

# TASI Lectures on Inflation

Daniel Baumann

*Department of Physics, Harvard University, Cambridge, MA 02138, USA*

---

## Abstract

In a series of five lectures I review inflationary cosmology. I begin with a description of the initial conditions problems of the Friedmann-Robertson-Walker (FRW) cosmology and then explain how inflation, an early period of accelerated expansion, solves these problems. Next, I describe how inflation transforms microscopic quantum fluctuations into macroscopic seeds for cosmological structure formation. I present in full detail the famous calculation for the primordial spectra of scalar and tensor fluctuations. I then define the inverse problem of extracting information on the inflationary era from observations of cosmic microwave background fluctuations. The current observational evidence for inflation and opportunities for future tests of inflation are discussed. Finally, I review the challenge of relating inflation to fundamental physics by giving an account of inflation in string theory.

The lectures are at the graduate level assuming some background in general relativity and quantum field theory.

## Lecture 1: Classical Dynamics of Inflation

The aim of this lecture is a first-principles introduction to the classical dynamics of inflationary cosmology. After a brief review of basic FRW cosmology we show that the conventional Big Bang theory leads to an initial conditions problem: the universe as we know it can only arise for very special and finely-tuned initial conditions. We then explain how inflation (an early period of accelerated expansion) solves this initial conditions problem and allows our universe to arise from generic initial conditions. We describe the necessary conditions for inflation and explain how inflation modifies the causal structure of spacetime to solve the Big Bang puzzles. Finally, we end this lecture with a discussion of the physical origin of the inflationary expansion.

## Lecture 2: Quantum Fluctuations during Inflation

In this lecture we review the famous calculation of the primordial fluctuation spectra generated by quantum fluctuations during inflation. We present the calculation in full detail and try to avoid ‘cheating’ and approximations. After a brief review of fundamental aspects of cosmological perturbation theory, we first give a qualitative summary of the basic mechanism by which inflation converts microscopic quantum fluctuations into macroscopic seeds for cosmological structure formation. As a pedagogical introduction to quantum field theory in curved spacetime we then review the quantization of the simple harmonic oscillator. We emphasize that a unique vacuum state is

chosen by demanding that the vacuum is the minimum energy state. We then proceed by giving the corresponding calculation for inflation in all detail. We calculate the power spectra of both scalar and tensor fluctuations.

### **Lecture 3: Contact with Observations**

In this lecture we describe the inverse problem of extracting information on the inflationary perturbation spectra from observations of the cosmic microwave background and the large-scale structure. We define the precise relations between the gauge-invariant scalar and tensor power spectra computed in the previous lecture and the observed CMB anisotropies and galaxy power spectra. We give the transfer functions that relate the primordial fluctuations to the late-time observables. We then use these results to discuss the current observational evidence for inflation. Finally, we indicate opportunities for future tests of inflation.

### **Lecture 4: Primordial Non-Gaussianity**

In this lecture we summarize key theoretical results in the study of primordial non-Gaussianity. Most results are stated without proof, but their significance for constraining the fundamental physical origin of inflation is explained. After introducing the bispectrum as a basic diagnostic of non-Gaussian statistics, we show that its momentum dependence is a powerful probe of the inflationary action. Large non-Gaussianity can only arise if inflaton interactions are significant during inflation. In single-field slow-roll inflation non-Gaussianity is therefore predicted to be unobservably small, while it can be significant in models with multiple fields, higher-derivative interactions or non-standard initial states. Finally, we end the lecture with a discussion of the observational prospects for detecting or constraining primordial non-Gaussianity.

### **Lecture 5: Inflation in String Theory**

We end this lecture series with a discussion of a slightly more advanced topic: inflation in string theory. We provide a pedagogical overview of the subject. Our central theme is the sensitivity of inflation to Planck-scale physics, which we argue provides both the primary motivation and the central theoretical challenge for realizing inflation in string theory. We illustrate these issues through two case studies of inflationary scenarios in string theory: warped D-brane inflation and axion monodromy inflation. Finally, we indicate opportunities for future progress both theoretically and observationally.

---

# Contents

<b>I</b>	<b>Introduction</b>	<b>9</b>
<b>1</b>	<b>The Microscopic Origin of Structure</b>	<b>9</b>
1.1	TASI 2009: <i>The Physics of the Large and the Small</i> . . . . .	9
1.2	Structure and Evolution of the Universe . . . . .	10
1.3	The First $10^{-10}$ Seconds . . . . .	12
<b>2</b>	<b>Outline of the Lectures</b>	<b>13</b>
<b>II</b>	<b>Lecture 1: Classical Dynamics of Inflation</b>	<b>15</b>
<b>3</b>	<b>Review: The Homogeneous Universe</b>	<b>15</b>
3.1	FRW Spacetime . . . . .	15
3.2	Kinematics: Conformal Time and Horizons . . . . .	16
3.3	Dynamics: Einstein Equations . . . . .	18
<b>4</b>	<b>Big Bang Puzzles</b>	<b>20</b>
4.1	The Cauchy Problem of the Universe . . . . .	21
4.2	Horizon Problem . . . . .	21
4.3	Flatness Problem . . . . .	22
4.4	On the Problem of Initial Conditions . . . . .	23
<b>5</b>	<b>A First Look at Inflation</b>	<b>23</b>
5.1	The Shrinking Hubble Sphere . . . . .	23
5.1.1	Comoving Horizon during Inflation . . . . .	23
5.1.2	Flatness Problem Revisited . . . . .	24
5.1.3	Horizon Problem Revisited . . . . .	24
5.2	Conditions for Inflation . . . . .	24
5.3	Conformal Diagram of Inflation . . . . .	26
<b>6</b>	<b>The Physics of Inflation</b>	<b>27</b>
6.1	Scalar Field Dynamics . . . . .	27
6.2	Slow-Roll Inflation . . . . .	29
6.3	Case Study: $m^2\phi^2$ Inflation . . . . .	31
6.4	Reheating . . . . .	32
6.5	Models of Inflation . . . . .	32
6.5.1	Single-Field Slow-Roll Inflation . . . . .	32
6.5.2	Beyond Single-Field Slow-Roll . . . . .	34
<b>7</b>	<b>Summary: Lecture 1</b>	<b>36</b>
<b>8</b>	<b>Problem Set: Lecture 1</b>	<b>38</b>

<b>III Lecture 2: Quantum Fluctuations during Inflation</b>	<b>40</b>
<b>9 Review: Cosmological Perturbations</b>	<b>41</b>
9.1 Generalities . . . . .	41
9.1.1 Linear Perturbations . . . . .	41
9.1.2 Gauge Choice . . . . .	41
9.1.3 Scalars, Vectors and Tensors . . . . .	42
9.2 The Inhomogeneous Universe . . . . .	43
9.2.1 Metric Perturbations . . . . .	43
9.2.2 Matter Perturbations . . . . .	44
9.2.3 Gauge-Invariant Variables . . . . .	44
9.2.4 Superhorizon (Non-)Evolution . . . . .	45
9.3 Statistics of Cosmological Perturbations . . . . .	46
<b>10 Preview: The Quantum Origin of Structure</b>	<b>47</b>
10.1 Quantum Zero-Point Fluctuations . . . . .	47
10.2 Horizon Exit and Re-Entry . . . . .	48
<b>11 Quantum Mechanics of the Harmonic Oscillator</b>	<b>48</b>
11.1 Action . . . . .	49
11.2 Canonical Quantization . . . . .	49
11.3 Non-Uniqueness of the Mode Functions . . . . .	50
11.4 Zero-Point Fluctuations in the Ground State . . . . .	51
<b>12 Quantum Fluctuations in de Sitter Space</b>	<b>51</b>
12.1 Summary of the Computational Strategy . . . . .	52
12.2 Scalar Perturbations . . . . .	52
12.2.1 Free Field Action . . . . .	53
12.2.2 Quantization . . . . .	53
12.2.3 Boundary Conditions and Bunch-Davies Vacuum . . . . .	54
12.2.4 Solution in de Sitter Space . . . . .	54
12.2.5 Power Spectrum in Quasi-de Sitter . . . . .	55
12.2.6 Spatially-Flat Gauge . . . . .	56
12.3 Tensor Perturbations . . . . .	56
12.3.1 Action . . . . .	56
12.3.2 Quantization . . . . .	57
12.3.3 Power Spectrum . . . . .	57
12.3.4 The Tensor-to-Scalar Ratio . . . . .	57
12.4 The Lyth Bound . . . . .	58
<b>13 Primordial Spectra</b>	<b>58</b>
13.1 Scale-Dependence . . . . .	58
13.2 Slow-Roll Results . . . . .	59
13.3 Case Study: $m^2\phi^2$ Inflation . . . . .	60

<b>14 Summary: Lecture 2</b>	<b>61</b>
<b>15 Problem Set: Lecture 2</b>	<b>62</b>
<b>IV Lecture 3: Contact with Observations</b>	<b>64</b>
<b>16 Connecting Observations to the Early Universe</b>	<b>64</b>
<b>17 Review: The Cosmic Microwave Background</b>	<b>66</b>
17.1 Temperature Anisotropies . . . . .	66
17.1.1 Harmonic Expansion . . . . .	66
17.1.2 Large Scales . . . . .	68
17.1.3 Non-Gaussianity . . . . .	69
17.2 Polarization . . . . .	70
17.2.1 Polarization via Thomson Scattering . . . . .	70
17.2.2 Characterization of the Radiation Field . . . . .	70
17.2.3 $E$ and $B$ -modes . . . . .	71
17.2.4 $E$ -modes and Scalars . . . . .	73
17.2.5 $B$ -modes and Tensors . . . . .	73
<b>18 Review: Large-Scale Structure</b>	<b>73</b>
18.1 Dark Matter Transfer Functions . . . . .	74
18.2 Galaxy Bias . . . . .	75
<b>19 Current Evidence for Inflation</b>	<b>76</b>
19.1 Flatness . . . . .	76
19.2 Coherent Phases and Superhorizon Fluctuations . . . . .	76
19.2.1 The Peaks of the TT Spectrum . . . . .	76
19.2.2 $\ell < 100$ in the TE Spectrum . . . . .	77
19.3 Scale-Invariant, Gaussian and Adiabatic . . . . .	79
19.3.1 Spectral Index . . . . .	79
19.3.2 Gaussianity . . . . .	80
19.3.3 Adiabaticity . . . . .	81
19.4 Testing Slow-Roll Models . . . . .	81
<b>20 Future Tests of Inflation</b>	<b>82</b>
20.1 Amplitude of Tensor Modes . . . . .	82
20.2 Scale Dependence of Scalar Modes . . . . .	83
20.3 Scale Dependence of Tensor Modes . . . . .	83
20.4 Non-Gaussianity . . . . .	83
20.5 Isocurvature Fluctuations . . . . .	84
<b>21 Summary: Lecture 3</b>	<b>86</b>

<b>V Lecture 4: Primordial Non-Gaussianity</b>	<b>87</b>
<b>22 Preliminaries</b>	<b>87</b>
22.1 The Bispectrum and Local Non-Gaussianity . . . . .	87
22.2 Shapes of Non-Gaussianity . . . . .	88
22.3 $f_{NL}$ : The Amplitude of Non-Gaussianity . . . . .	89
<b>23 Theoretical Expectations</b>	<b>89</b>
23.1 Single-Field Slow-Roll Inflation . . . . .	89
23.2 The Maldacena Theorem . . . . .	89
23.3 Models with Large Non-Gaussianity . . . . .	90
23.3.1 Higher-Derivative Interactions . . . . .	90
23.3.2 Multiple Fields . . . . .	92
23.3.3 Non-Standard Vacuum . . . . .	92
<b>24 Observational Prospects</b>	<b>92</b>
24.1 Cosmic Microwave Background . . . . .	92
24.2 Large-Scale Structure . . . . .	92
<b>25 Summary: Lecture 4</b>	<b>93</b>
<b>VI Lecture 5: Inflation in String Theory</b>	<b>94</b>
<b>26 Why Combine Two Speculative Ideas?</b>	<b>94</b>
<b>27 UV Sensitivity of Inflation</b>	<b>94</b>
27.1 Effective Field Theory and Inflation . . . . .	94
27.2 The Eta Problem . . . . .	96
27.2.1 Radiative Instability of the Inflaton Mass . . . . .	96
27.2.2 Supergravity Example . . . . .	96
27.3 The Lyth Bound . . . . .	98
27.4 Super-Planckian Fields and Flat Potentials . . . . .	98
27.4.1 No Shift Symmetry . . . . .	98
27.4.2 Shift Symmetry . . . . .	99
<b>28 Inflation in String Theory</b>	<b>99</b>
28.1 From String Compactifications to the Inflaton Action . . . . .	99
28.1.1 Elements of String Compactifications . . . . .	99
28.1.2 The Effective Inflaton Action . . . . .	101
28.2 Case Study: Warped D-brane Inflation . . . . .	101
28.2.1 D3-branes in Warped Throat Geometries . . . . .	102
28.2.2 The Field Range Bound . . . . .	103
28.2.3 The D3-brane Potential . . . . .	103
28.2.4 Sketch of the Supergravity Analysis . . . . .	104
28.2.5 Phenomenological Implications . . . . .	105

28.2.6	Summary and Perspective . . . . .	105
28.3	Case Study: Axion Monodromy Inflation . . . . .	106
28.3.1	Axions in String Theory . . . . .	106
28.3.2	Axion Inflation in String Theory . . . . .	107
28.3.3	Compactification Considerations . . . . .	109
28.3.4	Summary and Perspective . . . . .	109
<b>29</b>	<b>Outlook</b>	<b>110</b>
29.1	Theoretical Prospects . . . . .	110
29.2	Observational Signatures? . . . . .	111
<b>30</b>	<b>Summary: Lecture 5</b>	<b>112</b>
<b>VII</b>	<b>Conclusions</b>	<b>113</b>
<b>31</b>	<b>Future Prospects and Open Problems</b>	<b>113</b>
<b>32</b>	<b>Guide to Further Reading</b>	<b>114</b>
<b>VIII</b>	<b>Appendix</b>	<b>116</b>
<b>A</b>	<b>Cosmological Perturbation Theory</b>	<b>116</b>
A.1	The Perturbed Universe . . . . .	116
A.1.1	Metric Perturbations . . . . .	116
A.1.2	Stress-Energy Perturbations . . . . .	116
A.2	Scalars, Vectors and Tensors . . . . .	117
A.2.1	Helicity and SVT-Decomposition in Fourier Space . . . . .	117
A.2.2	Real Space SVT-Decomposition . . . . .	120
A.3	Scalars . . . . .	121
A.3.1	Metric Perturbations . . . . .	121
A.3.2	Matter Perturbations . . . . .	122
A.3.3	Einstein Equations . . . . .	122
A.3.4	Popular Gauges . . . . .	123
A.4	Vectors . . . . .	126
A.4.1	Metric Perturbations . . . . .	126
A.4.2	Matter Perturbations . . . . .	126
A.4.3	Einstein Equations . . . . .	127
A.5	Tensors . . . . .	127
A.5.1	Metric Perturbations . . . . .	127
A.5.2	Matter Perturbations . . . . .	127
A.5.3	Einstein Equations . . . . .	127
A.6	Statistics . . . . .	128
A.6.1	Fourier Conventions . . . . .	128
A.6.2	Two-Point Correlation Function . . . . .	128

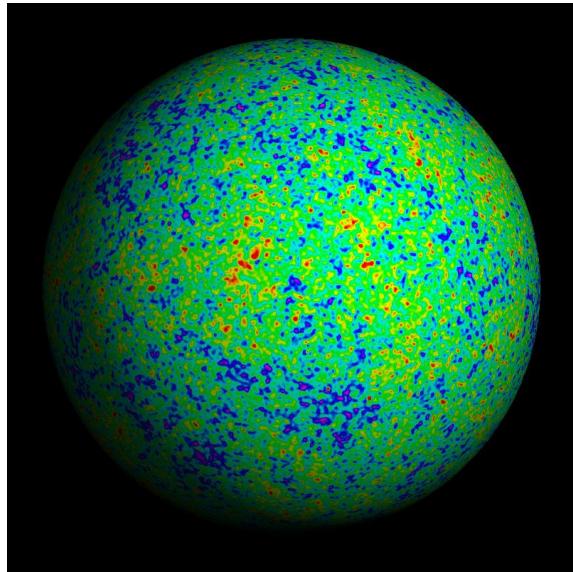
A.6.3	Power Spectrum . . . . .	128
A.6.4	Bispectrum . . . . .	129
<b>B</b>	<b>Free Field Action for <math>\mathcal{R}</math></b>	<b>130</b>
B.1	Slow-Roll Background . . . . .	130
B.2	ADM Formalism . . . . .	130
B.2.1	Comoving Gauge . . . . .	131
B.2.2	Constraint Equations . . . . .	131
B.2.3	The Free Field Action . . . . .	132
<b>C</b>	<b>Slow-Roll Inflation in the Hamilton-Jacobi Approach</b>	<b>133</b>
C.1	Hamilton-Jacobi Formalism . . . . .	133
C.2	Hubble Slow-Roll Parameters . . . . .	133
C.3	Slow-Roll Inflation . . . . .	134
C.4	Inflationary Attractor Solution . . . . .	135

# Part I

## Introduction

“I’m astounded by people who want to ‘know’ the Universe  
when it’s hard enough to find your way around Chinatown”

*Woody Allen*



**Figure 1:** Fluctuations in the Cosmic Microwave Background (CMB). What produced them?

## 1 The Microscopic Origin of Structure

### 1.1 TASI 2009: *The Physics of the Large and the Small*

The fluctuations in the temperature of the cosmic microwave background (CMB) (see Fig. 1) tell an amazing story. Measured now almost routinely by experiments like the Wilkinson Microwave Anisotropy Probe (WMAP), the temperature variations of the microwave sky bear testimony of minute fluctuations in the density of the primordial universe. These fluctuations grew via gravitational instability into the large-scale structures (LSS) that we observe in the universe today. The success in relating observations of the thermal afterglow of the Big Bang to the formation of structures billions of years later motivate us to ask an even bolder question: what is the fundamental microphysical origin of the CMB fluctuations? An answer to this question would provide us with nothing less than a fundamental understanding of the physical origin of all structure in the universe.

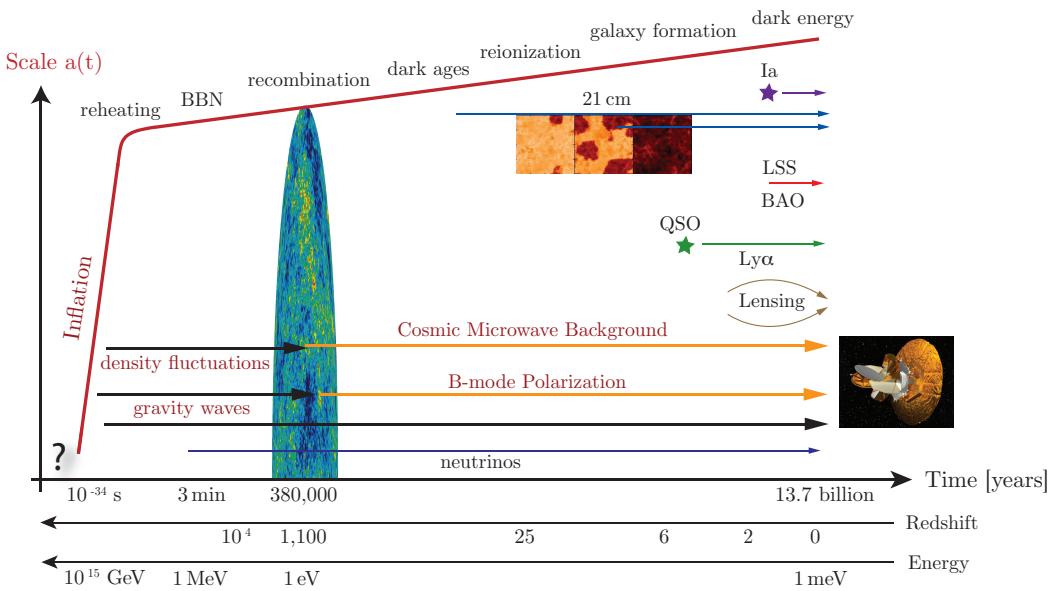
In these lectures, I will describe the currently leading working hypothesis that a period of cosmic *inflation* was integral part of this picture for the formation and evolution of structure. Inflation [1–3],

a period of exponential expansion in the very early universe, is believed to have taken place some  $10^{-34}$  seconds after the Big Bang singularity. Remarkably, inflation is thought to be responsible both for the large-scale homogeneity of the universe and the small fluctuations that were the seeds for the formation of structures like our own galaxy.

The central focus of this lecture series will be to explain in full detail the physical mechanism by which inflation transformed microscopic quantum fluctuations into macroscopic fluctuations in the energy density of the universe. In this sense inflation provides the most dramatic example for the theme of TASI 2009: the connection between the ‘physics of the large and the small’. We will calculate explicitly the statistical properties and the scale dependence of the spectrum of fluctuations produced by inflation. This result provides the input for all studies of cosmological structure formation and is one of the great triumphs of modern theoretical cosmology.

## 1.2 Structure and Evolution of the Universe

There is undeniable evidence for the expansion of the universe: the light from distant galaxies is systematically shifted towards the red end of the spectrum, the observed abundances of the light elements (H, He, and Li) matches the predictions of Big Bang Nucleosynthesis (BBN), and the only convincing explanation for the CMB is a relic radiation from a hot early universe.



