger: "Alternatives to the inflationary paradigm of structure formation". *Int. J. Mod. Phys. Conf. Ser.* 01, 67–79 (2011). DOI: 10.1142/S2010194511000109. arXiv: 0902.4731 [hep-th]

[25] R. H. Brandenberger: "Lectures on the Theory of Cosmological Perturbations". *Lect. Notes Phys.* 646, 127–167 (2004). DOI: 10.1007/978-3-540-40918-2_5. arXiv: hep-th/0306071 [hep-th]

[26] C. P. Burgess: "Quantum gravity in everyday life: General relativity as an effective field theory". *Living Rev. Rel.* 7, 5–56 (2004). DOI: 10.12942/lrr-2004-5. arXiv: gr-qc/0311082 [gr-qc]

[27] C. Burrage, R. H. Ribeiro, and D. Seery: "Large slow-roll corrections to the bispectrum of noncanonical inflation". *JCAP* 1107, 032 (2011). DOI: 10.1088/1475-7516/2011/07/032. arXiv: 1103.4126 [astro-ph.CO]

[28] C. T. Byrnes and K.-Y. Choi: "Review of Local Non-Gaussianity from Multi-Field inflation". *Adv. Astron.* 2010, 724525 (2010). DOI: 10.1155/2010/724525. arXiv: 1002.3110 [astro-ph.CO]

[29] C. T. Byrnes, K.-Y. Choi, and L. M. H. Hall: "Conditions for large non-Gaussianity in two-field slow-roll inflation". *JCAP* 0810, 008 (2008). DOI: 10.1088/1475-7516/2008/10/008. arXiv: 0807.1101 [astro-ph]

[30] D. Campo: "Quantum corrections during inflation and conservation of adiabatic perturbations". *Phys. Rev.* D81, 043535 (2010). DOI: 10.1103/PhysRevD.81.043535. arXiv: 0908.3642 [hep-th]

[31] S. M. Carroll: "Lecture Notes on General Relativity" (1997). arXiv: gr-qc/9712019 [gr-qc]

[32] X. Chen: "Primordial Non-Gaussianities from Inflation Models". *Adv. Astron.* 2010, 638979 (2010). DOI: 10.1155/2010/638979. arXiv: 1002.1416 [astro-ph.CO]

[33] X. Chen: "Running Non-Gaussianities in DBI inflation". *Phys. Rev.* D72, 123518 (2005). DOI: 10.1103/PhysRevD.72.123518. arXiv: astro-ph/0507053 [astro-ph]

[34] X. Chen, R. Easther, and E. A. Lim: "Large Non-Gaussianities in Single Field Inflation". *JCAP* 0706, 023 (2007). DOI: 10.1088/1475-7516/2007/06/023. arXiv: astro-ph/0611645 [astro-ph]

[35] X. Chen and Y. Wang: "Quasi-Single Field Inflation and Non-Gaussianities". *JCAP* 1004, 027 (2010). DOI: 10.1088/1475-7516/2010/04/027. arXiv: 0911.3380 [hep-th]

[36] X. Chen et al.: "Observational signatures and non-Gaussianities of general single field inflation". *JCAP* 0701, 002 (2007). DOI: 10.1088/1475-7516/2007/01/002. arXiv: hep-th/0605045 [hep-th]

[37] C. Cheung et al.: "On the Consistency Relation of the 3-point Function in Single Field Inflation". *JCAP* 0802, 021 (2008). DOI: 10.1088/1475-7516/2008/02/021. arXiv: 0709.0295 [hep-th]

[38] C. Cheung et al.: "The Effective Field Theory of Inflation". *JHEP* 03, 014 (2008). DOI: 10.1088/1126-6708/2008/03/014. arXiv: 0709.0293 [hep-th]

[39] M. Colless et al.: "The 2dF Galaxy Redshift Survey: spectra and redshifts". *Monthly Notices of the Royal Astronomical Society* 328.4, 1039–1063 (2001). DOI: 10.1046/j.1365-8711.2001.04902.x. URL: http://dx.doi.org/10.1046/j.1365-8711.2001.04902.x