**Figure 2:** History of the universe. In this schematic we present key events in the history of the universe and their associated time and energy scales. We also illustrate several cosmological probes that provide us with information about the structure and evolution of the universe. *Acronyms:* BBN (Big Bang Nucleosynthesis), LSS (Large-Scale Structure), BAO (Baryon Acoustic Oscillations), QSO (Quasi-Stellar Objects = Quasars), Ly $\alpha$  (Lyman-alpha), CMB (Cosmic Microwave Background), Ia (Type Ia supernovae), 21cm (hydrogen 21cm-transition).

Two principles characterize thermodynamics and particle physics in an expanding universe: i) interactions between particles freeze out when the interaction rate  $\Gamma$  drops below the expansion

rate  $H$ , and ii) broken symmetries in the laws of physics may be restored at high energies. Table 1 shows the thermal history of the universe and various phase transitions related to symmetry breaking events. In the following we will give a quick qualitative summary of these milestones in the evolution of our universe. We will emphasize which aspects of this cosmological story are based on established physics and which require more speculative ideas.

**Table 1:** Major Events in the History of the Universe

	Time	Energy	
Planck Epoch?	$< 10^{-43}$ s	$10^{18}$ GeV	
String Scale?	$\gtrsim 10^{-43}$ s	$\lesssim 10^{18}$ GeV	
Grand Unification?	$\sim 10^{-36}$ s	$10^{15}$ GeV	
Inflation?	$\gtrsim 10^{-34}$ s	$\lesssim 10^{15}$ GeV	
SUSY Breaking?	$< 10^{-10}$ s	$> 1$ TeV	
Baryogenesis?	$< 10^{-10}$ s	$> 1$ TeV	
Electroweak Unification	$10^{-10}$ s	1 TeV	
Quark-Hadron Transition	$10^{-4}$ s	$10^2$ MeV	
Nucleon Freeze-out	0.01 s	10 MeV	
Neutrino Decoupling	1 s	1 MeV	
BBN	3 min	0.1 MeV	
			Redshift
Matter-Radiation Equality	$10^4$ yrs	1 eV	$10^4$
Recombination	$10^5$ yrs	0.1 eV	1,100
Dark Ages	$10^5 - 10^8$ yrs		$> 25$
Reionization	$10^8$ yrs		25 – 6
Galaxy Formation	$\sim 6 \times 10^8$ yrs		$\sim 10$
Dark Energy	$\sim 10^9$ yrs		$\sim 2$
Solar System	$8 \times 10^9$ yrs		0.5
Albert Einstein born	$14 \times 10^9$ yrs	1 meV	0

From  $10^{-10}$  seconds to today the history of the universe is based on well understood and experimentally tested laws of particle physics, nuclear and atomic physics and gravity. We are therefore justified to have some confidence about the events shaping the universe during that time (see Figure 2 and Table 1).

Let us enter the universe at 100 GeV, the time of the electroweak phase transition ( $10^{-10}$  s). Above 100 GeV the electroweak symmetry is restored and the  $Z$  and  $W^\pm$  bosons are massless. Interactions are strong enough to keep quarks and leptons in thermal equilibrium. Below 100 GeV the symmetry between the electromagnetic and the weak forces is broken,  $Z$  and  $W^\pm$  bosons acquire mass and the cross-section of weak interactions decreases as the temperature of the universe drops. As a result, at 1 MeV, neutrinos decouple from the rest of the matter. Shortly after, at

1 second, the temperature drops below the electron rest mass and electrons and positrons annihilate efficiently. Only an initial matter-antimatter asymmetry of one part in a billion survives. The resulting photon-baryon fluid is in equilibrium. Around 0.1 MeV the strong interaction becomes important and protons and neutrons combine into the light elements (H, He, Li) during Big Bang Nucleosynthesis (BBN) ( $\sim 200$  s). The successful prediction of the H, He and Li abundances is one of the most striking consequences of the Big Bang theory. The matter and radiation densities are equal around 1 eV ( $10^{11}$  s). Charged matter particles and photons are strongly coupled in the plasma and fluctuations in the density propagate as cosmic ‘sound waves’. Around 0.1 eV (380,000 yrs) protons and electrons combine into neutral hydrogen atoms. Photons decouple and form the free-streaming cosmic microwave background. 13.7 billion years later these photons give us the earliest snapshot of the universe. Anisotropies in the CMB temperature give us the evidence for fluctuations in the primordial matter density.

These small density perturbations,  $\rho(\mathbf{x}, t) = \bar{\rho}(t)[1 + \delta(\mathbf{x}, t)]$ , grow via gravitational instability to form the large-scale structures observed in the late universe. A competition between the background pressure and the universal attraction of gravity determines the details of the growth of structure. During radiation domination the growth is slow,  $\delta \sim \ln a$  (where  $a(t)$  is the scale factor describing the expansion of space). Clustering becomes more efficient after matter dominates the background density (and the pressure drops to zero),  $\delta \sim a$ . Small scales become non-linear first,  $\delta \gtrsim 1$ , and form gravitationally bound objects that decouple from the overall expansion. This leads to a picture of hierarchical structure formation with small-scale structures (like stars and galaxies) forming first and then merging into larger structures (clusters and superclusters of galaxies). Around redshift  $z \sim 25$  ( $1 + z = a^{-1}$ ), high energy photons from the first stars begin to ionize the hydrogen in the inter-galactic medium. This process of ‘reionization’ is completed at  $z \approx 6$ . Meanwhile, the most massive stars run out of nuclear fuel and explode as ‘supernovae’. In these explosions the heavy elements (C, O, ...) necessary for the formation of life are created, leading to the slogan “we are all stardust”. At  $z \approx 1$ , a negative pressure ‘dark energy’ comes to dominate the universe. The background spacetime is accelerating and the growth of structure ceases,  $\delta \sim \text{const.}$

### 1.3 The First $10^{-10}$ Seconds

The history of the universe from  $10^{-10}$  seconds (1 TeV) to today is based on observational facts and tested physical theories like the Standard Model of particle physics, general relativity and fluid dynamics, *e.g.* the fundamental laws of high energy physics are well-established up to the energies reached by current particle accelerators ( $\sim 1$  TeV). Before  $10^{-10}$  seconds, the energy of the universe exceeds 1 TeV and we lose the comfort of direct experimental guidance. The physics of that era is therefore as speculative as it is fascinating.

To explain the fluctuations seen in the CMB temperature requires an input of primordial seed fluctuations. In these lectures we will explain the conjecture that these primordial fluctuations were generated in the very early universe ( $\sim 10^{-34}$  seconds) during a period of inflation. We will explain how microscopic quantum fluctuations in the energy density get stretched by the inflationary expansion to macroscopic scales, larger than the physical horizon at the time. After a perturbation exits the horizon no causal physics can affect it and it remains frozen with constant amplitude until it re-enters the horizon at a later time during the conventional (non-accelerating) Big Bang expansion. The fluctuations associated with cosmological structures re-enter the horizon when the universe is

about 100,000 years old, a short time before the decoupling of the CMB photons. Inside the horizon causal physics can affect the perturbation amplitudes and in fact leads to the acoustic peak structure of the CMB and the collapse of high-density fluctuations into galaxies and clusters of galaxies. Since we understand (and can calculate) the evolution of perturbations after they re-enter the horizon we can use the late time observations of the CMB and the LSS to infer the primordial input spectrum. Assuming this spectrum was produced by inflation, this gives us an observational probe of the physical conditions when the universe was  $10^{-34}$  seconds old. This fascinating opportunity to use cosmology to probe physics at the highest energies will be the subject of these lectures.

## 2 Outline of the Lectures

In Lecture 1 we introduce the *classical* background dynamics of inflation. We explain how inflation solves the horizon and flatness problems. We discuss the slow-roll conditions and reheating and speculate on the physical origin of the inflationary expansion. In Lecture 2 we describe how *quantum* fluctuations during inflation become the seeds for the formation of large-scale structures. We present in full detail the derivation of the inflationary power spectra of scalar and tensor perturbations,  $\mathcal{R}$  and  $h_{ij}$ . In Lecture 3 we relate the results of Lecture 2 to observations of the cosmic microwave background and the distribution of galaxies, *i.e.* we explain how to measure  $P_{\mathcal{R}}(k)$  and  $P_h(k)$  in the sky! We describe current observational constraints and emphasize future tests of inflation. In Lecture 4 we present key results in the study of non-Gaussianity of the primordial fluctuations. We explain how non-Gaussian correlations can provide important information on the inflationary action. We reserve Lecture 5 for the study of an advanced topic that is at the frontier of current research: inflation in string theory. We describe the main challenges of the subject and summarize recent advances.

To make each lecture self-contained, the necessary background material is presented in a short review section preceding the core of each lecture. Every lecture ends with a *summary* of the most important results. An important part of every lecture are problems and exercises that appear throughout the text and (for longer problems) as a separate *problem set* appended to the end of the lecture. The exercises were carefully chosen to complement the material of the lecture or to fill in certain details of the computations.

A number of appendices collect standard results from cosmological perturbation theory and details of the inflationary perturbation calculation. It is hoped that the appendices provide a useful reference for the reader. In fact, in recent research projects I frequently found myself consulting these appendices, so I hope that they can be equally useful for others.

## Notation

I have tried hard to keep the notation of these lectures coherent and consistent:

Throughout we will use the God-given natural units

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

We use the reduced Planck mass

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

and often set it equal to one. Our metric signature is  $(- +++)$ . Greek indices will take the values  $\mu, \nu = 0, 1, 2, 3$  and latin indices stand for  $i, j = 1, 2, 3$ . Our Fourier convention is

$$\mathcal{R}_{\mathbf{k}} = \int d^3x \mathcal{R}(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}},$$

so that the power spectrum is

$$\langle \mathcal{R}_{\mathbf{k}} \mathcal{R}_{\mathbf{k}'} \rangle = (2\pi)^3 P_{\mathcal{R}}(k) \delta(\mathbf{k} + \mathbf{k}'), \quad \Delta_{\mathcal{R}}^2(k) \equiv \frac{k^3}{2\pi^2} P_{\mathcal{R}}(k).$$

For conformal time we use the letter  $\tau$  (and caution the reader not confuse it with the astrophysical parameter for optical depth). We reserve the letter  $\eta$  for the second slow-roll parameter. Derivatives with respect to physical time are denoted by overdots. Derivatives with respect to conformal time are denoted by primes. Partial derivatives are denoted by commas, covariant derivatives by semi-colons.

## Acknowledgements

I am most grateful to Scott Dodelson and Csaba Csaki for the invitation to lecture at TASI 2009.

In past few years I have learned many things from my teachers Liam McAllister, Paul Steinhardt, and Matias Zaldarriaga and my collaborators Igor Klebanov, Anatoly Dymarsky, Shamit Kachru, Hiranya Peiris, Alberto Nicolis and Asantha Cooray. Thanks to all of them for generously sharing their insights with me.

Finally, I appreciated the input from the members of the CMBPol Inflation Working Group [4]. Their expert opinions on many of the topics described in these lectures were most valuable to me.

## Part II

# Lecture 1: Classical Dynamics of Inflation

---

### Abstract

The aim of this lecture is a first principles introduction to the classical dynamics of inflationary cosmology. After a brief review of basic FRW cosmology we show that the conventional Big Bang theory leads to an initial conditions problem: the universe as we know it can only arise for very special and finely-tuned initial conditions. We then explain how inflation (an early period of accelerated expansion) solves this initial conditions problem and allows our universe to arise from generic initial conditions. We describe the necessary conditions for inflation and explain how inflation modifies the causal structure of spacetime to solve the Big Bang puzzles. Finally, we end this lecture with a discussion of the physical origin of the inflationary expansion.

---

## 3 Review: The Homogeneous Universe

To begin, we review basic aspects of the homogeneous universe. Since this material was covered in Prof. Turner's lectures at TASI 2009 and is part of any textbook treatment of cosmology (*e.g.* [5–7]), we will be brief and recall many of the concepts via exercises for the reader. We will naturally focus on the elements most relevant for the study of inflation.

### 3.1 FRW Spacetime

Cosmology describes the structure and evolution of the universe on the largest scales. Assuming *homogeneity* and *isotropy*<sup>1</sup> on large scales one is lead to the Friedmann-Robertson-Walker (FRW) metric for the spacetime of the universe (see **Problem 1**):

$$ds^2 = -dt^2 + a^2(t) \left( \frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \right). \quad (1)$$

Here, the scale factor  $a(t)$  characterizes the relative size of spacelike hypersurfaces  $\Sigma$  at different times. The curvature parameter  $k$  is  $+1$  for positively curved  $\Sigma$ ,  $0$  for flat  $\Sigma$ , and  $-1$  for negatively curved  $\Sigma$ . Eqn. (1) uses comoving coordinates – the universe expands as  $a(t)$  increases, but galaxies/observers keep fixed coordinates  $r, \theta, \phi$  as long as there aren't any forces acting on them, *i.e.* in

---

<sup>1</sup>A *homogeneous* space is one which is translation invariant, or the same at every point. An *isotropic* space is one which is rotationally invariant, or the same in every direction. A space which is everywhere isotropic is necessarily homogeneous, but the converse is not true; *e.g.* a space with a uniform electric field is translationally invariant but not rotationally invariant.

the absence of peculiar motion. The corresponding physical distance is obtained by multiplying with the scale factor,  $R = a(t)r$ , and is time-dependent even for objects with vanishing peculiar velocities. By a coordinate transformation the metric (1) may be written as

$$ds^2 = -dt^2 + a^2(t) (\mathrm{d}\chi^2 + \Phi_k(\chi^2)(\mathrm{d}\theta^2 + \sin^2 \theta \mathrm{d}\phi^2)) , \quad (2)$$

where

$$r^2 = \Phi_k(\chi^2) \equiv \begin{cases} \sinh^2 \chi & k = -1 \\ \chi^2 & k = 0 \\ \sin^2 \chi & k = +1 \end{cases} . \quad (3)$$

For the FRW ansatz the evolution of the homogeneous universe boils down to the single function  $a(t)$ . Its form is dictated by the matter content of the universe via the Einstein field equations (see §3.3). An important quantity characterizing the FRW spacetime is the expansion rate

$$H \equiv \frac{\dot{a}}{a} . \quad (4)$$

The *Hubble parameter*  $H$  has unit of inverse time and is positive for an expanding universe (and negative for a collapsing universe). It sets the fundamental scale of the FRW spacetime, *i.e.* the characteristic timescale of the homogeneous universe is the Hubble time,  $t \sim H^{-1}$ , and the characteristic lengthscale is the Hubble length,  $d \sim H^{-1}$  (in units where  $c = 1$ ). The Hubble scale sets the scale for the age of the universe, while the Hubble length sets the size of the observable universe.

## 3.2 Kinematics: Conformal Time and Horizons

Having defined the metric for the average spacetime of the universe we can now study kinematical properties of the propagation of light and matter particles.

*Conformal time and null geodesics*

The causal structure of the universe is determined by the propagation of light in the FRW spacetime (1). Light follows null geodesics,  $ds^2 = 0$ . These are studied most easily if we define conformal time<sup>2</sup>

$$\tau = \int \frac{dt}{a(t)} , \quad (5)$$

for which the FRW metric becomes

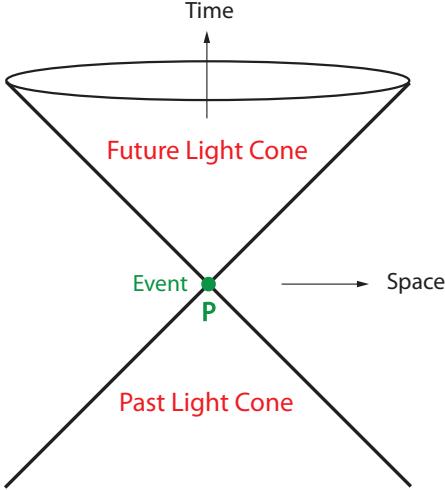
$$ds^2 = a(\tau)^2 [-d\tau^2 + (\mathrm{d}\chi^2 + \Phi_k(\chi^2)(\mathrm{d}\theta^2 + \sin^2 \theta \mathrm{d}\phi^2))] . \quad (6)$$

In an isotropic universe we may consider radial propagation of light as determined by the two-dimensional line element

$$ds^2 = a(\tau)^2 [-d\tau^2 + d\chi^2] . \quad (7)$$

The metric has factorized into a static Minkowski metric multiplied by a time-dependent conformal factor  $a(\tau)$ . The radial null geodesics of light therefore satisfy

$$\chi(\tau) = \pm\tau + \text{const} , \quad (8)$$



**Figure 3:** Light cones and causality. Photons travel along world lines of zero proper time,  $ds^2 = 0$ , called *null geodesics*. Massive particles travel along world lines with real proper time,  $ds^2 > 0$ , called *timelike geodesics*. Causally disconnected regions of spacetime are separated by *spacelike* intervals,  $ds^2 < 0$ . The set of all null geodesics passing through a given point (or *event*) in spacetime is called the *light cone*. The interior of the light cone, consisting of all null and timelike geodesics, defined the region of spacetime causally related to that event.

i.e. they correspond to straight lines at angles  $\pm 45^\circ$  in the  $\tau-\chi$  plane. If instead we had used physical time  $t$  to study light propagation, then the light cones for curved spacetimes would be curved.

#### Particle horizon

The maximum comoving distance light can propagate between an initial time  $t_i$  and some later time  $t$  is

$$\chi_p(\tau) = \tau - \tau_i = \int_{t_i}^t \frac{dt}{a(t)} . \quad (9)$$

This is called the (comoving) particle horizon. The initial time  $t_i$  is often taken to be the ‘origin of the universe’,  $t_i \equiv 0$ . The physical size of the particle horizon is

$$d_p(t) = a(t)\chi_p . \quad (10)$$

The particle horizon is of crucial importance to understanding the causal structure of the universe and it will be fundamental to our discussion of inflation. As we will see, the conventional Big Bang model ‘begins’ at a finite time in the past and at any time in the past the particle horizon was finite, limiting the distance over which spacetime region could have been in causal contact. This feature is at the heart of the ‘big bang puzzles’.

#### Event horizon

An event horizon defines the set of points from which signals sent at a given moment of time  $\tau$  will never be received by an observer in the future. In comoving coordinates these points satisfy

$$\chi > \chi_e = \int_{\tau}^{\tau_{\max}} d\tau = \tau_{\max} - \tau , \quad (11)$$

---

<sup>2</sup>Conformal time may be interpreted as a “clock” which slows down with the expansion of the universe.

where  $\tau_{\max}$  denotes the ‘final moment of time’ (this might be infinite or finite). The physical size of the event horizon is

$$d_e(t) = a(t)\chi_e. \quad (12)$$

### Angular diameter and luminosity distance

Conformal time or comoving distance relate in a simple way to angular diameter and luminosity distances that are important for the discussion CMB anisotropies and supernova distances, respectively. These details won’t concern us here, but may be found in the standard books [5–7].

### 3.3 Dynamics: Einstein Equations

The dynamics of the universe as characterized by the evolution of the scale factor of the FRW spacetime  $a(t)$  is determined by the Einstein equations

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}. \quad (13)$$

We will often work in units where  $8\pi G \equiv 1$ . For convenience we here recall the definition of the Einstein tensor

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R, \quad (14)$$

in terms of the Ricci tensor  $R_{\mu\nu}$  and the Ricci scalar  $R$ ,

$$R_{\mu\nu} = \Gamma_{\mu\nu,\alpha}^\alpha - \Gamma_{\mu\alpha,\nu}^\alpha + \Gamma_{\beta\alpha}^\alpha \Gamma_{\mu\nu}^\beta - \Gamma_{\beta\nu}^\alpha \Gamma_{\mu\alpha}^\beta, \quad R \equiv g^{\mu\nu} R_{\mu\nu}, \quad (15)$$

where

$$\Gamma_{\alpha\beta}^\mu \equiv \frac{g^{\mu\nu}}{2} [g_{\alpha\nu,\beta} + g_{\beta\nu,\alpha} - g_{\alpha\beta,\nu}]. \quad (16)$$

Commas denote partial derivatives, *e.g.*  $(\dots)_{,\mu} = \frac{\partial(\dots)}{\partial x^\mu}$ . We will continue to follow this notation in the rest of these lectures.

For the case of a perfect fluid the stress-energy tensor is

$$T_\nu^\mu = g^{\mu\alpha} T_{\alpha\nu} = (\rho + p) u^\mu u_\nu - p \delta_\nu^\mu, \quad (17)$$

where  $\rho$  and  $p$  are the proper energy density and pressure in the fluid rest frame and  $u^\mu$  is the 4-velocity of the fluid. The 4-velocity satisfies  $g_{\mu\nu} u^\mu u^\nu = -1$ . In a frame that is comoving with the fluid we may choose  $u^\mu = (1, 0, 0, 0)$ , *i.e.*

$$T_\nu^\mu = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & -p & 0 & 0 \\ 0 & 0 & -p & 0 \\ 0 & 0 & 0 & -p \end{pmatrix}. \quad (18)$$

The Einstein equations then take the form of two coupled, non-linear ordinary differential equations, the *Friedmann equations* (see **Problem 3**)

$$H^2 \equiv \left( \frac{\dot{a}}{a} \right)^2 = \frac{1}{3}\rho - \frac{k}{a^2},$$

(19)

and

$$\boxed{\dot{H} + H^2 = \frac{\ddot{a}}{a} = -\frac{1}{6}(\rho + 3p)} . \quad (20)$$

Here we defined the Hubble parameter  $H \equiv \partial_t \ln a$  as a measure of the expansion rate of the universe. Overdots denote derivatives with respect to physical time  $t$ . Notice, that in an expanding universe (*i.e.*  $\dot{a} > 0 \forall t$ ) filled with ordinary matter (*i.e.* matter satisfying the strong energy condition:  $\rho + 3p \geq 0$ ) equation (20) implies  $\ddot{a} < 0 \forall t$ . This indicates the existence of a singularity in the finite past:  $a(t \equiv 0) = 0$ . Of course, this conclusion relies on the assumption that general relativity and the Friedmann equations are applicable up to arbitrary high energies and that no exotic forms of matter become relevant at high energies. More likely the singularity signals the breakdown of general relativity.

Equations (19) and (20) may be combined into the continuity equation

$$\boxed{\frac{d\rho}{dt} + 3H(\rho + p) = 0} . \quad (21)$$

This may also be written as

$$\frac{d \ln \rho}{d \ln a} = -3(1+w) , \quad (22)$$

if we define the equation of state parameter

$$w \equiv \frac{p}{\rho} . \quad (23)$$

Eqn. (22) may be integrated to give

$$\rho \propto a^{-3(1+w)} . \quad (24)$$

Together with the Friedmann Equation (19) this leads to the time evolution of the scale factor

$$a \propto \begin{cases} t^{2/3(1+w)} & w \neq -1 , \\ e^{Ht} & w = -1 . \end{cases} \quad (25)$$

*i.e.*  $a(t) \propto t^{2/3}$ ,  $a(t) \propto t^{1/2}$  and  $a(t) \propto \exp(Ht)$ , for the scale factor of a flat ( $k = 0$ ) universe dominated by non-relativistic matter ( $w = 0$ ), radiation or relativistic matter ( $w = \frac{1}{3}$ ) and a cosmological constant ( $w = -1$ ), respectively.

**Table 2:** FRW solutions for a flat universe dominated by radiation, matter or a cosmological constant.

	$w$	$\rho(a)$	$a(t)$	$a(\tau)$
MD	0	$a^{-3}$	$t^{2/3}$	$\tau^2$
RD	$\frac{1}{3}$	$a^{-4}$	$t^{1/2}$	$\tau$
$\Lambda$	-1	$a^0$	$e^{Ht}$	$-\tau^{-1}$

If more than one matter species (baryons, photons, neutrinos, dark matter, dark energy, etc.) contributes significantly to the energy density and the pressure,  $\rho$  and  $p$  refer to the sum of all components

$$\rho \equiv \sum_i \rho_i , \quad p \equiv \sum_i p_i . \quad (26)$$

For each species  $i$  we define the *present* ratio of the energy density relative to the *critical energy density*  $\rho_{\text{crit}} \equiv 3H_0^2$

$$\Omega_i \equiv \frac{\rho_0^i}{\rho_{\text{crit}}} , \quad (27)$$

and the corresponding equations of state

$$w_i \equiv \frac{p_i}{\rho_i} . \quad (28)$$

Here and in the following the subscript '0' denotes evaluation of a quantity at the present time  $t_0$ . We normalize the scale factor such that  $a_0 = a(t_0) \equiv 1$ . This allows one to write the Friedmann equation (19) as

$$\left(\frac{H}{H_0}\right)^2 = \sum_i \Omega_i a^{-3(1+w_i)} + \Omega_k a^{-2}, \quad (29)$$

with  $\Omega_k \equiv -k/a_0^2 H_0^2$ , which implies the following consistency relation

$$\sum_i \Omega_i + \Omega_k = 1 . \quad (30)$$

The second Friedmann Equation (20) evaluated at  $t = t_0$  becomes

$$\frac{1}{a_0 H_0^2} \frac{d^2 a_0}{dt^2} = -\frac{1}{2} \sum_i \Omega_i (1 + 3w_i) . \quad (31)$$

Observations of the cosmic microwave background (CMB) and the large-scale structure (LSS) find that the universe is flat ( $\Omega_k \sim 0$ ) and composed of 4% atoms, 23% (cold) dark matter and 73% dark energy:  $\Omega_b = 0.04$ ,  $\Omega_d = 0.23$ ,  $\Omega_\Lambda = 0.72$ , with  $w_\Lambda \approx -1$ . It is also found that the universe has tiny ripples of adiabatic, scale-invariant, Gaussian density fluctuations. In the bulk of this lecture series I will describe how quantum fluctuations during inflation can explain the observed cosmological perturbations.

## 4 Big Bang Puzzles

It is somewhat of a philosophical question whether initial conditions form part of a physical theory or should be considered separately. The purpose of physics is to predict the future evolution of a system given a set of initial conditions (*e.g.* Newton's laws of gravity will predict the path of a projectile if we define its initial position and velocity.). It is therefore far from clear whether cosmology should predict or even just explain the initial conditions of the universe. On the other hand, it would be very disappointing if only very special and finely-tuned initial conditions would lead to the universe as we see it, making the observed universe a very 'improbable accident'.

In this section we will explain that the conventional Big Bang theory requires precisely such a fine-tuned set of initial conditions to allow the universe to evolve to its current state. One of the major achievements of inflation is that it explains the initial conditions of the universe. Via inflation, the universe we know and love grew out of generic initial conditions.

## 4.1 The Cauchy Problem of the Universe

To specify the initial condition of the universe we consider a spatial slice of constant time  $\Sigma$  (we here won't worry about the gauge-dependence of the choice of  $\Sigma$ ). On  $\Sigma$  we define the positions and velocities of all matter particles. The laws of gravity and fluid dynamics are then used to evolve the system forward in time.

- Initial Homogeneity

We describe the spatial distribution of matter by its density and pressure as a function of coordinate  $\mathbf{x}$ , *i.e.*  $\rho(\mathbf{x})$  and  $p(\mathbf{x})$ . In the previous section we assumed homogeneity and isotropy of the universe. Why is this a good assumption? Inhomogeneities are gravitationally unstable and therefore grow with time. Observations of the cosmic microwave background show that the inhomogeneities were much smaller in the past (at last-scattering) than today. One thus expects that these inhomogeneities were even smaller at earlier times. How do we explain the smoothness of the early universe?

This is particularly surprising since we will show in §4.2 that in the conventional Big Bang picture the early universe (*e.g.* the CMB at last-scattering) consisted of a large number of causally-disconnected regions of space. In the Big Bang theory, there is no dynamical reason to explain why these causally-separated patches show such similar physical conditions. The homogeneity problem is therefore often called the *horizon problem*.

- Initial Velocities

In addition to specifying the initial density distribution, the complete characterization of the Cauchy problem of the universe requires the fluid velocities at every point in space. As we will see, to ensure that the universe remains homogeneous at late times requires the initial fluid velocities to take very precise values. If the initial velocities are just slightly too small, the universe recollapses within a fraction of a second. If they are just slightly too big, the universe expands too rapidly and quickly becomes nearly empty. The fine-tuning of initial velocities is made more dramatic by considering it in combination with the horizon problem. The fluid velocities need to be fine-tuned across causally-separated regions of space.

Since the difference between the potential energy and the kinetic energy defines the local curvature of a region of space (see **Exercise 1**), this fine-tuning of initial velocities is often called the *flatness problem*.

## 4.2 Horizon Problem

In the previous section, we defined the *comoving (particle) horizon*,  $\tau$ , as the causal horizon or the maximum distance a light ray can travel between time 0 and time  $t$

$$\boxed{\tau \equiv \int_0^t \frac{dt'}{a(t')} = \int_0^a \frac{da}{Ha^2} = \int_0^a d \ln a \left( \frac{1}{aH} \right)}. \quad (32)$$

Here, we have expressed the comoving horizon as an integral of the *comoving Hubble radius*,  $(aH)^{-1}$ ,

$$(aH)^{-1} = H_0^{-1} a^{\frac{1}{2}(1+w)}. \quad (33)$$

The comoving Hubble radius plays a crucial role in inflation. During the conventional Big Bang expansion ( $w \gtrsim 0$ )  $(aH)^{-1}$  grows monotonically and the comoving horizon  $\tau$  or the *fraction of the universe in causal contact increases with time*

$$\tau \propto a^{\frac{1}{2}(1+3w)}. \quad (34)$$

In particular, for radiation-dominated (RD) and matter-dominated (MD) universes we find

$$\tau = \int_0^a \frac{da}{Ha^2} \propto \begin{cases} a & \text{RD} \\ a^{1/2} & \text{MD} \end{cases}. \quad (35)$$

This means that the comoving horizon grows monotonically with time which implies that comoving scales entering the horizon today have been far outside the horizon at CMB decoupling. But the near-homogeneity of the CMB tells us that the universe was extremely homogeneous at the time of last-scattering on scales encompassing many regions that *a priori* are causally independent. How is this possible?

### 4.3 Flatness Problem

**Exercise 1 (Flatness and Kinetic Energy)** *Show that the curvature parameter*

$$\Omega_k \equiv \Omega - 1 = \frac{\rho - \rho_{\text{crit}}}{\rho_{\text{crit}}}, \quad \text{where } \rho_{\text{crit}} \equiv 3H^2$$

*may be interpreted as the difference between the average potential energy and the average kinetic energy of a region of space.*

Spacetime in general relativity is dynamical, curving in response to matter in the universe. Why then is the universe so closely approximated by flat Euclidean space? To quantify the problem we consider the Friedmann equation

$$H^2 = \frac{1}{3}\rho(a) - \frac{k}{a^2}, \quad (36)$$

written as

$$\boxed{1 - \Omega(a) = \frac{-k}{(aH)^2}}, \quad (37)$$

where

$$\Omega(a) \equiv \frac{\rho(a)}{\rho_{\text{crit}}(a)}, \quad \rho_{\text{crit}}(a) \equiv 3H(a)^2. \quad (38)$$

Notice that  $\Omega(a)$  is now defined to be time-dependent, whereas the  $\Omega$ 's in the previous sections were constants,  $\Omega(a_0)$ . In standard cosmology the comoving Hubble radius,  $(aH)^{-1}$ , grows with time and from equation (37) the quantity  $|\Omega - 1|$  must thus diverge with time.  $\Omega = 1$  is an *unstable fixed point*. Therefore, in standard Big Bang cosmology without inflation, the near-flatness observed today,  $\Omega(a_0) \sim 1$ , requires an extreme fine-tuning of  $\Omega$  close to 1 in the early universe. More specifically, one finds that the deviation from flatness at Big Bang nucleosynthesis (BBN), during the GUT era and at the Planck scale, respectively has to satisfy the following conditions

$$|\Omega(a_{\text{BBN}}) - 1| \leq \mathcal{O}(10^{-16}) \quad (39)$$

$$|\Omega(a_{\text{GUT}}) - 1| \leq \mathcal{O}(10^{-55}) \quad (40)$$

$$|\Omega(a_{\text{pl}}) - 1| \leq \mathcal{O}(10^{-61}). \quad (41)$$

Another way of understanding the flatness problem is from the following differential equation

$$\frac{d\Omega}{d \ln a} = (1 + 3w)\Omega(\Omega - 1). \quad (42)$$

Equation (42) is derived by differentiating (37) and using the continuity equation (22). This makes it apparent that  $\Omega = 1$  is an unstable fixed point if the strong energy condition is satisfied

$$\frac{d|\Omega - 1|}{d \ln a} > 0 \Leftrightarrow 1 + 3w > 0. \quad (43)$$

Again, why is  $\Omega(a_0) \sim \mathcal{O}(1)$  and not much smaller or much larger?

## 4.4 On the Problem of Initial Conditions

We should emphasize that the flatness and horizon problems are *not* strict inconsistencies in the standard cosmological model. If one assumes that the initial value of  $\Omega$  was extremely close to unity and that the universe began homogeneously over superhorizon distances (but with just the right level of inhomogeneity to explain structure formation) then the universe will continue to evolve homogeneously in agreement with observations. The flatness and horizon problems are therefore really just severe shortcomings in the predictive power of the Big Bang model. The dramatic flatness of the early universe cannot be predicted by the standard model, but must instead be assumed in the initial conditions. Likewise, the striking large-scale homogeneity of the universe is not explained or predicted by the model, but instead must simply be assumed. A theory that explains these initial conditions dynamically seems very attractive.

# 5 A First Look at Inflation

## 5.1 The Shrinking Hubble Sphere

In §4.2 and §4.3 we emphasized the fundamental role of the comoving Hubble radius,  $(aH)^{-1}$ , in the horizon and flatness problems of the standard Big Bang cosmology. Both problems arise since in the conventional cosmology the comoving Hubble radius is strictly increasing. This suggest that all the Big Bang puzzles are solved by a beautifully simple idea: *invert the behavior of the comoving Hubble radius*, i.e. make it *decrease sufficiently in the very early universe*.

### 5.1.1 Comoving Horizon during Inflation

The evolution of the comoving horizon is of such crucial importance to the whole idea of inflation that it is worth being explicit about a few important points.

Recall the definition of the comoving horizon (= conformal time) as a logarithmic integral of the comoving Hubble radius

$$\boxed{\tau = \int_0^a d \ln a' \frac{1}{a' H(a')}}. \quad (44)$$

Let us emphasize a subtle distinction between the comoving horizon  $\tau$  and the comoving Hubble radius  $(aH)^{-1}$  [5]:

If particles are separated by distances greater than  $\tau$ , they *never* could have communicated with one another; if they are separated by distances greater than  $(aH)^{-1}$ , they cannot talk to each other *now!* This distinction is crucial for the solution to the horizon problem which relies on the following: It is possible that  $\tau$  is much larger than  $(aH)^{-1}$  now, so that particles cannot communicate today but were in causal contact early on. From equation (44) we see that this might happen if the comoving Hubble radius in the early universe was much larger than it is now so that  $\tau$  got most of its contribution from early times. Hence, we require a phase of decreasing Hubble radius. Since  $H$  is approximately constant while  $a$  grows exponentially during inflation we find that the comoving Hubble radius decreases during inflation just as advertised.

Besides solving the Big Bang puzzles the decreasing comoving horizon during inflation is the key feature required for the quantum generation of cosmological perturbations described in the second lecture. I will describe how quantum fluctuations are generated on subhorizon scales, but exit the horizon once the Hubble radius becomes smaller than their comoving wavelength. In physical coordinates this corresponds to the superluminal expansion stretching perturbations to acausal distances. They become classical superhorizon density perturbations which re-enter the horizon in the subsequent Big Bang evolution and then gravitationally collapse to form the large-scale structure in the universe.

With this understanding of how the comoving horizon and the comoving Hubble radius evolve during inflation it is now almost trivial to explain how inflation solves the Big Bang puzzles.

### 5.1.2 Flatness Problem Revisited

Recall the Friedmann Equation (37)

$$|1 - \Omega(a)| = \frac{1}{(aH)^2}. \quad (45)$$

If the comoving Hubble radius decreases this *drives the universe toward flatness* (rather than away from it). This solves the flatness problem!  $\Omega = 1$  is an attractor during inflation.

### 5.1.3 Horizon Problem Revisited

A decreasing comoving horizon means that large scales entering the present universe were inside the horizon before inflation (see Figure 2). Causal physics before inflation therefore established spatial homogeneity. The uniformity of the CMB is not a mystery.

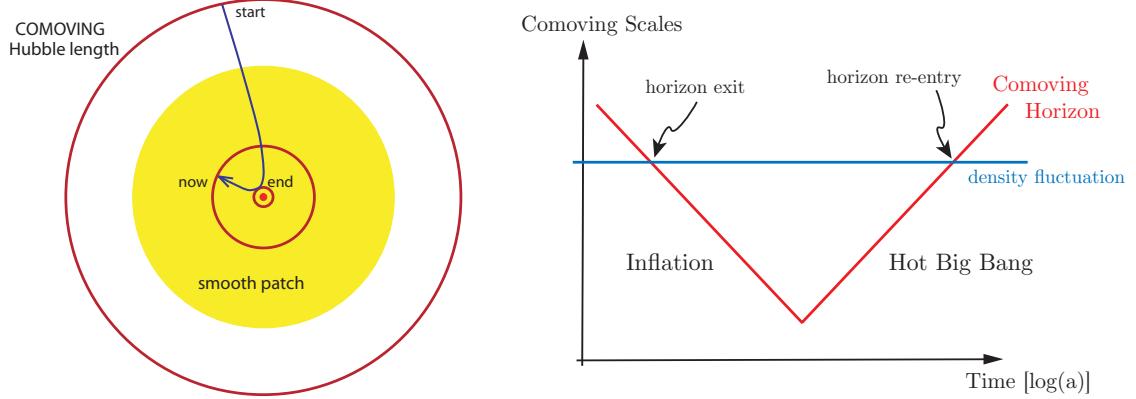
## 5.2 Conditions for Inflation

Via the Friedmann Equations a shrinking comoving Hubble radius can be related to the acceleration and the pressure of the universe

$$\frac{d}{dt} \left( \frac{H^{-1}}{a} \right) < 0 \quad \Rightarrow \quad \frac{d^2a}{dt^2} > 0 \quad \Rightarrow \quad \rho + 3p < 0.$$

(46)

The three equivalent conditions for inflation therefore are:



**Figure 4:** *Left:* Evolution of the comoving Hubble radius,  $(aH)^{-1}$ , in the inflationary universe. The comoving Hubble sphere shrinks during inflation and expands after inflation. Inflation is therefore a mechanism to ‘zoom-in’ on a smooth initial subhorizon patch. *Right:* Solution of the horizon problem. Scales of cosmological interest were larger than the Hubble radius until  $a \sim 10^{-5}$ . However, very early on, before inflation operated, all scales of interest were smaller than the Hubble radius and therefore susceptible to microphysical processing. Similarly, at very late time, scales of cosmological interest came back within the Hubble radius.

- **Decreasing comoving horizon**

The shrinking Hubble sphere is defined as

$$\frac{d}{dt} \left( \frac{1}{aH} \right) < 0. \quad (47)$$

We used this as our fundamental definition of inflation since it most directly relates to the flatness and horizon problems.

- **Accelerated expansion**

From the relation

$$\frac{d}{dt} (aH)^{-1} = \frac{-\ddot{a}}{(aH)^2} \quad (48)$$

we see immediately that a shrinking comoving Hubble radius implies accelerated expansion

$$\frac{d^2 a}{dt^2} > 0. \quad (49)$$

This explains why inflation is often defined as a period of accelerated expansion. The second time derivative of the scale factor may of course be related to the first time derivative of the Hubble parameter  $H$

$$\frac{\ddot{a}}{a} = H^2(1 - \varepsilon), \quad \text{where } \varepsilon \equiv -\frac{\dot{H}}{H^2}. \quad (50)$$

Acceleration therefore corresponds to

$$\varepsilon = -\frac{\dot{H}}{H^2} = -\frac{d \ln H}{dN} < 1$$

(51)

Here, we have defined  $dN = Hdt = d\ln a$ , which measures the number of  $e$ -folds  $N$  of inflationary expansion. Eqn. (51) therefore means that the fractional change of the Hubble parameter per  $e$ -fold is small.

- **Negative pressure**

What stress-energy can source acceleration? Consulting equation (20) we infer that  $\ddot{a} > 0$  requires

$$p < -\frac{1}{3}\rho, \quad (52)$$

*i.e.* negative pressure or a violation of the strong energy condition. How this can arise in a physical theory will be explained in §6.2.

### 5.3 Conformal Diagram of Inflation

A truly illuminating way of visualizing inflation is with the aid of a conformal spacetime diagram. Recall from §3 the flat FRW metric in conformal time  $d\tau = dt/a(t)$

$$ds^2 = a^2(\tau) [-d\tau^2 + d\mathbf{x}^2]. \quad (53)$$

Also recall that in conformal coordinates null geodesics ( $ds^2 = 0$ ) are always at  $45^\circ$  angles,  $d\tau = \pm\sqrt{d\mathbf{x}^2} \equiv \pm dr$ . Since light determines the causal structure of spacetime this provides a nice way to study horizons in inflationary cosmology.