[40] P. Creminelli et al.: "The (not so) Squeezed Limit of the Primordial 3-point Function". *JCAP* 1111, 038 (2011). DOI: 10.1088/1475-7516/2011/11/038. arXiv: 1106.1462 [astro-ph.CO]

[41] A. Gabrielli et al.: *Statistical Physics for Cosmic Structures*. Springer Science & Business Media, 2006

[42] J. Garcia-Bellido and D. Wands: "Constraints from inflation on scalar - tensor gravity theories". *Phys. Rev.* D52, 6739–6749 (1995). DOI: 10.1103/PhysRevD.52.6739. arXiv: gr-qc/9506050 [gr-qc]

[43] J. Garcia-Bellido and D. Wands: "Metric perturbations in two field inflation". *Phys. Rev.* D53, 5437–5445 (1996). DOI: 10.1103/PhysRevD.53.5437. arXiv: astro-ph/9511029 [astro-ph]

[44] D. S. Gorbunov and V. A. Rubakov: *Introduction to the Theory of the Early Universe: Cosmological Perturbations and Inflationary Theory*. World Scientific, 2011

[45] C. Gordon et al.: "Adiabatic and entropy perturbations from inflation". *Phys. Rev.* D63, 023506 (2001). DOI: 10.1103/PhysRevD.63.023506. arXiv: astro-ph/0009131 [astro-ph]

[46] E. Gourgoulhon: "3+1 formalism and bases of numerical relativity" (2007). arXiv: gr-qc/0703035 [GR-QC]

[47] A. H. Guth: "Inflationary Universe: A possible Solution to the Horizon and Flatness Problems". *Physical Review D* 23.2, 347 (1981)

[48] M. C. Guzzetti et al.: "Gravitational Waves from Inflation". *Riv. Nuovo Cim.* 39.9, 399–495 (2016). DOI: 10.1393/ncr/i2016-10127-1. arXiv: 1605.01615 [astro-ph.CO]

[49] G. Hinshaw et al.: "Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: cosmological parameter results". *The Astrophysical Journal Supplement Series* 208.2, 19 (2013). DOI: 10.1088/0067-0049/208/2/19. URL: http://iopscience.iop.org/article/10.1088/0067-0049/208/2/19/meta

[50] R. Holman and A. J. Tolley: "Enhanced Non-Gaussianity from Excited Initial States". *JCAP* 0805, 001 (2008). DOI: 10.1088/1475-7516/2008/05/001. arXiv: 0710.1302 [hep-th]

[51] E. Hubble: "A Relation between Distance and Radial Velocity among Extra-galactic Nebulae". *Proceedings of the National Academy of Sciences* 15.3, 168–173 (1929)

[52] K. Izumi and S. Mukohyama: "Trispectrum from Ghost Inflation". *JCAP* 1006, 016 (2010). DOI: 10.1088/1475-7516/2010/06/016. arXiv: 1004.1776 [hep-th]

[53] D. I. Kaiser and A. T. Todhunter: "Primordial Perturbations from Multifield Inflation with Nonminimal Couplings". *Phys. Rev.* D81, 124037 (2010). DOI: 10.1103/PhysRevD.81.124037. arXiv: 1004.3805 [astro-ph.CO]

[54] E. Komatsu: "Hunting for Primordial Non-Gaussianity in the Cosmic Microwave Background". *Class. Quant. Grav.* 27, 124010 (2010). DOI: 10.1088/0264-9381/27/12/124010. arXiv: 1003.6097 [astro-ph.CO]

[55] E. Komatsu and D. N. Spergel: "Acoustic Signatures in the Primary Microwave Background Bispectrum". *Phys. Rev.* D63, 063002 (2001). DOI: 10.1103/PhysRevD.63.063002. arXiv: astro-ph/0005036 [astro-ph]

[56] H. Kurki-Suonio: "Cosmological Perturbation Theory, Part 1" (2015)

[57] H. Kurki-Suonio: "Cosmological Perturbation Theory, Part 2" (2015)

[58] D. Langlois and S. Renaux-Petel: "Perturbations in generalized multi-field inflation". *JCAP* 0804, 017 (2008). DOI: 10.1088/1475-7516/2008/04/017. arXiv: 0801.1085 [hep-th]

[59] G. Lematre: "Un Univers homogne de Masse Constante et de Rayon Croissant Rendant Compte de la Vitesse Radiale des Nbuleuses Extra-galactiques". *Annales de la Socit scientifique de Bruxelles*. Vol. 47. 1927, 49–59

[60] A. R. Liddle and D. H. Lyth: *Cosmological Inflation and Large Scale Structure*. 2000. ISBN: 9780521828499

[61] A. R. Liddle and D. H. Lyth: "COBE, Gravitational Waves, Inflation and Extended Inflation". *Phys. Lett.* B291, 391–398 (1992). DOI: 10.1016/0370-2693(92)91393-N. arXiv: astro-ph/9208007 [astro-ph]

[62] E. A. Lim: "Advanced Cosmology: Primordial non-Gaussianities". *Cambridge DAMTP Advanced Cosmology Lectures* (2012)

[63] D. H. Lyth, C. Ungarelli, and D. Wands: "The Primordial Density Perturbation in the Curvaton Scenario". *Phys. Rev.* D67, 023503 (2003). DOI: 10.1103/PhysRevD.67.023503. arXiv: astro-ph/0208055 [astro-ph]

[64] J. M. Maldacena: "Non-Gaussian Features of Primordial Fluctuations in Single Field Inflationary Models". *JHEP* 05, 013 (2003). DOI: 10.1088/1126-6708/2003/05/013. arXiv: astro-ph/0210603 [astro-ph]

[65] K. A. Malik and D. Wands: "Cosmological Perturbations". *Phys. Rept.* 475, 1–51 (2009). DOI: 10.1016/j.physrep.2009.03.001. arXiv: 0809.4944 [astro-ph]

[66] K. A. Malik: "Cosmological Perturbations in an Inflationary Universe". PhD thesis. Portsmouth U., 2001. arXiv: astro-ph/0101563 [astro-ph]

[67] V. Mukhanov: *Physical Foundations of Cosmology*. Cambridge: Cambridge Univ. Press, 2005. URL: https://cds.cern.ch/record/991646

[68] V. F. Mukhanov: "Quantum theory of Gauge-invariant Cosmological Perturbations". *Zh. Eksp. Teor. Fiz* 94.1 (1988)

[69] V. F. Mukhanov, H. A. Feldman, and R. H. Brandenberger: "Theory of Cosmological Perturbations. Part 1. Classical Perturbations. Part 2. Quantum Theory of Perturbations. Part 3. Extensions". *Phys. Rept.* 215, 203–333 (1992). DOI: 10.1016/0370-1573(92)90044-Z

[70] E. Pajer, G. L. Pimentel, and J. V. S. Van Wijck: "The Conformal Limit of Inflation in the Era of CMB Polarimetry". *JCAP* 1706.06, 009 (2017). DOI: 10.1088/1475-7516/2017/06/009. arXiv: 1609.06993 [hep-th]

[71] C. Patrignani et al.: "Review of Particle Physics". *Chin. Phys.* C40.10 (2016)

[72] M. E. Peskin: *An introduction to quantum field theory*. CRC Press, 2018

[73] P. Peter and J. Uzan: *Primordial cosmology*. Oxford University Press, 2013

[74] A. Riotto: "Inflation and the Theory of Cosmological Perturbations". *ICTP Lect. Notes Ser.* 14, 317–413 (2003). arXiv: hep-ph/0210162 [hep-ph]

[75] M. Sasaki: "Large Scale Quantum Fluctuations in the Inflationary Universe". *Progress of Theoretical Physics* 76.5, 1036–1046 (1986). DOI: 10.1143/PTP.76.1036. eprint: /oup/backfile/content_public/journal/ptp/76/5/10.1143/ptp.76.1036/2/76-5-1036.pdf. URL: +%20http://dx.doi.org/10.1143/PTP.76.1036