During matter or radiation domination the scale factor evolves as

$$a(\tau) \propto \begin{cases} \tau & \text{RD} \\ \tau^2 & \text{MD} \end{cases} \quad (54)$$

If the universe had always been dominated by matter or radiation, this would imply the existence of the *Big Bang singularity* at  $\tau = 0$

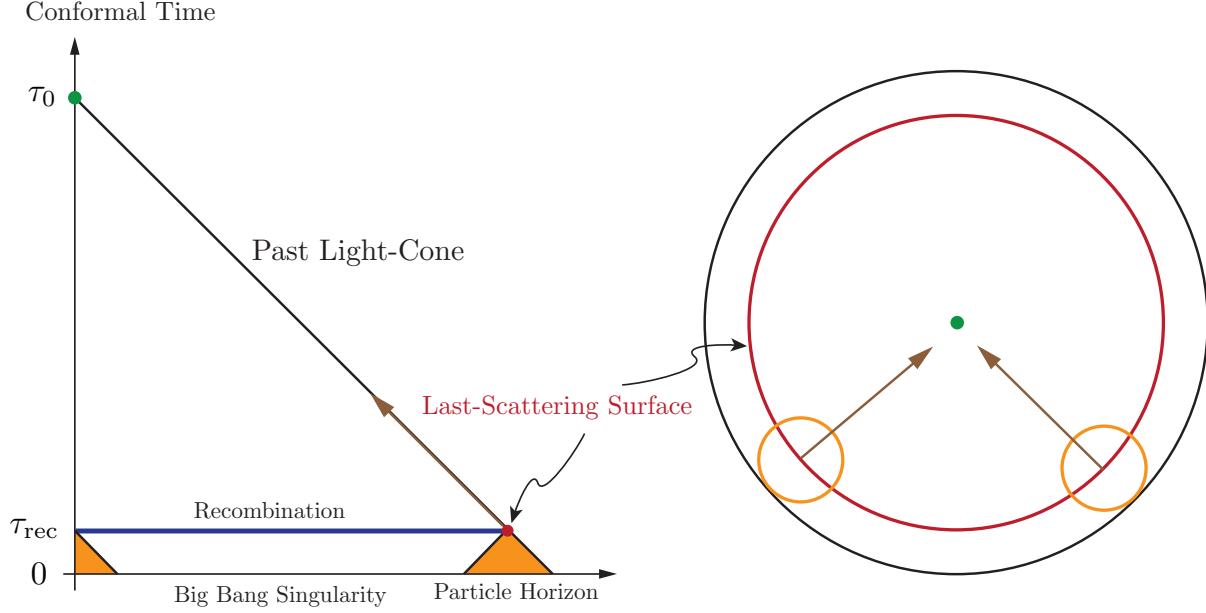
$$a(\tau \equiv 0) = 0. \quad (55)$$

The conformal diagram corresponding to standard Big Bang cosmology is given in Figure 5. The horizon problem is apparent. Each spacetime point in the conformal diagram has an associated past light cone which defines its causal past. Two points on a given  $\tau = \text{constant}$  surface are in causal contact if their past light cones intersect at the Big Bang,  $\tau = 0$ . This means that the surface of last-scattering ( $\tau_{\text{CMB}}$ ) consisted of many causally disconnected regions that won't be in thermal equilibrium. The uniformity of the CMB on large scales hence becomes a serious puzzle.

During inflation ( $H \approx \text{const.}$ ), the scale factor is

$$a(\tau) = -\frac{1}{H\tau}, \quad (56)$$

and the singularity,  $a = 0$ , is pushed to the infinite past,  $\tau \rightarrow -\infty$ . The scale factor (56) becomes infinite at  $\tau = 0$ ! This is because we have assumed de Sitter space with  $H = \text{const.}$ , which means that inflation will continue forever with  $\tau = 0$  corresponding to the infinite future  $t \rightarrow +\infty$ . In reality, inflation ends at some finite time, and the approximation (56) although valid at early times, breaks down near the end of inflation. So the surface  $\tau = 0$  is not the Big Bang, but the end of inflation. The initial singularity has been pushed back arbitrarily far in conformal time  $\tau \ll 0$ , and light cones can extend through the apparent Big Bang so that apparently disconnected points are in causal contact. This is summarized in the conformal diagram in Figure 6.



**Figure 5:** Conformal diagram of Big Bang cosmology. The CMB at last-scattering (recombination) consists of  $10^5$  causally disconnected regions!

## 6 The Physics of Inflation

Inflation is a very unfamiliar physical phenomenon: within a fraction a second the universe grew exponential at an accelerating rate. In Einstein gravity this requires a negative pressure source or a nearly constant energy density. In this section we describe the physical conditions under which this can arise.

### 6.1 Scalar Field Dynamics

The simplest models of inflation involve a single scalar field  $\phi$ , the *inflaton*. The dynamics of a scalar field (minimally) coupled to gravity is governed by the action

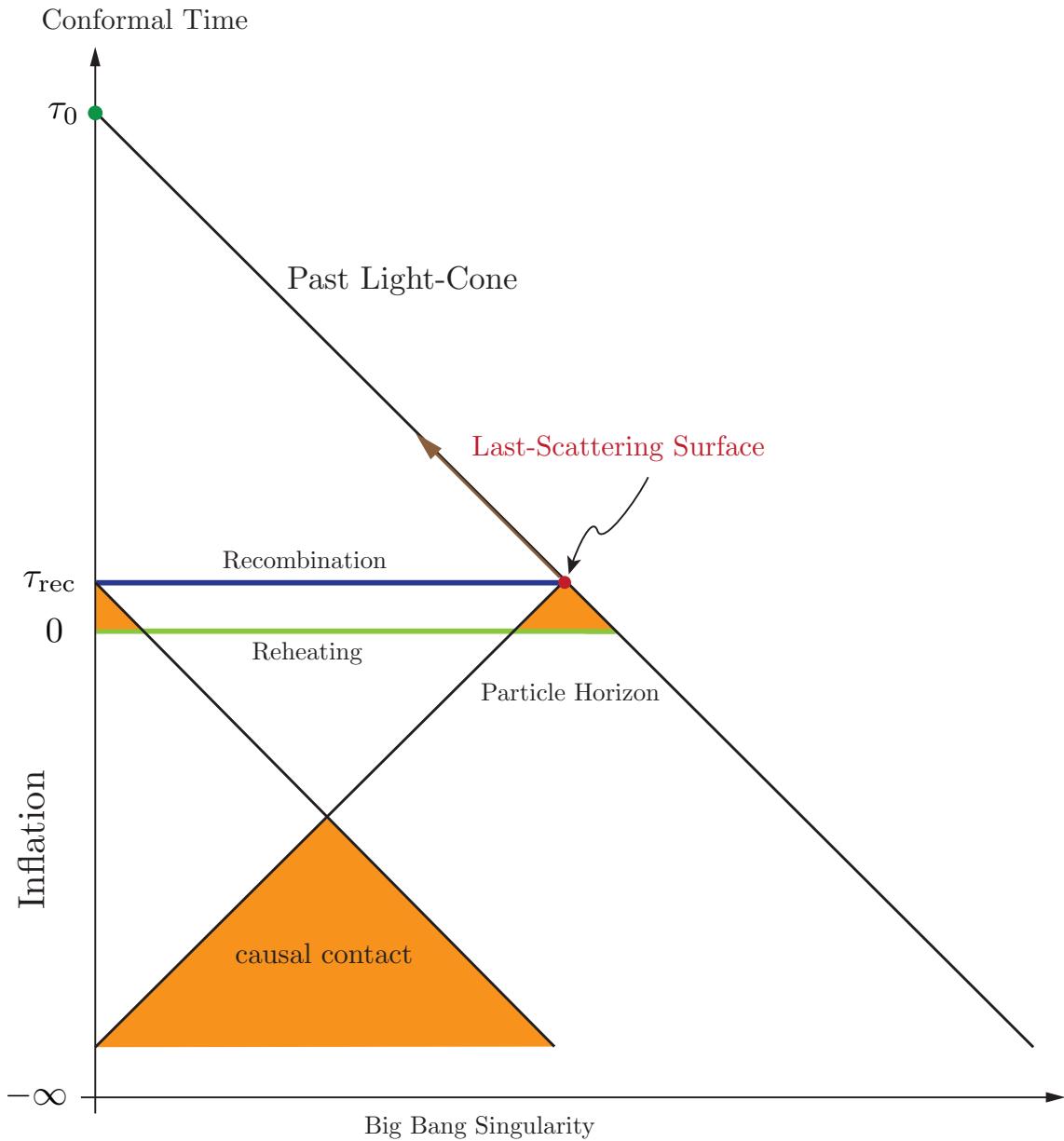
$$S = \int d^4x \sqrt{-g} \left[ \frac{1}{2}R + \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi) \right] = S_{EH} + S_\phi. \quad (57)$$

The action (57) is the sum of the gravitational Einstein-Hilbert action  $S_{EH}$  and the action of a scalar field  $S_\phi$  with canonical kinetic term. The potential  $V(\phi)$  describes the self-interaction of the scalar field. The energy-momentum tensor for the scalar field is

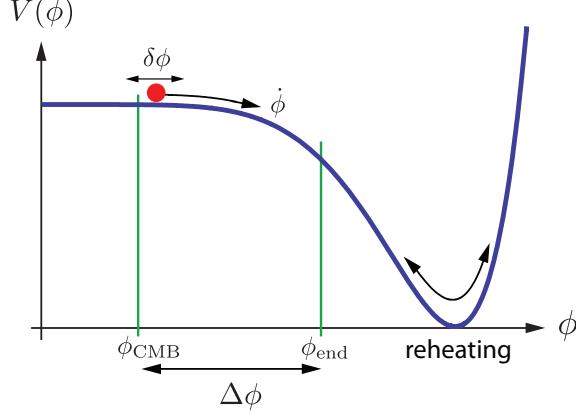
$$T_{\mu\nu}^{(\phi)} \equiv -\frac{2}{\sqrt{-g}}\frac{\delta S_\phi}{\delta g^{\mu\nu}} = \partial_\mu\phi\partial_\nu\phi - g_{\mu\nu} \left( \frac{1}{2}\partial^\sigma\phi\partial_\sigma\phi + V(\phi) \right). \quad (58)$$

The field equation of motion is

$$\frac{\delta S_\phi}{\delta\phi} = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}\partial^\mu\phi) + V_{,\phi} = 0. \quad (59)$$



**Figure 6:** Conformal diagram of inflationary cosmology. Inflation extends conformal time to negative values! The end of inflation creates an “apparent” Big Bang at  $\tau = 0$ . There is, however, no singularity at  $\tau = 0$  and the light cones intersect at an earlier time if inflation lasts for at least 60  $e$ -folds.



**Figure 7:** Example of an inflaton potential. Acceleration occurs when the potential energy of the field,  $V(\phi)$ , dominates over its kinetic energy,  $\frac{1}{2}\dot{\phi}^2$ . Inflation ends at  $\phi_{\text{end}}$  when the kinetic energy has grown to become comparable to the potential energy,  $\frac{1}{2}\dot{\phi}^2 \approx V$ . CMB fluctuations are created by quantum fluctuations  $\delta\phi$  about 60 e-folds before the end of inflation. At reheating, the energy density of the inflaton is converted into radiation.

Assuming the FRW metric (1) for  $g_{\mu\nu}$  and restricting to the case of a homogeneous field  $\phi(t, \mathbf{x}) \equiv \phi(t)$ , the scalar energy-momentum tensor takes the form of a perfect fluid (18) with

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi), \quad (60)$$

$$p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi). \quad (61)$$

The resulting equation of state

$$w_\phi \equiv \frac{p_\phi}{\rho_\phi} = \frac{\frac{1}{2}\dot{\phi}^2 - V}{\frac{1}{2}\dot{\phi}^2 + V} \quad (62)$$

shows that a scalar field can lead to negative pressure ( $w_\phi < 0$ ) and accelerated expansion ( $w_\phi < -1/3$ ) if the potential energy  $V$  dominates over the kinetic energy  $\frac{1}{2}\dot{\phi}^2$ . The dynamics of the (homogeneous) scalar field and the FRW geometry is determined by

$$\ddot{\phi} + 3H\dot{\phi} + V_{,\phi} = 0 \quad \text{and} \quad H^2 = \frac{1}{3} \left( \frac{1}{2}\dot{\phi}^2 + V(\phi) \right). \quad (63)$$

## 6.2 Slow-Roll Inflation

The acceleration equation for a universe dominated by a homogeneous scalar field can be written as follows

$$\frac{\ddot{a}}{a} = -\frac{1}{6}(\rho_\phi + 3p_\phi) = H^2(1 - \varepsilon), \quad (64)$$

where

$$\varepsilon \equiv \frac{3}{2}(w_\phi + 1) = \frac{1}{2} \frac{\dot{\phi}^2}{H^2}. \quad (65)$$

The so-called *slow-roll parameter*  $\varepsilon$  may be related to the evolution of the Hubble parameter

$$\boxed{\varepsilon = -\frac{\dot{H}}{H^2}} = -\frac{d \ln H}{dN}, \quad (66)$$

where  $dN = Hdt$ . Accelerated expansion occurs if  $\varepsilon < 1$ . The de Sitter limit,  $p_\phi \rightarrow -\rho_\phi$ , corresponds to  $\varepsilon \rightarrow 0$ . In this case, the potential energy dominates over the kinetic energy

$$\dot{\phi}^2 \ll V(\phi). \quad (67)$$

Accelerated expansion will only be sustained for a sufficient period of time if the second time derivative of  $\phi$  is small enough

$$|\ddot{\phi}| \ll |3H\dot{\phi}|, |V_{,\phi}|. \quad (68)$$

This requires smallness of a second slow-roll parameter

$$\eta = -\frac{\ddot{\phi}}{H\dot{\phi}} = \varepsilon - \frac{1}{2\varepsilon} \frac{d\varepsilon}{dN}. \quad (69)$$

The slow-roll conditions,  $\varepsilon, |\eta| < 1$ , may also be expressed as conditions on the shape of the inflationary potential

$$\boxed{\epsilon_v(\phi) \equiv \frac{M_{\text{pl}}^2}{2} \left( \frac{V_{,\phi}}{V} \right)^2}, \quad (70)$$

and

$$\boxed{\eta_v(\phi) \equiv M_{\text{pl}}^2 \frac{V_{,\phi\phi}}{V}}. \quad (71)$$

Here, we temporarily reintroduced the Planck mass to make  $\epsilon_v$  and  $\eta_v$  manifestly dimensionless. In the following we will set  $M_{\text{pl}}$  to one again. In the slow-roll regime

$$\epsilon_v, |\eta_v| \ll 1 \quad (72)$$

the background evolution is

$$H^2 \approx \frac{1}{3}V(\phi) \approx \text{const}, \quad (73)$$

$$\dot{\phi} \approx -\frac{V_{,\phi}}{3H}, \quad (74)$$

and the spacetime is approximately *de Sitter*

$$a(t) \sim e^{Ht}, \quad \text{where } H \approx \text{const}. \quad (75)$$

The parameters  $\epsilon_v$  and  $\eta_v$  are called the *potential slow-roll parameters* to distinguish them from the *Hubble slow-roll parameters*  $\varepsilon$  and  $\eta$ . In the slow-roll approximation the Hubble and potential slow-roll parameters are related as follows

$$\varepsilon \approx \epsilon_v, \quad \eta \approx \eta_v - \epsilon_v. \quad (76)$$

Inflation ends when the slow-roll conditions (72) are violated

$$\varepsilon(\phi_{\text{end}}) \equiv 1, \quad \epsilon_v(\phi_{\text{end}}) \approx 1. \quad (77)$$

The number of *e*-folds before inflation ends is

$$\begin{aligned} N(\phi) &\equiv \ln \frac{a_{\text{end}}}{a} \\ &= \int_t^{t_{\text{end}}} H dt = \int_\phi^{\phi_{\text{end}}} \frac{H}{\dot{\phi}} d\phi \approx \int_{\phi_{\text{end}}}^\phi \frac{V}{V_{,\phi}} d\phi, \end{aligned} \quad (78)$$

where we used the slow-roll results (73) and (74). The result (78) may also be written as

$$N(\phi) = \int_{\phi_{\text{end}}}^{\phi} \frac{d\phi}{\sqrt{2\varepsilon}} \approx \int_{\phi_{\text{end}}}^{\phi} \frac{d\phi}{\sqrt{2\epsilon_v}}. \quad (79)$$

To solve the horizon and flatness problems requires that the total number of inflationary  $e$ -folds exceeds about 60

$$N_{\text{tot}} \equiv \ln \frac{a_{\text{end}}}{a_{\text{start}}} \gtrsim 60. \quad (80)$$

The precise value depends on the details of reheating after inflation. The fluctuations observed in the CMB are created  $N_{\text{cmb}} \approx 40 - 60$   $e$ -folds before the end of inflation (the precise value again depending on the details of reheating and the post-inflationary thermal history of the universe). The following integral constraint gives the corresponding field value  $\phi_{\text{cmb}}$

$$\int_{\phi_{\text{end}}}^{\phi_{\text{cmb}}} \frac{d\phi}{\sqrt{2\epsilon_v}} = N_{\text{cmb}} \approx 40 - 60. \quad (81)$$

### 6.3 Case Study: $m^2\phi^2$ Inflation

As an example, let us give the slow-roll analysis of arguably the simplest model of inflation: single field inflation driven by a mass term

$$V(\phi) = \frac{1}{2}m^2\phi^2. \quad (82)$$

The slow-roll parameters are

$$\epsilon_v(\phi) = \eta_v(\phi) = 2 \left( \frac{M_{\text{pl}}}{\phi} \right)^2. \quad (83)$$

To satisfy the slow-roll conditions  $\epsilon_v, |\eta_v| < 1$ , we need to consider super-Planckian values for the inflaton

$$\phi > \sqrt{2}M_{\text{pl}} \equiv \phi_{\text{end}}. \quad (84)$$

The relation between the inflaton field value and the number of  $e$ -folds before the end of inflation is

$$N(\phi) = \frac{\phi^2}{4M_{\text{pl}}^2} - \frac{1}{2}. \quad (85)$$

Fluctuations observed in the CMB are created at

$$\phi_{\text{cmb}} = 2\sqrt{N_{\text{cmb}}} M_{\text{pl}} \sim 15M_{\text{pl}}. \quad (86)$$

In the next lecture we will come back to this example when we compute the fluctuation spectrum generated by  $m^2\phi^2$  inflation.

**Exercise 2 ( $m^2\phi^2$  Inflation)** Verify the above slow-roll results for  $m^2\phi^2$  inflation.

## 6.4 Reheating

After inflation ends the scalar field begins to oscillate around the minimum of the potential. During this phase of coherent oscillations the scalar field acts like pressureless matter

$$\frac{d\bar{\rho}_\phi}{dt} + 3H\bar{\rho}_\phi = 0. \quad (87)$$

**Exercise 3 (Coherent Scalar Field Oscillations)** *Confirm (87) from the equations of motion for  $\phi$ .*

The coupling to other particles leads to a decay of the inflaton energy

$$\frac{d\bar{\rho}_\phi}{dt} + (3H + \Gamma_\phi)\bar{\rho}_\phi = 0. \quad (88)$$

The coupling parameter  $\Gamma_\phi$  depends on complicated and model-dependent physical processes that we do not have the time to review. Eventually, the inflationary energy density is converted into standard model degrees of freedom and the hot big bang commences.

Reheating is a rich and complicated subject to which we couldn't do justice to in these lectures. We refer the interested reader to the review by Bassett et al. [8] for more details.

## 6.5 Models of Inflation

The fundamental microscopic origin of inflation is still a mystery. Basic questions like: what is the inflaton? what is the shape of the inflaton potential? and why did the universe start in a high energy state? remain unanswered. The challenge to explain the physics of inflation is considerable. Inflation is believed to have occurred at an enormous energy scale  $\sim 10^{15}$  GeV, far out of reach of terrestrial particle accelerators. Any description of the inflationary era therefore requires a huge extrapolation of the known laws of physics, and until recently, only a phenomenological parameterization of the inflationary dynamics was possible.<sup>3</sup> In this approach, a suitable inflationary potential function  $V(\phi)$  is postulated (see Figures 7 and 8 for two popular examples) and the experimental predictions are computed from that. As we will see in the next lecture, details of the primordial fluctuation spectra will depend on the precise shape of the inflaton potential.

### 6.5.1 Single-Field Slow-Roll Inflation

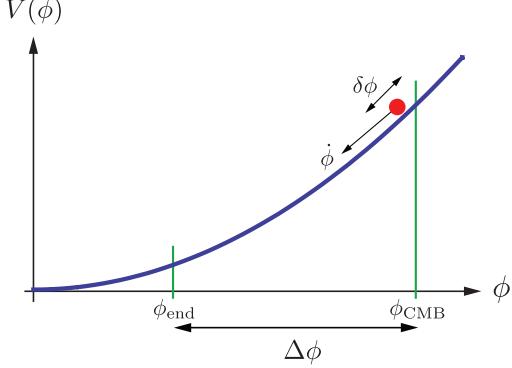
The definition of an inflationary model amounts to a specification of the inflaton action (potential and kinetic terms) and its coupling to gravity. So far we have phrased our discussion of inflation in terms of the simplest models, single-field slow-roll inflation, characterized by the following action

$$S = \int d^4x \sqrt{-g} \left[ \frac{1}{2}R + \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi) \right]. \quad (89)$$

The dynamics of the inflaton field, from the time when CMB fluctuations were created (see Lecture 2) at  $\phi_{\text{cmb}}$  to the end of inflation at  $\phi_{\text{end}}$ , is determined by the shape of the inflationary potential  $V(\phi)$ . The different possibilities for  $V(\phi)$  can be classified in a useful way by determining whether they allow the inflaton field to move over a large or small distance  $\Delta\phi \equiv \phi_{\text{cmb}} - \phi_{\text{end}}$  in Planck units.

---

<sup>3</sup>Recently, progress has been made both in a systematic effective field theory description of inflation [9, 10] and in top-down derivations of inflationary potentials from string theory [11].



**Figure 8:** Large-field inflation. In an important class of inflationary models the inflationary dynamics is driven by a single monomial term in the potential,  $V(\phi) \propto \phi^p$ . In these models the inflaton field evolves over a super-Planckian range during inflation,  $\Delta\phi > M_{\text{pl}}$ , and a large amplitude of gravitational waves is produced by quantum mechanical fluctuations (see **Lecture 2**).

### 1. Small-Field Inflation

In small-field models the field moves over a small (sub-Planckian) distance:  $\Delta\phi < M_{\text{pl}}$ . This is relevant for future observations because small-field models predict that the amplitude of the gravitational waves produced during inflation is too small to be detected (see Lecture 2). The potentials that give rise to such small-field evolution often arise in mechanisms of *spontaneous symmetry breaking*, where the field rolls off an unstable equilibrium toward a displaced vacuum (see Fig. 7). A simple example is the Higgs-like potential

$$V(\phi) = V_0 \left[ 1 - \left( \frac{\phi}{\mu} \right)^2 \right]^2. \quad (90)$$

More generally, small-field models can be locally approximated by the following expansion

$$V(\phi) = V_0 \left[ 1 - \left( \frac{\phi}{\mu} \right)^p \right] + \dots \quad (91)$$

where the dots represent higher-order terms that become important near the end of inflation and during reheating.

Historically, a famous inflationary potential is the *Coleman-Weinberg potential* [2, 3]

$$V(\phi) = V_0 \left[ \left( \frac{\phi}{\mu} \right)^4 \left( \ln \left( \frac{\phi}{\mu} \right) - \frac{1}{4} \right) + \frac{1}{4} \right], \quad (92)$$

which arises as the potential for radiatively-induced symmetry breaking in electroweak and grand unified theories. Although the original values of the parameters  $V_0$  and  $\mu$  based on the  $SU(5)$  theory are incompatible with the small amplitude of inflationary fluctuations, the Coleman-Weinberg potential remains a popular phenomenological model.

### 2. Large-Field Inflation

In large-field models the inflaton field starts at a large field value and then evolves to a minimum at the origin  $\phi = 0$ . If the field evolution is super-Planckian,  $\Delta\phi > M_{\text{pl}}$ , the gravitational waves produced by inflation should be observed in the near future.

The prototypical large-field model is *chaotic inflation* where a single monomial term dominates the potential (see Fig. 8)

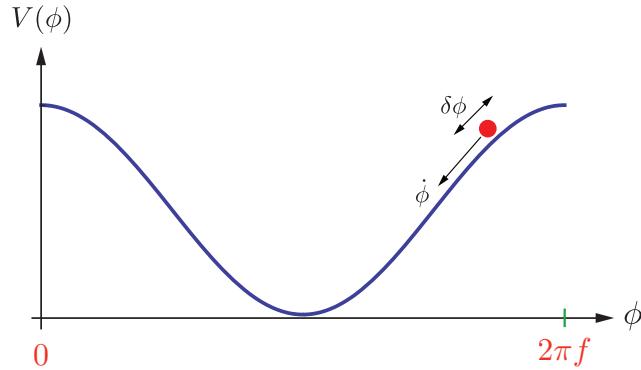
$$V(\phi) = \lambda_p \phi^p. \quad (93)$$

For such a potential the slow-roll parameters are small for super-Planckian field values,  $\phi \gg M_{\text{pl}}$  (notice that the slow-roll conditions are independent of the coupling constant  $\lambda_p$ ). However, to arrange for a small amplitude of density fluctuations (see Lecture 2) the inflaton self-coupling has to be very small,  $\lambda_p \ll 1$ . This condition automatically guarantees that the potential energy (density) is sub-Planckian,  $V \ll M_{\text{pl}}^4$ , and quantum gravity effects are not necessarily important (see §27).

One of the most elegant inflationary models is *natural inflation* where the potential takes the following form (see Fig. 9)

$$V(\phi) = V_0 \left[ \cos\left(\frac{\phi}{f}\right) + 1 \right]. \quad (94)$$

This potential often arises if the inflaton field is taken to be an axion. Depending on the parameter  $f$  the model can be of the small-field or large-field type. However, it is particularly attractive to consider natural inflation for large-field variations,  $2\pi f > M_{\text{pl}}$ , since for axions a shift symmetry can be employed to protect the potential from correction terms even over large field ranges (see §27).



**Figure 9:** Natural Inflation. If the periodicity  $2\pi f$  is super-Planckian the model can naturally support a large gravitational wave amplitude.

### 6.5.2 Beyond Single-Field Slow-Roll

The possibilities for getting inflationary expansion are (maybe frustratingly) varied. Inflation is a paradigm, a framework for a theory of the early universe, but it is not a unique theory. A large number of phenomenological models have been proposed with different theoretical motivations and observational predictions. For the majority of these lectures we will focus on the simplest single-field slow-roll models that we just described. However, in this short section we want to relieve ourselves of the sin of not mentioning the broader landscape of inflationary model-building.

The simplest inflationary actions (89) may be extended in a number of obvious ways:

1. *Non-minimal coupling to gravity.*

The action (89) is called minimally coupled in the sense that there is no direct coupling between the inflaton field and the metric. In principle, we could imagine a non-minimal coupling between the inflaton and the graviton, however, in practice, non-minimally coupled theories can be transformed to minimally coupled form by a field redefinition.

2. *Modified gravity.*

Similarly, we could entertain the possibility that the Einstein-Hilbert part of the action is modified at high energies. However, the simplest examples for this UV modification of gravity, so-called  $f(R)$  theories, can again be transformed into a minimally coupled scalar field with potential  $V(\phi)$ .

3. *Non-canonical kinetic term.*

The action (89) has a canonical kinetic term

$$\mathcal{L}_\phi = X - V(\phi), \quad X \equiv \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi. \quad (95)$$

Inflation can then only occur if the potential  $V(\phi)$  is very flat. More generally, however, we could imagine that the high-energy theory has fields with non-canonical kinetic terms

$$\mathcal{L}_\phi = F(\phi, X) - V(\phi), \quad (96)$$

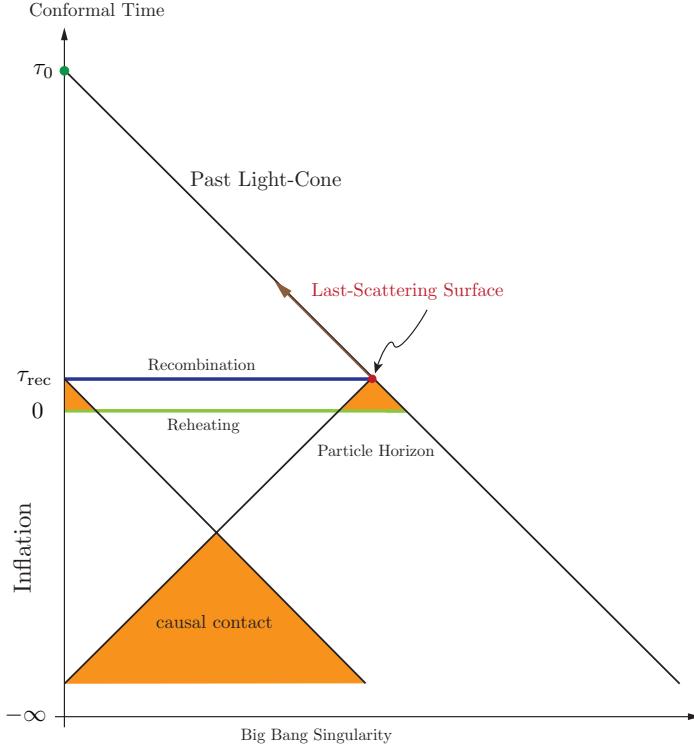
where  $F(\phi, X)$  is some function of the inflaton field and its derivatives. For actions such as (96) it is possible that inflation is driven by the kinetic term and occurs even in the presence of a steep potential.

4. *More than one field.*

If we allow more than one field to be dynamically relevant during inflation, then the possibilities for the inflationary dynamics (and the mechanisms for the production of fluctuations) expand dramatically and the theory loses a lot of its predictive power.

## 7 Summary: Lecture 1

The initial conditions for the conventional FRW cosmology seem highly tuned. Both the horizon problem and the flatness problem can be traced back to the fact that during the standard Big Bang evolution the comoving Hubble radius,  $(aH)^{-1}$ , grows monotonically with time. During inflation on the other hand the comoving Hubble radius is temporarily decreasing. This changes the causal



**Figure 10:** Conformal diagram of inflationary cosmology. Inflation extends conformal time to negative values! The end of inflation creates an “apparent” Big Bang at  $\tau = 0$ . There is, however, no singularity at  $\tau = 0$  and the light cones intersect at an earlier time if inflation lasts for at least 60  $e$ -folds.

structure of the early universe making the horizon problem a fiction of extrapolating the conventional FRW expansion back to arbitrarily early times. From the Einstein equations one may show that a shrinking Hubble radius corresponds to accelerated expansion as it occurs if the universe is filled with a negative pressure component. The three equivalent conditions for inflation therefore are

$$\frac{d(aH)^{-1}}{dt} < 0 \quad \Rightarrow \quad \frac{d^2a}{dt^2} > 0 \quad \Rightarrow \quad p < -\frac{\rho}{3} .$$

A negative pressure fluid can be modeled by scalar field  $\phi$ , the inflaton, with the following action

$$S = \int d^4x \sqrt{-g} \left[ \frac{1}{2}R + \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi) \right]$$

This will lead to inflation if the slow-roll conditions are satisfied

$$\epsilon_v = \frac{M_{pl}^2}{2} \left( \frac{V_{,\phi}}{V} \right)^2 , \quad \eta_v = M_{pl}^2 \frac{V_{,\phi\phi}}{V} , \quad \epsilon_v, |\eta_v| < 1 .$$

The number of  $e$ -folds of inflationary expansion then is

$$N(\phi) = \int_{\phi_{\text{end}}}^{\phi} \frac{d\phi}{\sqrt{2\epsilon_v}}. \quad (97)$$

The total number of  $e$ -folds needs to be at least 60 to solve the horizon problem. Even in the restricted framework for single-field slow-roll inflation described by the above action, there are a multitude of inflationary models characterized by different choices for the inflationary potential  $V(\phi)$ .

## 8 Problem Set: Lecture 1

**Problem 1 (Homogeneous and Isotropic Spaces)** *Homogeneous and isotropic spaces are characterized by translational and rotational invariance. Convince yourself that in three dimensions there exist only three types of homogeneous and isotropic spaces with simple topology:*

- i) flat space
- ii) a three-dimensional sphere with constant positive curvature
- iii) a three-dimensional hyperbolic space with constant negative curvature.

*It is easier to visualize the two-dimensional analogues. Consider the embedding of a two-dimensional sphere in a three-dimensional Euclidean space*

$$x^2 + y^2 + z^2 = a^2, \quad (98)$$

*where  $a$  is the radius of the sphere. Show that the induced metric on the surface of the sphere is*

$$d\ell_2^2 = \frac{dr'^2}{1 - (r'^2/a^2)} + r'^2 d\phi^2, \quad (99)$$

*where  $x = r' \cos \phi$  and  $y = r' \sin \phi$ . The limit  $a^2 \rightarrow \infty$  corresponds to flat space (a plane). Negative  $a^2$  corresponds to a space with constant negative curvature. It cannot be embedded in three-dimensional Euclidean space. (Consider the embedding of  $x^2 + y^2 - z^2 = -a^2$  in a space with metric  $d\ell_2^2 = dx^2 + dy^2 - dz^2$  instead.)*

*By rescaling the radial coordinate the metric can be brought into the form*

$$d\ell_2^2 = |a^2| \left( \frac{dr^2}{1 - kr^2} + r^2 d\phi^2 \right) \quad (100)$$

*where  $k = +1$  for the sphere ( $a^2 > 0$ ),  $k = -1$  for the hyperbolic space ( $a^2 < 0$ ) and  $k = 0$  for the plane ( $a^2 = 0$ ).*

*Generalize the above argument to the embedding of homogeneous and isotropic three-dimensional spaces in four-dimensional Euclidean space. Show that their metric is*

$$d\ell_3^2 = a^2 \left( \frac{dr^2}{1 - kr^2} + r^2 d\Omega^2 \right), \quad d\Omega^2 \equiv d\theta^2 + \sin^2 \theta d\phi^2. \quad (101)$$

*or*

$$d\ell_3^2 = a^2 \left[ d\chi^2 + \begin{pmatrix} \sinh^2 \chi \\ \chi^2 \\ \sin^2 \chi \end{pmatrix} d\Omega^2 \right] \quad \begin{array}{ll} k = -1 \\ k = 0 \\ k = +1 \end{array} \quad (102)$$

*Convince yourself that the only time-dependent four-dimensional spacetime that preserves homogeneity and isotropy of space is the FRW metric*

$$ds^2 = dt^2 - a^2(t) \left( \frac{dr^2}{1 - kr^2} + r^2 d\Omega^2 \right). \quad (103)$$

**Problem 2 (Conformal Time)** Derive some simple expressions for the conformal time  $\tau$  as a function of  $a$ .

1. Show that  $\tau \propto a^{1/2}$  in a matter-dominated universe and  $\tau \propto a$  in one dominated by radiation.
2. Consider a universe with only matter and radiation, with equality at  $a_{\text{eq}}$ . Show that

$$\tau = \frac{2}{\sqrt{\Omega_m H_0^2}} \left[ \sqrt{a + a_{\text{eq}}} - \sqrt{a_{\text{eq}}} \right]. \quad (104)$$

What is the conformal time today? At decoupling?

3. Use your favorite software (say Mathematica or Maple) to compute the conformal time numerically for our universe (filled with dark energy, matter and radiation). Compute the conformal time today and at decoupling. What is the percentage error between this result and the analytical result for a matter/radiation only universe (Eqn. (104))?

**Problem 3 (Friedmann Equations)** Derive the Ricci tensor and the Ricci scalar for the FRW spacetime (1)

$$R_{00} = -3 \frac{\ddot{a}}{a}, \quad R_{ij} = \delta_{ij} \left[ 2\dot{a}^2 + a\ddot{a} + 2\frac{k}{a^2} \right], \quad R = g^{\mu\nu} R_{\mu\nu} = 6 \left[ \frac{\ddot{a}}{a} + \left( \frac{\dot{a}}{a} \right)^2 + \frac{k}{a^2} \right]. \quad (105)$$

Confirm that the 00-component of the Einstein equation (13) gives the Friedmann equation

$$\left( \frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}. \quad (106)$$

Confirm that the trace of the Einstein equation (13) gives the acceleration equation

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p). \quad (107)$$

Show that the two Friedmann equations imply the continuity equation

$$\dot{\rho} = -3H(\rho + p). \quad (108)$$

Derive the continuity equation from

$$\nabla_\mu T^\mu_\nu = 0. \quad (109)$$

(Hint: Contract (109) with  $U^\mu$ , use the energy-momentum tensor for a perfect fluid and the properties of the 4-velocity. No need for Christoffel symbols!)

**Problem 4 ( $\lambda\phi^4$  Inflation)** Derive the slow-roll dynamics for  $\lambda\phi^4$  inflation.

**Problem 5 (The Phase Space of  $m\phi^2$  Inflation)** Read about the attractor behavior of  $m^2\phi^2$  inflation in Mukhanov's book [6].

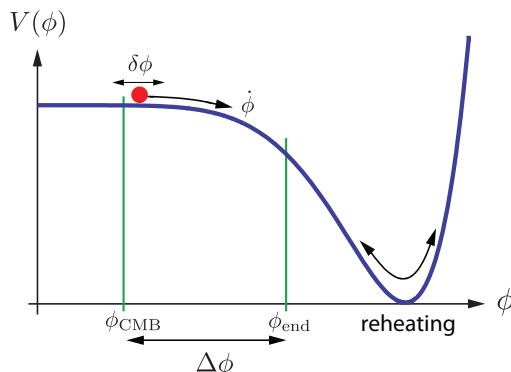
## Part III

# Lecture 2: Quantum Fluctuations during Inflation

### Abstract

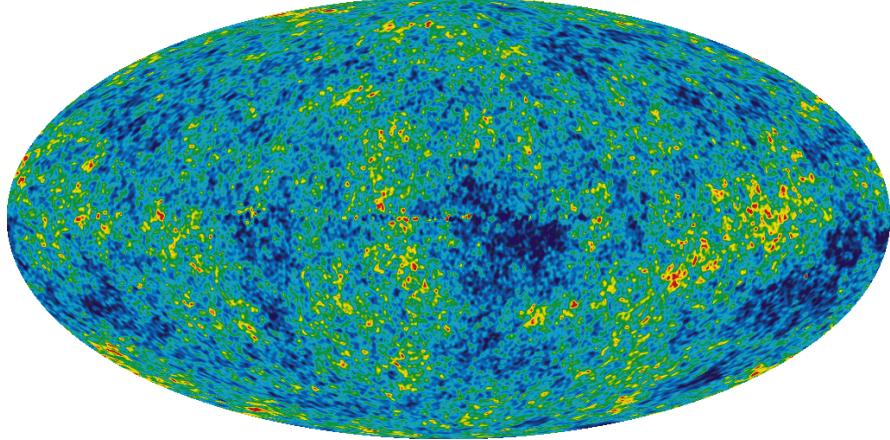
In this lecture we present the famous calculation of the primordial fluctuation spectra generated by quantum fluctuations during inflation. We present the calculation in full detail and try to avoid ‘cheating’ and approximations. After a brief review of fundamental aspects of cosmological perturbation theory, we first give a qualitative summary of the basic mechanism by which inflation converts microscopic quantum fluctuations into macroscopic seeds for cosmological structure formation. As a pedagogical introduction to quantum field theory in curved spacetime we then review the quantization of the simple harmonic oscillator. We emphasize that a unique vacuum state is chosen by demanding that the vacuum is the minimum energy state. We then proceed by giving the corresponding calculation for inflation in all detail. We calculate the power spectra of both scalar and tensor fluctuations and discuss their dependence on scale.

In the last lecture we studied the classical ( $\hbar = 0$ ) dynamics of a scalar field rolling down a potential with speed  $\dot{\phi}$  (see Fig. 11). In this lecture we study the effects of quantum ( $\hbar \neq 0$ ) fluctuations around the classical background evolution  $\bar{\phi}(t)$ . These fluctuations lead to a local time delay in the time inflation ends, *i.e.* different parts of the universe will end inflation at slightly different times. For instance, for the potential shown in Fig. 11 regions acquiring a negative frozen fluctuation  $\delta\phi$  remain potential-dominated longer than regions with positive  $\delta\phi$ . Different parts of the universe therefore undergo slightly different evolutions. This induces relative density fluctuations  $\delta\rho(t, \mathbf{x})$ .



**Figure 11:** Quantum fluctuations  $\delta\phi(t, \mathbf{x})$  around the classical background evolution  $\bar{\phi}(t)$ .

In this lecture we will discuss the technical details underlying this basic picture for the quantum origin of large-scale structure.



**Figure 12:** Observations of the CMB prove that the early universe wasn't perfectly homogeneous, but also show that the inhomogeneities were small and can therefore be analyzed as linear perturbations around a homogenous background.

## 9 Review: Cosmological Perturbations

In this lecture we present in detail the generation of cosmological perturbations from quantum fluctuations during inflation. This discussion will require some background in cosmological perturbation theory which we now briefly review. More details may be found in Appendix A.

### 9.1 Generalities

#### 9.1.1 Linear Perturbations

Observations of the CMB (Fig. 12) explain the success of cosmological perturbation theory. At the time of decoupling the universe was very nearly homogeneously with small inhomogeneities at the  $10^{-5}$  level. A natural strategy therefore is to split all quantities  $X(t, \mathbf{x})$  (metric  $g_{\mu\nu}$  and matter fields  $T_{\mu\nu} \rightarrow \phi, \rho, p$ , etc.) into a homogeneous background  $\bar{X}(t)$  that depends only on cosmic time and a space-dependent perturbation

$$\delta X(t, \mathbf{x}) \equiv X(t, \mathbf{x}) - \bar{X}(t). \quad (110)$$

Because the perturbations are small,  $\delta X \ll \bar{X}$ , expanding the Einstein equation at linear order in perturbations approximates the full non-linear solution to very high accuracy

$$\delta G_{\mu\nu} = 8\pi G \delta T_{\mu\nu}. \quad (111)$$

#### 9.1.2 Gauge Choice

A crucial subtlety in the study of cosmological perturbations is the fact that the split into background and perturbations, Eqn. (110), is not unique, but depends on the choice of coordinates or the *gauge*

*choice*.<sup>4</sup> When we described the homogeneous universe in **Lecture 1** we introduced coordinates  $t$  and  $\mathbf{x}$  to define the FRW metric. The spacelike hypersurfaces of constant time  $t$  defined the *slicing* to the four-dimensional spacetime, while the timelike worldlines of constant  $\mathbf{x}$  defined the *threading*. The FRW threading corresponds to the motion of comoving observers who see zero momentum density at their location. These observers are free-falling and the expansion defined by them is isotropic. The slicing is orthogonal to the threading with each spacelike slice corresponding to a homogeneous universe. These features made our coordinate choice so distinguished that we never worried about other coordinates (in which the universe would not look homogeneous and isotropic). However, now that we are considering perturbations it is important to realize that the slicing and threading of the perturbed spacetime is not unique. Furthermore, when describing an inhomogeneous spacetime there is often not a preferred coordinate choice. When we make a gauge choice to define the slicing and threading of the spacetime we implicitly also define the perturbations. If we aren't careful this gauge dependence of perturbations can lead to some confusion. To demonstrate this fact most dramatically consider an *unperturbed* homogeneous and isotropic universe, where  $\rho(t, \mathbf{x}) = \rho(t)$ . We now show that a change of the time coordinate can introduce *fictitious* perturbations  $\delta\rho$ . Consider a new time coordinate  $\tilde{t} = t + \delta t(t, \mathbf{x})$ . In general, the energy density on the new time-slice will not be homogeneous,  $\tilde{\rho}(\tilde{t}, \mathbf{x}) = \rho(t(\tilde{t}, \mathbf{x}))$ . These perturbations in the energy density aren't physical, but entirely due to the choice of new time-slicing. Similarly, we can remove a *real* perturbation in the energy density by choosing the hypersurface of constant time to coincide with the hypersurface of constant energy density. Then  $\delta\tilde{\rho} = 0$  although there are real inhomogeneities. To resolve ambiguities between real and fake perturbations in general relativity, we need to consider the complete set of perturbations, *i.e.* we need both the matter field perturbations and the metric perturbations and by a gauge transformation we can trade one for the other. To avoid misinterpretation of fictitious gauge modes it will also be useful to study gauge-invariant combinations of perturbations. Fluctuations of gauge-invariant quantities can by definition not be removed by a coordinate transformation.