[76] D. Seery and J. E. Lidsey: "Primordial non-Gaussianities from multiple-field inflation". *JCAP* 0509, 011 (2005). DOI: 10.1088/1475-7516/2005/09/011. arXiv: astro-ph/0506056 [astro-ph]

[77] D. Seery and J. E. Lidsey: "Primordial non-Gaussianities in single field inflation". *JCAP* 0506, 003 (2005). DOI: 10.1088/1475-7516/2005/06/003. arXiv: astro-ph/0503692 [astro-ph]

[78] D. Seery, M. S. Sloth, and F. Vernizzi: "Inflationary trispectrum from graviton exchange". *JCAP* 0903, 018 (2009). DOI: 10.1088/1475-7516/2009/03/018. arXiv: 0811.3934 [astro-ph]

[79] L. Senatore: "Lectures on Inflation". *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings (TASI 2015): Boulder, CO, USA, June 1-26, 2015*. 2017, 447–543. DOI: 10.1142/9789813149441_0008. arXiv: 1609.00716 [hep-th]. URL: https://inspirehep.net/record/1485069/files/arXiv:1609.00716.pdf

[80] L. Senatore and M. Zaldarriaga: "The constancy of ζ in Single-clock Inflation at all Loops". *JHEP* 09, 148 (2013). DOI: 10.1007/JHEP09(2013)148. arXiv: 1210.6048 [hep-th]

[81] D. Tong: "Advanced Cosmology: Primordial non-Gaussianities and detection of stochastic gravitational waves, DAMPT Cambridge" (2012). URL: http://www.damtp.cam.ac.uk/user/eal40/teach/AdvCos

[82] D. Tong: "Lectures on Quantum Field Theory, DAMPT Cambridge" (2006). URL: http://www.damtp.cam.ac.uk/user/tong/qft.html

[83] S. Tsujikawa and H. Yajima: "New constraints on multifield inflation with nonminimal coupling". *Phys. Rev.* D62, 123512 (2000). DOI: 10.1103/PhysRevD.62.123512. arXiv: hep-ph/0007351 [hep-ph]

[84] L. Verde: "A Practical Guide to Basic Statistical Techniques for Data Analysis in Cosmology" (2007). arXiv: 0712.3028 [astro-ph]

[85] F. Vernizzi and D. Wands: "Non-gaussianities in two-field inflation". *JCAP* 0605, 019 (2006). DOI: 10.1088/1475-7516/2006/05/019. arXiv: astro-ph/0603799 [astro-ph]

[86] L. Wang and A. Mazumdar: "Cosmological Perturbations from a Spectator Field during Inflation". *JCAP* 1305, 012 (2013). DOI: 10.1088/1475-7516/2013/05/012. arXiv: 1302.2637 [astro-ph.CO]

[87] M. Wang: "Primordial Gravitational Waves from Cosmic Inflation, DAMPT Cambridge" (2017)

[88] Y. Wang: "Inflation, Cosmic Perturbations and Non-Gaussianities". *Commun. Theor. Phys.* 62, 109–166 (2014). DOI: 10.1088/0253-6102/62/1/19. arXiv: 1303.1523 [hep-th]

[89] S. Weinberg: "Gravitation and cosmology: principles and applications of the general theory of relativity" (1972)

[90] S. Weinberg: "Adiabatic modes in cosmology". *Phys. Rev.* D67, 123504 (2003). DOI: 10.1103/PhysRevD.67.123504. arXiv: astro-ph/0302326 [astro-ph]

[91] S. Weinberg: "Can non-adiabatic perturbations arise after single-field inflation?" *Phys. Rev.* D70, 043541 (2004). DOI: 10.1103/PhysRevD.70.043541. arXiv: astro-ph/0401313 [astro-ph]

[92] S. Weinberg: *Cosmology*. Oxford University Press, 2008

[93] S. Weinberg: "Effective Field Theory for Inflation". *Phys. Rev.* D77, 123541 (2008). DOI: 10.1103/PhysRevD.77.123541. arXiv: 0804.4291 [hep-th]

[94] T. de Wild: "Aspects of Inflationary Cosmology". *University of Groningen* (2017)