### 9.1.3 Scalars, Vectors and Tensors

The spatially flat, homogeneous and isotropic background spacetime possesses a great deal of symmetry. These symmetries allow a decomposition of the metric and the stress-energy perturbations into independent scalar (S), vector (V) and tensor (T) components. This *SVT decomposition* is most easily described in Fourier space

$$X_{\mathbf{k}}(t) = \int d^3\mathbf{x} X(t, \mathbf{x}) e^{i\mathbf{k}\cdot\mathbf{x}}, \quad X \equiv \delta\phi, \delta g_{\mu\nu}, \text{etc.} \quad (112)$$

We note that *translation invariance* of the linear equations of motion for the perturbations means that the different Fourier modes do not interact (see Appendix A for the proof). Different Fourier

---

<sup>4</sup>The perturbation  $\Delta X$  in any relevant quantity, say represented by a tensor field  $X$ , is defined as the difference between the value  $X$  has in the physical spacetime (the perturbed spacetime), and the value  $X_0$  the same quantity has in the given (unperturbed) background spacetime. However, it is a basic fact of differential geometry that, in order to make the comparison of tensors meaningful, they can thus be compared only after a prescription for identifying points of these spacetimes is given. A *gauge choice* is precisely this, *i.e.* a one-to-one correspondence (map) between the background and the physical spacetime. A change of this map is then a gauge transformation, and the freedom one has in choosing it gives rise to an arbitrariness in the value of the perturbation of  $X$  at any given spacetime point, unless  $X$  is gauge-invariant.

modes can therefore be studied independently. This often simplifies the differential equations for the perturbations. Next we consider rotations around a single Fourier wavevector  $\mathbf{k}$ . A perturbation is said to have helicity  $m$  if its amplitude is multiplied by  $e^{im\psi}$  under rotation of the coordinate system around the wavevector by an angle  $\psi$

$$X_{\mathbf{k}} \rightarrow e^{im\psi} X_{\mathbf{k}}. \quad (113)$$

Scalar, vector and tensor perturbations have helicity 0,  $\pm 1$  and  $\pm 2$ , respectively.<sup>5</sup> The importance of the SVT decomposition is that the perturbations of each type evolve independently (at the linear level) and can therefore be treated separately (see Appendix A for the proof). This considerably simplifies the study of cosmological perturbations.

After these general remarks, let us now become more specific and explicitly define the perturbations around the homogenous and isotropic FRW universe.

## 9.2 The Inhomogeneous Universe

### 9.2.1 Metric Perturbations

During inflation we define perturbations around the homogeneous background solutions for the inflaton  $\bar{\phi}(t)$  and the metric  $\bar{g}_{\mu\nu}(t)$ ,

$$\phi(t, \mathbf{x}) = \bar{\phi}(t) + \delta\phi(t, \mathbf{x}), \quad g_{\mu\nu}(t, \mathbf{x}) = \bar{g}_{\mu\nu}(t) + \delta g_{\mu\nu}(t, \mathbf{x}) \quad (114)$$

where

$$\begin{aligned} ds^2 &= g_{\mu\nu} dx^\mu dx^\nu \\ &= -(1 + 2\Phi)dt^2 + 2aB_i dx^i dt + a^2[(1 - 2\Psi)\delta_{ij} + E_{ij}]dx^i dx^j. \end{aligned} \quad (115)$$

In real space, the SVT decomposition of the metric perturbations (115) is<sup>6</sup>

$$B_i \equiv \partial_i B - S_i, \quad \text{where } \partial^i S_i = 0, \quad (116)$$

and

$$E_{ij} \equiv 2\partial_{ij}E + 2\partial_{(i}F_{j)} + h_{ij}, \quad \text{where } \partial^i F_i = 0, \quad h_i^i = \partial^i h_{ij} = 0. \quad (117)$$

The vector perturbations  $S_i$  and  $F_i$  aren't created by inflation (and in any case decay with the expansion of the universe). For this reason we ignore vector perturbations in these lectures. Our focus will be on scalar and tensor fluctuations which are observed as density fluctuations and gravitational waves in the late universe.

Tensor fluctuations are gauge-invariant, but scalar fluctuations change under a change of coordinates. Consider the gauge transformation

$$t \rightarrow t + \alpha \quad (118)$$

$$x^i \rightarrow x^i + \delta^{ij}\beta_j. \quad (119)$$

---

<sup>5</sup>Should this abstract definition of scalar, vector and tensor perturbations in terms of their helicities be confusing, the reader may want to test those rules on the explicit metric and stress-energy perturbations introduced in the next section.

<sup>6</sup>SVT decomposition in real space corresponds to the distinctive transformation properties of scalars, vectors and tensors on spatial hypersurfaces.

Under these coordinate transformations the scalar metric perturbations transform as

$$\Phi \rightarrow \Phi - \dot{\alpha} \quad (120)$$

$$B \rightarrow B + a^{-1}\alpha - a\dot{\beta} \quad (121)$$

$$E \rightarrow E - \beta \quad (122)$$

$$\Psi \rightarrow \Psi + H\alpha. \quad (123)$$

**Exercise 4 (Linear gauge transformations)** *Derive the gauge transformations of the scalar metric perturbations (120)–(123). Hint: use invariance of the spacetime interval,*

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = \tilde{g}_{\mu\nu}d\tilde{x}^\mu d\tilde{x}^\nu. \quad (124)$$

### 9.2.2 Matter Perturbations

During inflation the inflationary energy is the dominant contribution to the stress-energy of the universe, so that the inflaton perturbations  $\delta\phi$  backreact on the spacetime geometry. This coupling between matter perturbations and metric perturbations is described by the Einstein equations (see Appendix A).

After inflation, the perturbations to the total stress-energy tensor of the universe are

$$T_0^0 = -(\bar{\rho} + \delta\rho) \quad (125)$$

$$T_i^0 = (\bar{\rho} + \bar{p})v_i \quad (126)$$

$$T_0^i = -(\bar{\rho} + \bar{p})(v^i + B^i) \quad (127)$$

$$T_j^i = \delta_j^i(\bar{p} + \delta p) + \Sigma_j^i. \quad (128)$$

The anisotropic stress  $\Sigma_j^i$  is gauge-invariant while the density, pressure and momentum density ( $(\delta q)_i \equiv (\bar{\rho} + \bar{p})v_i$ ) transform as follows

$$\delta\rho \rightarrow \delta\rho - \dot{\bar{\rho}}\alpha \quad (129)$$

$$\delta p \rightarrow \delta p - \dot{\bar{p}}\alpha \quad (130)$$

$$\delta q \rightarrow \delta q + (\bar{\rho} + \bar{p})\alpha. \quad (131)$$

### 9.2.3 Gauge-Invariant Variables

As we explained above, to avoid the pitfall of fictitious gauge modes, it is useful to introduce gauge-invariant combinations of metric and matter perturbations. An important gauge-invariant scalar quantity is the *curvature perturbation on uniform-density hypersurfaces* [12]

$$\boxed{-\zeta \equiv \Psi + \frac{H}{\dot{\bar{\rho}}}\delta\rho}. \quad (132)$$

Geometrically,  $\zeta$  measures the spatial curvature of constant-density hypersurfaces,  $R^{(3)} = 4\nabla^2\Psi/a^2$ . An important property of  $\zeta$  is that it remains constant outside the horizon for adiabatic matter perturbations, *i.e.* perturbation that satisfy

$$\boxed{\delta p_{en} \equiv \delta p - \frac{\dot{\bar{p}}}{\dot{\bar{\rho}}}\delta\rho} = 0. \quad (133)$$

Notice that the definition of  $\delta p_{en}$  is gauge-invariant. In the single-field inflation models studied in this lecture the condition (291) is always satisfied, so the perturbation  $\zeta_{\mathbf{k}}$  doesn't evolve outside the horizon,  $k \ll aH$ .

In a gauge defined by spatially flat hypersurfaces,  $\Psi$ , the perturbations  $\zeta$  is the dimensionless density perturbation  $\frac{1}{3}\delta\rho/(\bar{\rho} + \bar{p})$ . Taking into account appropriate transfer functions to describe the sub-horizon evolution of the fluctuations, CMB and large-scale structure (LSS) observations can therefore be related to the primordial value of  $\zeta$  (see **Lecture 3**). During slow-roll inflation

$$-\zeta \approx \Psi + \frac{H}{\dot{\phi}}\delta\phi. \quad (134)$$

Another gauge-invariant scalar is the *comoving curvature perturbation*

$$\mathcal{R} \equiv \Psi - \frac{H}{\bar{\rho} + \bar{p}}\delta q, \quad (135)$$

where  $\delta q$  is the scalar part of the 3-momentum density  $T_i^0 = \partial_i\delta q$ . During inflation  $T_i^0 = -\dot{\phi}\partial_i\delta\phi$  and hence

$$\mathcal{R} = \Psi + \frac{H}{\dot{\phi}}\delta\phi. \quad (136)$$

Geometrically,  $\mathcal{R}$  measures the spatial curvature of comoving (or constant- $\phi$ ) hypersurfaces.

The linearized Einstein equations relate  $\zeta$  and  $\mathcal{R}$  as follows (see Appendix A)

$$-\zeta = \mathcal{R} + \frac{k^2}{(aH)^2} \frac{2\bar{\rho}}{3(\bar{\rho} + \bar{p})} \Psi_B \quad (137)$$

where

$$\Psi_B \equiv \psi + a^2 H(\dot{E} - B/a). \quad (138)$$

$\zeta$  and  $\mathcal{R}$  are therefore equal on superhorizon scales,  $k \ll aH$ .  $\zeta$  and  $\mathcal{R}$  are also equal during slow-roll inflation, *cf.* Eqs. (134) and (136). The correlation functions of  $\zeta$  and  $\mathcal{R}$  are therefore equal at horizon crossing and both  $\zeta$  and  $\mathcal{R}$  are conserved on superhorizon scales. In this lecture we will compute the primordial spectrum of  $\mathcal{R}$  at horizon crossing.

Finally, a gauge-invariant measure of inflaton perturbations is the *inflaton perturbation on spatially flat slices*

$$Q \equiv \delta\phi + \frac{\dot{\phi}}{H}\Psi. \quad (139)$$

**Exercise 5 (Gauge-invariant perturbations)** *Using the linear gauge transformations for the metric and matter perturbations, confirm that  $\zeta$ ,  $\mathcal{R}$  and  $Q$  are gauge-invariant.*

#### 9.2.4 Superhorizon (Non-)Evolution

The Einstein equations (see Appendix A) give the evolution equation for the gauge-invariant curvature perturbation

$$\dot{\mathcal{R}} = -\frac{H}{\bar{\rho} + \bar{p}}\delta p_{en} + \mathcal{O}\left(\frac{k^2}{(aH)^2}\right). \quad (140)$$

For adiabatic matter perturbations ( $\delta p_{en} = 0$ ),  $\mathcal{R}$  is conserved on superhorizon scales,  $k < aH$ .

**Exercise 6 (Separate universe approach)** *Read about the separate universe approach [13] for proving conservation of the curvature perturbation  $\mathcal{R}$  on superhorizon scales.*

### 9.3 Statistics of Cosmological Perturbations

A crucial statistical measure of the primordial scalar fluctuations is the power spectrum of  $\mathcal{R}$  (or  $\zeta$ )<sup>7</sup>

$$\langle \mathcal{R}_\mathbf{k} \mathcal{R}_{\mathbf{k}'} \rangle = (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}') P_{\mathcal{R}}(k), \quad \Delta_s^2 \equiv \Delta_{\mathcal{R}}^2 = \frac{k^3}{2\pi^2} P_{\mathcal{R}}(k). \quad (141)$$

Here,  $\langle \dots \rangle$  defines the ensemble average of the fluctuations. The scale-dependence of the power spectrum is defined by the scalar spectral index (or tilt)

$$n_s - 1 \equiv \frac{d \ln \Delta_s^2}{d \ln k}, \quad (142)$$

where scale-invariance corresponds to the value  $n_s = 1$ . We may also define the running of the spectral index by

$$\alpha_s \equiv \frac{dn_s}{d \ln k}. \quad (143)$$

The power spectrum is often approximated by a power law form

$$\Delta_s^2(k) = A_s(k_*) \left( \frac{k}{k_*} \right)^{n_s(k_*) - 1 + \frac{1}{2}\alpha_s(k_*) \ln(k/k_*)}, \quad (144)$$

where  $k_*$  is an arbitrary reference or pivot scale.

If  $\mathcal{R}$  is Gaussian then the power spectrum contains all the statistical information. Primordial non-Gaussianity is encoded in higher-order correlation functions of  $\mathcal{R}$ . In single-field slow-roll inflation the non-Gaussianity is predicted to be small, [14, 15], but non-Gaussianity can be significant in multi-field models or in single-field models with non-trivial kinetic terms and/or violation of the slow-roll conditions. We will return to primordial non-Gaussianity in **Lecture 3**. In this lecture we restrict our computation to Gaussian fluctuations and the associated power spectra.

The power spectrum for the two polarization modes of  $h_{ij}$ , *i.e.*  $h \equiv h^+, h^\times$ , is defined as

$$\langle h_\mathbf{k} h_{\mathbf{k}'} \rangle = (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}') P_h(k), \quad \Delta_h^2 = \frac{k^3}{2\pi^2} P_h(k). \quad (145)$$

We define the power spectrum of tensor perturbations as the sum of the power spectra for the two polarizations

$$\Delta_t^2 \equiv 2\Delta_h^2. \quad (146)$$

Its scale-dependence is defined analogously to (142) but for historical reasons without the  $-1$ ,

$$n_t \equiv \frac{d \ln \Delta_t^2}{d \ln k}, \quad (147)$$

*i.e.*

$$\Delta_t^2(k) = A_t(k_*) \left( \frac{k}{k_*} \right)^{n_t(k_*)}. \quad (148)$$

---

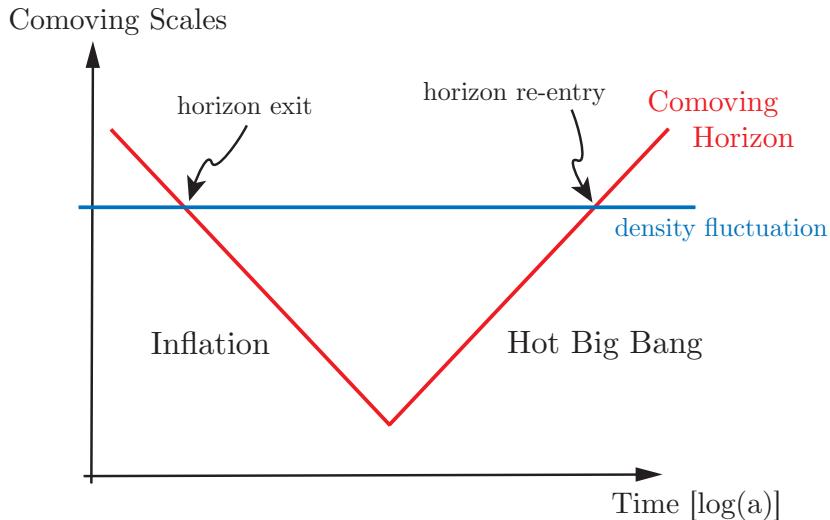
<sup>7</sup>The normalization of the dimensionless power spectrum  $\Delta_{\mathcal{R}}^2(k)$  is chosen such that the variance of  $\mathcal{R}$  is  $\langle \mathcal{R} \mathcal{R} \rangle = \int_0^\infty \Delta_{\mathcal{R}}^2(k) d \ln k$ .

## Aim of this Lecture

It will be the aim of this lecture to compute the power spectra of scalar and tensor fluctuations,  $P_R(k)$  and  $P_h(k)$ , from first principles. This is one of the most important calculations in modern theoretical cosmology, so to understand it will be well worth our efforts.

## 10 Preview: The Quantum Origin of Structure

In the last lecture we discussed the classical evolution of the inflaton field. Something remarkable happens when one considers quantum fluctuations of the inflaton: inflation combined with quantum mechanics provides an elegant mechanism for generating the initial seeds of all structure in the universe. In other words, quantum fluctuations during inflation are the source of the primordial power spectra of scalar and tensor fluctuations,  $P_s(k)$  and  $P_t(k)$ . In this section we sketch the mechanism by which inflation relates microscopic physics to macroscopic observables. In §12 we present the full calculation.



**Figure 13:** Creation and evolution of perturbations in the inflationary universe. Fluctuations are created quantum mechanically on subhorizon scales. While comoving scales,  $k^{-1}$ , remain constant the comoving Hubble radius during inflation,  $(aH)^{-1}$ , shrinks and the perturbations exit the horizon. Causal physics cannot act on superhorizon perturbations and they freeze until horizon re-entry at late times.

### 10.1 Quantum Zero-Point Fluctuations

As we explain quantitatively in §12 quantum fluctuations during inflation induce a non-zero variance for fluctuations in all light fields (like the inflaton or the metric). This is very similar to the variance in the amplitude of a harmonic oscillator induced by zero-point fluctuations; see §11. Fluctuations are created on all length scales, *i.e.* with a spectrum of wavenumbers  $k$ . Cosmologically relevant fluctuations start their lives inside the horizon (Hubble radius),  $k \gg aH$ . However, while the

comoving wavenumber is constant the comoving Hubble radius shrinks during inflation (recall this is how we ‘defined’ inflation!), so eventually all fluctuations *exit the horizon*.

## 10.2 Horizon Exit and Re-Entry

Cosmological inhomogeneity is characterized by the intrinsic curvature of spatial hypersurfaces defined with respect to the matter,  $\mathcal{R}$  or  $\zeta$ . Both  $\mathcal{R}$  and  $\zeta$  have the attractive feature that they remain constant outside the horizon, *i.e.* when  $k < aH$ . In particular, their amplitude is *not* affected by the unknown physical properties of the universe shortly after inflation (recall that we know next to nothing about the details of reheating; it is the constancy of  $\mathcal{R}$  and  $\zeta$  outside the horizon that allows us to nevertheless predict cosmological observables). After inflation, the comoving horizon grows, so eventually all fluctuations will re-enter the horizon. After horizon re-entry,  $\mathcal{R}$  or  $\zeta$  determine the perturbations of the cosmic fluid resulting in the observed CMB anisotropies and the LSS.

In [Lecture 1](#) we explained the evolution of the comoving horizon during inflation and in the standard FRW evolution after inflation. In this lecture ([Lecture 2](#)) we will compute the primordial power spectrum of comoving curvature fluctuations  $\mathcal{R}$  at horizon exit. In the next lecture ([Lecture 3](#)) we will compute the relation of  $\mathcal{R}$  fluctuations to fluctuations in cosmological observables after horizon re-entry. Together these three lectures therefore provide a complete account of both the generation and the observational consequences of the quantum fluctuations produced by inflation. It is a beautiful story. Let us begin to unfold it.

## 11 Quantum Mechanics of the Harmonic Oscillator

“The career of a young theoretical physicist consists of treating the harmonic oscillator in ever-increasing levels of abstraction.”

*Sidney Coleman*

The computation of quantum fluctuations generated during inflation is algebraically quite intensive and it is therefore instructive to start with a simpler example which nevertheless contains most of the relevant physics. We therefore warm up by considering the quantization of a one-dimensional simple harmonic oscillator. Harmonic oscillators are one of the few physical systems that physicists know how to solve exactly. Fortunately, almost all more complicated physical systems can be represented by a collection of simple harmonic oscillators with different amplitudes and frequencies. This is of course what Fourier analysis is all about. We will show below that free fields in curved spacetime (and de Sitter space in particular) are similar to collections of harmonic oscillators with time-dependent frequencies. The detailed treatment of the quantum harmonic oscillator in this section will therefore not be in vain, but will provide important intuition for the inflationary calculation.

## 11.1 Action

The classical action of a harmonic oscillator with time-dependent frequency is

$$S = \int dt \left( \frac{1}{2} \dot{x}^2 - \frac{1}{2} \omega^2(t) x^2 \right) \equiv \int dt L, \quad (149)$$

where for convenience we have set the particle mass  $m \equiv 1$ . For concreteness one may wish to consider a particle of mass  $m$  on a spring which is heated by an external source so that its spring constant depends on time,  $k(t)$ , where  $\omega^2 = k/m$ . The classical equation of motion follows from variation of the action with respect to the particle coordinate  $x$

$$\frac{\delta S}{\delta x} = 0 \Rightarrow \boxed{\ddot{x} + \omega^2(t) x = 0}. \quad (150)$$

## 11.2 Canonical Quantization

Canonical quantization of the system proceeds in the standard way: We define the momentum conjugate to  $x$

$$p \equiv \frac{dL}{dx} = \dot{x}, \quad (151)$$

which agrees with the standard notion of particle momentum  $p = mv$ . We then promote the classical variables  $x, p$  to quantum operators  $\hat{x}, \hat{p}$  and impose the canonical commutator

$$\boxed{[\hat{x}, \hat{p}] = i\hbar}, \quad (152)$$

where  $[\hat{x}, \hat{p}] \equiv \hat{x}\hat{p} - \hat{p}\hat{x}$ . The equation of motion implies that the commutator holds at all times if imposed at some initial time. In particular, for our present example

$$[x(t), \dot{x}(t)] = i\hbar. \quad (153)$$

Note that we are in the Heisenberg picture where operators vary in time while states are time-independent. The operator  $\hat{x}$  is then expanded in terms of creation and annihilation operators

$$\boxed{\hat{x} = v(t) \hat{a} + v^*(t) \hat{a}^\dagger}, \quad (154)$$

where the (complex) mode function satisfies the classical equation of motion

$$\ddot{v} + \omega^2(t)v = 0. \quad (155)$$

The commutator (152) becomes

$$\langle v, v \rangle [\hat{a}, \hat{a}^\dagger] = 1, \quad (156)$$

where the bracket notation is defined by

$$\langle v, w \rangle \equiv \frac{i}{\hbar} (v^* \partial_t w - (\partial_t v^*) w). \quad (157)$$

Let us assume that the solution  $v$  is chosen so that the real number  $\langle v, v \rangle$  is positive. The function  $v$  can then be rescaled such that  $\langle v, v \rangle \equiv 1$  and hence

$$\boxed{[\hat{a}, \hat{a}^\dagger] = 1}. \quad (158)$$

Eqn. (158) is the standard relation for the harmonic oscillator raising and lowering operators. We have hence identified the following annihilation and creation operators

$$\hat{a} = \langle v, \hat{x} \rangle \quad (159)$$

$$\hat{a}^\dagger = -\langle v^*, \hat{x} \rangle \quad (160)$$

and can define the vacuum state  $|0\rangle$  via the prescription

$$\boxed{\hat{a}|0\rangle = 0}, \quad (161)$$

i.e. the vacuum is annihilated by  $\hat{a}$ . Excited states of the system are created by repeated application of creation operators

$$|n\rangle \equiv \frac{1}{\sqrt{n!}} (\hat{a}^\dagger)^n |0\rangle. \quad (162)$$

These states are eigenstates of the number operator  $\hat{N} = \hat{a}^\dagger \hat{a}$  with eigenvalue  $n$ , i.e.

$$\hat{N}|n\rangle = n|n\rangle. \quad (163)$$

### 11.3 Non-Uniqueness of the Mode Functions

So far the solution  $v(t)$  is arbitrary, except for the normalization  $\langle v, v \rangle = 1$ . A change in  $v(t)$  would be accompanied by a change in  $\hat{a} = \langle v, \hat{x} \rangle$  that keeps the solution  $x(t)$  unchanged. For the simple harmonic oscillator with time-dependent frequency  $\omega(t)$  (and for quantum fields in curved spacetime) there is in fact *no* unique choice for the mode function  $v(t)$ . Hence, there is *no* unique decomposition of  $\hat{x}$  into annihilation and creation operators and *no* unique notion of the vacuum. Different choices for the solution  $v(t)$  give different vacuum solutions. This problem and its standard (but not uncontested) resolution in the case of inflation will be discussed in more detail below.

In the present case we can make progress by considering the special case of a constant frequency harmonic oscillator  $\omega(t) = \omega$ . In that case a special choice of  $v(t)$  is selected if we require that the vacuum state  $|0\rangle$  be the ground state of the Hamiltonian. Let's see how this comes about. For general  $v$  we have

$$\begin{aligned} \hat{H} &= \frac{1}{2}\hat{p}^2 + \frac{1}{2}\omega^2\hat{x}^2 \\ &= \frac{1}{2} \left[ (\dot{v}^2 + \omega^2 v^2)\hat{a}\hat{a} + (\dot{v}^2 + \omega^2 v^2)^* \hat{a}^\dagger \hat{a}^\dagger + (|\dot{v}|^2 + \omega^2 |v|^2)(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger \hat{a}) \right]. \end{aligned} \quad (164)$$

Using  $\hat{a}|0\rangle = 0$  and  $[\hat{a}, \hat{a}^\dagger] = 1$  we hence find the following action of the Hamiltonian operator on the vacuum state

$$\hat{H}|0\rangle = \frac{1}{2}(\dot{v}^2 + \omega^2 v^2)^* \hat{a}^\dagger \hat{a}^\dagger |0\rangle + \frac{1}{2}(|\dot{v}|^2 + \omega^2 |v|^2)|0\rangle. \quad (165)$$

The requirement that  $|0\rangle$  be an eigenstate of  $\hat{H}$  means that the first term must vanish which implies the condition

$$\dot{v} = \pm i\omega v. \quad (166)$$

For such a function  $v$  the norm is

$$\langle v, v \rangle = \mp \frac{2\omega}{\hbar} |v|^2 \quad (167)$$

and positivity of the normalization condition  $\langle v, v \rangle > 0$  selects the minus sign in Eqn. (166)

$$\dot{v} = -i\omega v. \quad (168)$$

This yields the normalized *positive frequency solution* ( $\langle v, v \rangle = 1$ ) to the equation of motion

$$v(t) = \sqrt{\frac{\hbar}{2\omega}} e^{-i\omega t}. \quad (169)$$

With this choice of mode function  $v$  the Hamiltonian becomes

$$\hat{H} = \hbar\omega \left( \hat{N} + \frac{1}{2} \right), \quad (170)$$

for which the vacuum  $|0\rangle$  is the state of minimum energy  $\hbar\omega/2$ . If any function other than (169) is chosen to expand the position operator, then the state annihilated by  $\hat{a}$  is *not* the ground state of the oscillator.

## 11.4 Zero-Point Fluctuations in the Ground State

Consider the mean square expectation value of the position operator  $\hat{x}$  in the ground state  $|0\rangle$

$$\begin{aligned} \langle |\hat{x}|^2 \rangle &\equiv \langle 0 | \hat{x}^\dagger \hat{x} | 0 \rangle \\ &= \langle 0 | (v^* \hat{a}^\dagger + v \hat{a})(v \hat{a} + v^* \hat{a}^\dagger) | 0 \rangle \\ &= |v(\omega, t)|^2 \langle 0 | \hat{a} \hat{a}^\dagger | 0 \rangle \\ &= |v(\omega, t)|^2 \langle 0 | [\hat{a}, \hat{a}^\dagger] | 0 \rangle \\ &= |v(\omega, t)|^2. \end{aligned} \quad (171)$$

This characterizes the zero-point fluctuations of the position in the vacuum state as the square of the mode function

$$\langle |\hat{x}|^2 \rangle = |v(\omega, t)|^2 = \frac{\hbar}{2\omega}. \quad (172)$$

This is (almost) all we need to know about quantum mechanics to compute the fluctuation spectrum created by inflation. However, first we need to do quite a bit of work to derive the mode equation for the scalar mode of cosmological perturbations, *i.e.* the analogue of (155).

## 12 Quantum Fluctuations in de Sitter Space

We have finally come to the highlight of this lecture: the full computation of the quantum-mechanical fluctuations generated during inflation and their relation to cosmological perturbations. Our calculation follows closely the treatment by Maldacena [16].

## 12.1 Summary of the Computational Strategy

The last two sections might have bored you, but they provided important background for the computation of inflationary fluctuations. We have defined the gauge-invariant curvature perturbation  $\mathcal{R}$ . It is conserved outside of the horizon, so we can compute it at horizon exit and remain ignorant about the subhorizon physics during and after reheating until horizon re-entry of a given  $\mathcal{R}$ -mode. We have recalled the quantization of the simple harmonic oscillator, so by writing the equation of motion for  $\mathcal{R}$  in simple harmonic oscillator form we are in the position to study the quantization of scalar fluctuations during inflation.

Here is a summary of the steps we will perform in the following sections:

1. We expand the action for single-field slow-roll inflation to second order in fluctuations. Specifically, we derive the second-order expansion of the action in terms of  $\mathcal{R}$ . The action approach guarantees the correct normalization for the quantization of fluctuations.
2. From the action we derive the equation of motion for  $\mathcal{R}$  and show that it is of SHO form.
3. The mode equations for  $\mathcal{R}$  will be hard to solve exactly so we consider several approximate solutions valid during slow-roll evolution.
4. We promote the classical field  $\mathcal{R}$  to a quantum operator and quantize it. Quantization will lead to a boundary condition on the mode functions in the ultraviolet (*i.e.* on small scales when the mode is deep inside the horizon). This doesn't fix the mode function completely.
5. We define the vacuum state by matching our solutions to the Minkowski vacuum on small (subhorizon) scales. This fixes the mode functions completely and their large-scale limit is hence determined.
6. For this Bunch-Davies vacuum state we then compute the power spectrum of curvature fluctuations at horizon crossing. In **Lecture 3** we will relate the power spectrum at horizon crossing during inflation to the angular power spectrum of CMB fluctuations at recombination.

Enough talking, let's compute!

## 12.2 Scalar Perturbations

We consider single-field slow-roll models of inflation defined by the action

$$S = \frac{1}{2} \int d^4x \sqrt{-g} [R - (\nabla\phi)^2 - 2V(\phi)] , \quad (173)$$

in units where  $M_{\text{pl}}^{-2} \equiv 8\pi G = 1$ . To fix time and spatial reparameterizations we choose the following gauge for the dynamical fields  $g_{ij}$  and  $\phi$

$$\delta\phi = 0 , \quad g_{ij} = a^2[(1 - 2\mathcal{R})\delta_{ij} + h_{ij}] , \quad \partial_i h_{ij} = h_i^j = 0 . \quad (174)$$

In this gauge the inflaton field is unperturbed and all scalar degrees of freedom are parameterized by the metric fluctuation  $\mathcal{R}(t, \mathbf{x})$ . An important property of  $\mathcal{R}$  is that it remains constant outside

the horizon. We can therefore restrict our computation to correlation functions of  $\mathcal{R}$  at horizon crossing. The remaining metric perturbations  $\Phi$  and  $B$  are related to  $\mathcal{R}$  by the Einstein equations (in the ADM formalism (see Appendix B) these are pure constraint equations).

### 12.2.1 Free Field Action

With quite some effort (see Appendix B) one may expand the action (173) to second order in  $\mathcal{R}$

$$S_{(2)} = \frac{1}{2} \int d^4x a^3 \frac{\dot{\phi}^2}{H^2} \left[ \dot{\mathcal{R}}^2 - a^{-2} (\partial_i \mathcal{R})^2 \right]. \quad (175)$$

Defining the *Mukhanov variable*

$$v \equiv z\mathcal{R}, \quad \text{where} \quad z^2 \equiv a^2 \frac{\dot{\phi}^2}{H^2} = 2a^2\varepsilon \quad (176)$$

and transitioning to conformal time  $\tau$  leads to the action for a canonically normalized scalar

$$S_{(2)} = \frac{1}{2} \int d\tau d^3x \left[ (v')^2 + (\partial_i v)^2 + \frac{z''}{z} v^2 \right]. \quad (177)$$

**Exercise 7 (Mukhanov action)** *Confirm (177). Hint: use integration by parts.*

We define the Fourier expansion<sup>8</sup> of the field  $v$

$$v(\tau, \mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} v_{\mathbf{k}}(\tau) e^{i\mathbf{k}\cdot\mathbf{x}}, \quad (178)$$

where

$$v_k'' + \left( k^2 - \frac{z''}{z} \right) v_k = 0. \quad (179)$$

Here, we have dropped the vector notation  $\mathbf{k}$  on the subscript, since (179) depends only on the magnitude of  $k$ . The Mukhanov equation (179) is hard to solve in full generality since the function  $z$  depends on the background dynamics. For a given inflationary background one may solve (179) numerically. However, to gain a more intuitive understanding of the solutions we will discuss approximate analytical solutions in the pure de Sitter limit (§12.2.4) and in the slow-roll approximation (**Problem 7**).

### 12.2.2 Quantization

The quantization of the field  $v$  is performed in completely analogy with our treatment of the quantum harmonic oscillator in §11.

As before we promote the field  $v$  and its conjugate momentum  $v'$  to quantum operator

$$v \rightarrow \hat{v} = \int \frac{d\mathbf{k}^3}{(2\pi)^3} \left[ v_{\mathbf{k}}(\tau) \hat{a}_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} + v_{\mathbf{k}}^*(\tau) \hat{a}_{\mathbf{k}}^\dagger e^{-i\mathbf{k}\cdot\mathbf{x}} \right]. \quad (180)$$

---

<sup>8</sup>Note that the coordinates  $\mathbf{x}$  are comoving coordinates, and the wavevector  $\mathbf{k}$ , which does not redshift with expansion. The proper wavevector is  $\mathbf{k}_{\text{prop}} = \mathbf{k}/a(\tau)$ . Therefore, the comoving wavenumber  $\mathbf{k}$  is not itself dynamical, but is just a set of constants labeling a particular Fourier component.

Alternatively, the Fourier components  $v_{\mathbf{k}}$  are promoted to operators and expressed via the following decomposition

$$v_{\mathbf{k}} \rightarrow \hat{v}_{\mathbf{k}} = v_k(\tau) \hat{a}_{\mathbf{k}} + v_{-k}^*(\tau) \hat{a}_{-\mathbf{k}}^\dagger, \quad (181)$$

where the creation and annihilation operators  $\hat{a}_{-\mathbf{k}}^\dagger$  and  $\hat{a}_{\mathbf{k}}$  satisfy the canonical commutation relation

$$[\hat{a}_{\mathbf{k}}, \hat{a}_{\mathbf{k}'}^\dagger] = (2\pi)^3 \delta(\mathbf{k} - \mathbf{k}') \quad (182)$$

if and only if the mode functions are normalized as follows

$$\langle v_k, v_k \rangle \equiv \frac{i}{\hbar} (v_k^* v'_k - v_k^{*\prime} v_k) = 1, \quad (\dots)' \equiv \partial_\tau(\dots). \quad (183)$$

Eqn. (183) provides one of the boundary conditions on the solutions of (179). The second boundary conditions that fixes the mode functions completely comes from *vacuum selection*.

### 12.2.3 Boundary Conditions and Bunch-Davies Vacuum

We must choose a vacuum state for the fluctuations,

$$\hat{a}_{\mathbf{k}} |0\rangle = 0, \quad (184)$$

which corresponds to specifying an additional boundary conditions for  $v_k$  (see *e.g.* Chapter 3 in Birrel and Davies [17]). The standard choice is the Minkowski vacuum of a comoving observer in the far past (when all comoving scales were far inside the Hubble horizon),  $\tau \rightarrow -\infty$  or  $|k\tau| \gg 1$  or  $k \gg aH$ . In this limit the mode equation (179) becomes

$$v''_k + k^2 v_k = 0. \quad (185)$$

This is the equation of a simple harmonic oscillator with time-independent frequency! For this case we showed that a unique solution (169) exists if we require the vacuum to be the minimum energy state. Hence we impose the initial condition

$$\lim_{\tau \rightarrow -\infty} v_k = \frac{e^{-ik\tau}}{\sqrt{2k}}.$$

(186)

### 12.2.4 Solution in de Sitter Space

Consider the de Sitter limit  $\varepsilon \rightarrow 0$  ( $H = \text{const.}$ ) and

$$\frac{z''}{z} = \frac{a''}{a} = \frac{2}{\tau^2}. \quad (187)$$

In a de Sitter background we therefore wish to solve the mode equation

$$v''_k + \left( k^2 - \frac{2}{\tau^2} \right) v_k = 0.$$

(188)

**Exercise 8 (de Sitter Mode Functions)** Verify by direct substitution that an exact solution to (188) is

$$v_k = \alpha \frac{e^{-ik\tau}}{\sqrt{2k}} \left( 1 - \frac{i}{k\tau} \right) + \beta \frac{e^{ik\tau}}{\sqrt{2k}} \left( 1 + \frac{i}{k\tau} \right). \quad (189)$$

The free parameters  $\alpha$  and  $\beta$  characterize the non-uniqueness of the mode functions. However, we may fix  $\alpha$  and  $\beta$  to unique values by considering the quantization condition (183) together with the subhorizon limit,  $|k\tau| \gg 1$ , equation (186). This fixes  $\alpha = 1$ ,  $\beta = 0$  and leads to the unique Bunch-Davies mode functions

$$v_k = \frac{e^{-ik\tau}}{\sqrt{2k}} \left( 1 - \frac{i}{k\tau} \right). \quad (190)$$

### 12.2.5 Power Spectrum in Quasi-de Sitter

We then compute the power spectrum of  $\hat{\psi}_k \equiv a^{-1}\hat{v}_k$

$$\langle \hat{\psi}_k(\tau) \hat{\psi}_{k'}(\tau) \rangle = (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}') \frac{|v_k(\tau)|^2}{a^2} \quad (191)$$

$$= (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}') \frac{H^2}{2k^3} (1 + k^2\tau^2). \quad (192)$$

On superhorizon scales,  $|k\tau| \ll 1$ , this approaches a constant

$$\langle \hat{\psi}_k(\tau) \hat{\psi}_{k'}(\tau) \rangle \rightarrow (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}') \frac{H^2}{2k^3}. \quad (193)$$

or

$$\Delta_\psi^2 = \left( \frac{H}{2\pi} \right)^2. \quad (194)$$

The de Sitter result for  $\psi = v/a$  (193) allows us to compute the power spectrum of  $\mathcal{R} = \frac{H}{\dot{\phi}}\psi$  at horizon crossing,  $a(t_*)H(t_*) = k$ ,

$$\langle \mathcal{R}_k(t) \mathcal{R}_{k'}(t) \rangle = (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}') \frac{H_*^2}{2k^3} \frac{H_*^2}{\dot{\phi}_*^2}. \quad (195)$$

We define the dimensionless power spectrum  $\Delta_{\mathcal{R}}(k)$  by

$$\langle \mathcal{R}_k \mathcal{R}_{k'} \rangle = (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}') P_{\mathcal{R}}(k), \quad \Delta_{\mathcal{R}}^2(k) \equiv \frac{k^3}{2\pi^2} P_{\mathcal{R}}(k), \quad (196)$$

such that the real space variance of  $\mathcal{R}$  is  $\langle \mathcal{R} \mathcal{R} \rangle = \int_0^\infty \Delta_{\mathcal{R}}^2(k) d\ln k$ . This gives

$$\Delta_{\mathcal{R}}^2(k) = \frac{H_*^2}{(2\pi)^2} \frac{H_*^2}{\dot{\phi}_*^2}. \quad (197)$$

Since  $\mathcal{R}$  approaches a constant on super-horizon scales the spectrum at horizon crossing determines the future spectrum until a given fluctuation mode re-enters the horizon.

The fact that we compute the power spectrum at a specific instant (horizon crossing,  $a_*H_* = k$ ) implicitly extends the result for the de Sitter background to a slowly time-evolving quasi-de Sitter space. Different modes exit the horizon as slightly different times when  $a_*H_*$  has a different value. This procedure gives the correct result for the power spectrum during slow-roll inflation (we prove this more rigorously in **Problem 7**). For non-slow roll inflation the background evolution will have to be tracked more precisely and the Mukhanov Equation typically has to be integrated numerically.

### 12.2.6 Spatially-Flat Gauge

In the previous sections we used the comoving gauge ( $\delta\phi = 0$ ) to compute the scalar power spectrum. A popular alternative to obtain the same result is to use *spatially-flat gauge*. In spatially-flat gauge, perturbations in  $\mathcal{R}$  are related to perturbations in the inflaton field value<sup>9</sup>  $\delta\phi$ , *cf.* Eqn. (136) with  $\Psi = 0$

$$\mathcal{R} = H \frac{\dot{\delta\phi}}{\dot{\phi}} \equiv -H\delta t. \quad (198)$$

The power spectrum of  $\mathcal{R}$  and the power spectrum of inflaton fluctuations  $\delta\phi$  are therefore related as follows

$$\langle \mathcal{R}_\mathbf{k} \mathcal{R}_{\mathbf{k}'} \rangle = \left( \frac{H}{\dot{\phi}} \right)^2 \langle \delta\phi_\mathbf{k} \delta\phi_{\mathbf{k}'} \rangle. \quad (199)$$

Finally, in the case of slow-roll inflation, quantum fluctuations of a light scalar field ( $m_\phi \ll H$ ) in quasi-de Sitter space ( $H \approx \text{const.}$ ) scale with the Hubble parameter  $H$

$$\langle \delta\phi_\mathbf{k} \delta\phi_{\mathbf{k}'} \rangle = (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}') \frac{2\pi^2}{k^3} \left( \frac{H}{2\pi} \right)^2. \quad (200)$$

Inflationary quantum fluctuations therefore produce the following power spectrum for  $\mathcal{R}$

$$\Delta_{\mathcal{R}}^2(k) = \frac{H_*^2}{(2\pi)^2} \frac{H_*^2}{\dot{\phi}_*^2}. \quad (201)$$

## 12.3 Tensor Perturbations

Having discussed the quantization of scalar perturbation in some details, the corresponding calculation for tensor perturbations will appear almost trivial.

### 12.3.1 Action

By expansion of the Einstein-Hilbert action one may obtain the second order action for tensor fluctuations is

$$S_{(2)} = \frac{M_{\text{pl}}^2}{8} \int d\tau dx^3 a^2 [(h'_{ij})^2 - (\partial_l h_{ij})^2]. \quad (202)$$

Here, we have reintroduced explicit factors of  $M_{\text{pl}}$  to make  $h_{ij}$  manifestly dimensionless. Up to a normalization factor of  $\frac{M_{\text{pl}}}{2}$  this is the same as the action for a massless scalar field in an FRW universe.

We define the following Fourier expansion

$$h_{ij} = \int \frac{d^3 k}{(2\pi)^3} \sum_{s=+,\times} \epsilon_{ij}^s(k) h_{\mathbf{k}}^s(\tau) e^{i\mathbf{k}\cdot\mathbf{x}}, \quad (203)$$

where  $\epsilon_{ii} = k^i \epsilon_{ij} = 0$  and  $\epsilon_{ij}^s(k) \epsilon_{ij}^{s'}(k) = 2\delta_{ss'}$ .

$$S_{(2)} = \sum_s \int d\tau dk \frac{a^2}{4} M_{\text{pl}}^2 [h_{\mathbf{k}}^s h_{\mathbf{k}}^{s'} - k^2 h_{\mathbf{k}}^s h_{\mathbf{k}}^s] \quad (204)$$

---

<sup>9</sup>Intuitively, the curvature perturbation  $\mathcal{R}$  is related to a spatially varying time-delay  $\delta t(\mathbf{x})$  for the end of inflation. This time-delay is induced by the inflaton fluctuation  $\delta\phi$ .

We define the canonically normalized field

$$v_{\mathbf{k}}^s \equiv \frac{a}{2} M_{\text{pl}} h_{\mathbf{k}}^s \quad (205)$$

to get

$$\boxed{S_{(2)} = \sum_s \frac{1}{2} \int d\tau d^3\mathbf{k} \left[ (v_{\mathbf{k}}^s')^2 - \left( k^2 - \frac{a''}{a} \right) (v_{\mathbf{k}}^s)^2 \right]}, \quad (206)$$

where

$$\frac{a''}{a} = \frac{2}{\tau^2} \quad (207)$$

holds in de Sitter space. You will recognize this as effectively two copies of the action (177).

### 12.3.2 Quantization

Each polarization of the gravitational wave is therefore just a renormalized massless field in de Sitter space

$$h_{\mathbf{k}}^s = \frac{2}{M_{\text{pl}}} \psi_{\mathbf{k}}^s, \quad \psi_{\mathbf{k}}^s \equiv \frac{v_{\mathbf{k}}}{a}. \quad (208)$$

Since we computed the power spectrum of  $\psi = v/a$  in the previous section,  $\Delta_{\psi}^2 = (H/2\pi)^2$  we can simply right down the answer for  $\Delta_h^2$ , the power spectrum for a single polarization,

$$\Delta_h^2(k) = \frac{4}{M_{\text{pl}}^2} \left( \frac{H_{\star}}{2\pi} \right)^2. \quad (209)$$

### 12.3.3 Power Spectrum

The dimensionless power spectrum of tensor fluctuations therefore is

$$\boxed{\Delta_t^2 = 2\Delta_h^2(k) = \frac{2}{\pi^2} \frac{H_{\star}^2}{M_{\text{pl}}^2}}. \quad (210)$$

### 12.3.4 The Tensor-to-Scalar Ratio

Tensor fluctuations are often normalized relative to the (measured) amplitude of scalar fluctuations,  $\Delta_s^2 \equiv \Delta_R^2 \sim 10^{-10}$ . The tensor-to-scalar ratio is

$$r \equiv \frac{\Delta_t^2(k)}{\Delta_s^2} . \quad (211)$$

Since  $\Delta_s^2$  is fixed and  $\Delta_t^2 \propto H^2 \approx V$ , the tensor-to-scalar ratio is a direct measure of the energy scale of inflation

$$V^{1/4} \sim \left( \frac{r}{0.01} \right)^{1/4} 10^{16} \text{ GeV}. \quad (212)$$

## 12.4 The Lyth Bound

Note from equations (197) and (210) that the tensor-to-scalar ratio relates directly to the evolution of the inflaton as a function of  $e$ -folds  $N$

$$r = \frac{8}{M_{\text{pl}}^2} \left( \frac{d\phi}{dN} \right)^2. \quad (213)$$

The total field evolution between the time when CMB fluctuations excited the horizon at  $N_{\text{cmb}}$  and the end of inflation at  $N_{\text{end}}$  can therefore be written as the following integral

$$\frac{\Delta\phi}{M_{\text{pl}}} = \int_{N_{\text{end}}}^{N_{\text{cmb}}} dN \sqrt{\frac{r}{8}}. \quad (214)$$

During slow-roll  $r(N)$  (or  $\varepsilon$ ) doesn't evolve much and one may obtain the following approximate relation [18]

$$\frac{\Delta\phi}{M_{\text{pl}}} = \mathcal{O}(1) \times \left( \frac{r_\star}{0.01} \right)^{1/2} \quad (215)$$

where  $r_\star \equiv r(N_{\text{cmb}})$  is the tensor-to-scalar ratio on CMB scales.  $r > 0.01$  therefore correlates with  $\Delta\phi > M_{\text{pl}}$  or *large-field inflation*.

## 13 Primordial Spectra

The results for the power spectra of the scalar and tensor fluctuations created by inflation are

$$\Delta_s^2(k) \equiv \Delta_{\mathcal{R}}^2(k) = \frac{1}{8\pi^2} \frac{H^2}{M_{\text{pl}}^2} \frac{1}{\varepsilon} \Big|_{k=aH}, \quad (216)$$

$$\Delta_t^2(k) \equiv 2\Delta_h^2(k) = \frac{2}{\pi^2} \frac{H^2}{M_{\text{pl}}^2} \Big|_{k=aH}, \quad (217)$$

where

$$\varepsilon = -\frac{d \ln H}{dN}. \quad (218)$$

The horizon crossing condition  $k = aH$  makes (216) and (217) functions of the comoving wavenumber  $k$ . The tensor-to-scalar ratio is

$$r \equiv \frac{\Delta_t^2}{\Delta_s^2} = 16\varepsilon_\star. \quad (219)$$

### 13.1 Scale-Dependence

The scale dependence of the spectra follows from the time-dependence of the Hubble parameter and is quantified by the spectral indices

$$n_s - 1 \equiv \frac{d \ln \Delta_s^2}{d \ln k}, \quad n_t \equiv \frac{d \ln \Delta_t^2}{d \ln k}. \quad (220)$$

We split this into two factors

$$\frac{d \ln \Delta_s^2}{d \ln k} = \frac{d \ln \Delta_s^2}{dN} \times \frac{dN}{d \ln k} \quad (221)$$

The derivative with respect to  $e$ -folds is

$$\frac{d \ln \Delta_s^2}{dN} = 2 \frac{d \ln H}{dN} - \frac{d \ln \epsilon_H}{dN}. \quad (222)$$

The first term is just  $-2\varepsilon$  and the second term may be evaluated with the following result from Appendix C

$$\frac{d \ln \varepsilon}{dN} = 2(\varepsilon - \eta), \quad \text{where } \eta = -\frac{d \ln H_{,\phi}}{dN}. \quad (223)$$

The second factor in (221) is evaluated by recalling the horizon crossing condition  $k = aH$ , or

$$\ln k = N + \ln H. \quad (224)$$

Hence

$$\frac{dN}{d \ln k} = \left[ \frac{d \ln k}{dN} \right]^{-1} = \left[ 1 + \frac{d \ln H}{dN} \right]^{-1} \approx 1 + \varepsilon. \quad (225)$$

To first order in the Hubble slow-roll parameters we therefore find

$$n_s - 1 = 2\eta_\star - 4\varepsilon_\star. \quad (226)$$

Similarly, we find

$$n_t = -2\varepsilon_\star. \quad (227)$$

Any deviation from perfect scale-invariance ( $n_s = 1$  and  $n_t = 0$ ) is a indirect probe of the inflationary dynamics as quantified by  $\varepsilon$  and  $\eta$ .

## 13.2 Slow-Roll Results

In the slow-roll approximation the Hubble and potential slow-roll parameters are related as follows

$$\varepsilon \approx \epsilon_v, \quad \eta \approx \eta_v - \epsilon_v \quad (228)$$

The scalar and tensor spectra are then expressed purely in terms of  $V(\phi)$  and  $\epsilon$  (or  $V_{,\phi}$ )

$$\Delta_s^2(k) \approx \frac{1}{24\pi^2} \frac{V}{M_{\text{pl}}^4} \frac{1}{\epsilon_v} \Big|_{k=aH}, \quad \Delta_t^2(k) \approx \frac{2}{3\pi^2} \frac{V}{M_{\text{pl}}^4} \Big|_{k=aH}. \quad (229)$$

The scalar spectral index is

$$\boxed{n_s - 1 = 2\eta_v^\star - 6\epsilon_v^\star}. \quad (230)$$

The tensor spectral index is

$$\boxed{n_t = -2\epsilon_v^\star}, \quad (231)$$

and the tensor-to-scalar ratio is

$$\boxed{r = 16\epsilon_v^\star}. \quad (232)$$

We see that single-field slow-roll models satisfy a consistency condition between the tensor-to-scalar ratio  $r$  and the tensor tilt  $n_t$

$$\boxed{r = -8n_t}. \quad (233)$$

In the slow-roll approximation measurements of the scalar and tensor spectra relate directly to the shape of the potential  $V(\phi)$ , *i.e.*  $H$  is a measure of the scale of the potential,  $\epsilon_v$  of its first derivative  $V_{,\phi}$ ,  $\eta_v$  of its second derivative  $V_{,\phi\phi}$ , etc. Measurements of the amplitude and the scale-dependence of the cosmological perturbations therefore encode information about the potential driving the inflationary expansion. This allows to reconstruct a power series expansion of the potential around  $\phi_{\text{cmb}}$  (corresponding to the time when CMB fluctuations exited the horizon).

### 13.3 Case Study: $m^2\phi^2$ Inflation

Recall from **Lecture 1** the slow-roll parameters for  $m^2\phi^2$  inflation evaluated at  $\phi_\star = \phi_{\text{cmb}}$ , *i.e.*  $N_{\text{cmb}} \sim 60$  e-folds before the end of inflation

$$\epsilon_v^\star = \eta_v^\star = 2 \left( \frac{M_{\text{pl}}}{\phi_{\text{cmb}}} \right)^2 = \frac{1}{2N_{\text{cmb}}}. \quad (234)$$

To satisfy the normalization of scalar fluctuations,  $\Delta_s^2 \sim 10^{-10}$ , we need to fix the inflaton mass to  $m \sim 10^{-6} M_{\text{pl}}$ . To see this note that Eqn. (229) implies

$$\Delta_s^2 = \frac{m^2}{M_{\text{pl}}^2} \frac{N_{\text{cmb}}^2}{3}. \quad (235)$$

The scalar spectral index  $n_s$  and the tensor-to-scalar ratio  $r$  evaluated at CMB scales are

$$\boxed{n_s} = 1 + 2\eta_v^\star - 6\epsilon_v^\star = 1 - \frac{2}{N_{\text{cmb}}} \approx \boxed{0.96}, \quad (236)$$

and

$$\boxed{r} = 16\epsilon_v^\star = \frac{8}{N_{\text{cmb}}} \approx \boxed{0.1}. \quad (237)$$

These predictions of one of the simplest inflationary models are something to look out for in the near future.

## 14 Summary: Lecture 2

A defining characteristic of inflation is the behavior of the comoving Hubble radius  $1/(aH)$  which shrinks quasi-exponentially. A mode with comoving wavenumber  $k$  is called super-horizon when  $k < aH$ , and sub-horizon when  $k > aH$ . The inflaton is taken to be in a vacuum state, defined such that sub-horizon modes approach the Minkowski vacuum for  $k \gg aH$ . After a mode exits the horizon, it is described by a classical probability distribution with variance given by the power spectrum evaluated at horizon crossing

$$P_s(k) = \frac{H^2}{2k^3} \frac{H^2}{\dot{\phi}^2} \Big|_{k=aH} .$$

Inflation also produces fluctuations in the tensor part of the spatial metric. This corresponds to a spectrum of gravitational waves with power spectrum

$$P_t(k) = \frac{4}{k^3} \frac{H^2}{M_{pl}^2} \Big|_{k=aH} .$$

For slow-roll models the scalar and tensor spectra are expressed purely in terms of  $V(\phi)$  and  $\epsilon$  (or  $V_{,\phi}$ )

$$\Delta_s^2(k) \approx \frac{1}{24\pi^2} \frac{V}{M_{pl}^4} \frac{1}{\epsilon_v} \Big|_{k=aH} , \quad \Delta_t^2(k) \approx \frac{2}{3\pi^2} \frac{V}{M_{pl}^4} \Big|_{k=aH} ,$$

where  $\Delta^2(k) \equiv \frac{k^3}{2\pi^2} P(k)$ . The scale dependence is given by

$$\begin{aligned} n_s - 1 &\equiv \frac{d \ln \Delta_s^2}{d \ln k} = 2\eta_v - 6\epsilon_v \\ n_t &\equiv \frac{d \ln \Delta_t^2}{d \ln k} = -2\epsilon_v . \end{aligned}$$

The tensor-to-scalar ratio is

$$r \equiv \frac{\Delta_t^2}{\Delta_s^2} = 16\epsilon_v .$$

By the Lyth bound,  $r$  relates directly to total field excursion during inflation

$$\frac{\Delta\phi}{M_{pl}} \approx \left( \frac{r}{0.01} \right)^{1/2} .$$

A large value for  $r$  therefore correlates both with a high scale for the inflationary energy and a super-Planckian field evolution.

## 15 Problem Set: Lecture 2

**Problem 6 (Vacuum Selection)** *Read about the Unruh effect in your favorite resource for QFT in curved spacetime.*

**Problem 7 (Slow-Roll Mode Functions)** *In this problem we compute the mode functions and the power spectrum of curvature perturbations to first order in the slow-roll approximation.*

Recall the mode equation

$$v_k'' + \left( k^2 - \frac{z''}{z} \right) v_k = 0, \quad z^2 = 2a^2\varepsilon. \quad (238)$$

1. Show that

$$\frac{z''}{z} = \frac{\nu^2 - 1/4}{\tau^2}, \quad \nu \approx \frac{3}{2} + 3\varepsilon - \eta, \quad (239)$$

at first order in the slow-roll parameter

$$\varepsilon \equiv -\frac{\dot{H}}{H^2}, \quad \eta \equiv 2\varepsilon - \frac{\dot{\varepsilon}}{2H\varepsilon}. \quad (240)$$

The solution can then be expressed as a linear combination of Hankel functions

$$v_k(\tau) = x^{1/2} \left[ c_1 H_\nu^{(1)}(x) + c_2 H_\nu^{(2)}(x) \right], \quad x \equiv k|\tau|. \quad (241)$$

In the far past,  $x = k|\tau| \rightarrow \infty$ , the Hankel function have the asymptotic limit

$$H_\nu^{(1,2)}(x) \rightarrow \sqrt{\frac{2}{\pi x}} \exp \left[ \pm i \left( x - \frac{\nu\pi}{2} - \frac{\pi}{4} \right) \right] \quad (242)$$

2. Show that the boundary condition (186) implies

$$v_k = a_1 (\pi x / 4k)^{1/2} H_\nu^{(1)}(x), \quad (243)$$

where

$$a_1 = \exp[i(2\nu + 1)\pi/4] \quad (244)$$

is a  $k$ -independent complex phase factor.

3. Compute the power spectrum of  $\mathcal{R} = v/z$  at large scales,  $k \ll aH$ .

Hint: Use the identity

$$H_\nu^{(1)}(k|\tau|) \rightarrow \frac{i}{\pi} \Gamma(\nu) \left( \frac{k|\tau|}{2} \right)^{-\nu}, \quad \text{for } k\tau \rightarrow 0 \quad (245)$$

and  $\Gamma(3/2) = \sqrt{\pi}/2$ .

Show that you reproduce the result of perfect de Sitter in the limit  $\varepsilon = \eta = 0$ .

4. Read of the scale-dependence of the spectrum.

**Problem 8 (Predictions of  $\lambda\phi^4$  Inflation)** Determine the predictions of an inflationary model with a quartic potential

$$V(\phi) = \lambda\phi^4. \quad (246)$$

1. Compute the slow-roll parameters  $\epsilon$  and  $\eta$  in terms of  $\phi$ .
2. Determine  $\phi_{\text{end}}$ , the value of the field at which inflation ends.
3. To determine the spectrum, you will need to evaluate  $\epsilon$  and  $\eta$  at horizon crossing,  $k = aH$  (or  $-k\tau = 1$ ). Choose the wavenumber  $k$  to be equal to  $a_0 H_0$ , roughly the horizon today. Show that the requirement  $-k\tau = 1$  then corresponds to

$$e^{60} = \int_0^N dN' \frac{e^{N'}}{H(N')/H_{\text{end}}}, \quad (247)$$

where  $H_{\text{end}}$  is the Hubble rate at the end of inflation, and  $N$  is defined to be the number of  $e$ -folds before the end of inflation

$$N \equiv \ln \left( \frac{a_{\text{end}}}{a} \right). \quad (248)$$

4. Take this Hubble rate to be a constant in the above with  $H/H_{\text{end}} = 1$ . This implies that  $N \approx 60$ . Turn this into an expression for  $\phi$ . This simplest way to do this is to note that  $N = \int_t^{t_{\text{end}}} dt' H(t')$  and assume that  $H$  is dominated by potential energy. Show that this mode leaves the horizon when  $\phi = 22M_{\text{pl}}$ .
5. Determine the predicted value of  $n_s$ ,  $r$  and  $n_t$ . Compare these predictions to the latest WMAP5 data (see Lecture 3).
6. Estimate the scalar amplitude in terms of  $\lambda$ . Set  $\Delta_s^2 \approx 10^{-9}$ . What value does this imply for  $\lambda$ ?

This model illustrates many of the features of generic inflationary models. In it, (i) the field is of order – even greater than – the Planck scale, but (ii) the energy scale  $V$  is much smaller than Planckian because of (iii) the very small coupling constant.

## Part IV

# Lecture 3: Contact with Observations

---

### Abstract

In this lecture we describe the inverse problem of extracting information on the inflationary perturbation spectra from observations of the cosmic microwave background and the large-scale structure. We define the precise relations between the scalar and tensor power spectra computed in the previous lecture and the observed CMB anisotropies and the galaxy power spectrum. We describe the transfer functions that relate the primordial fluctuations to the late-time observables. We then use these results to discuss the current observational evidence for inflation. Finally, we indicate opportunities for future tests of inflation.

---

## 16 Connecting Observations to the Early Universe

“It doesn’t matter how beautiful your theory is, it doesn’t matter how smart you are or what your name is. If it doesn’t agree with experiment, it’s wrong.”

*Richard Feynman*

In the last lecture we computed the power spectra of the primordial scalar and tensor fluctuations  $\mathcal{R}$  and  $h$  at horizon exit. In this lecture we relate these results to observations of the cosmic microwave background (CMB) and the large-scale structure (LSS). Making this correspondence explicit is crucial for constraining the inflationary predictions.

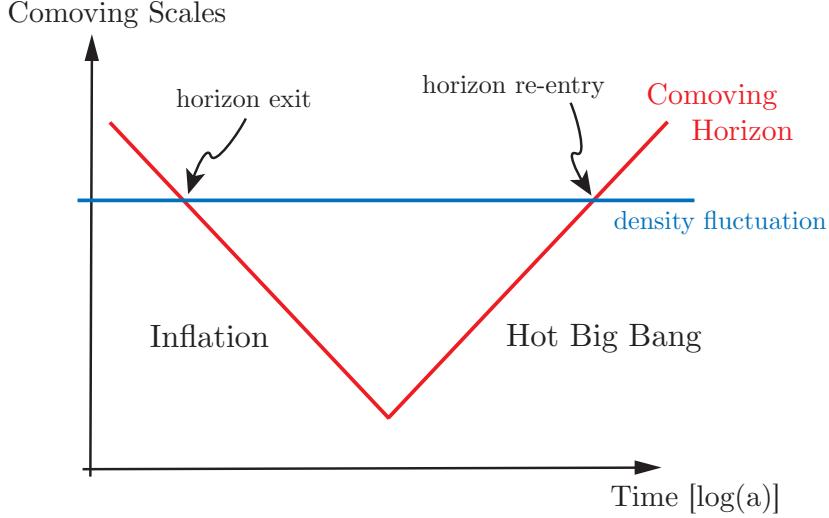
The curvature perturbation  $\mathcal{R}$  and the gravitational wave amplitude  $h$  both freeze at constant values once the mode exits the horizon,  $k = a(\tau_*)H(\tau_*)$ . In the previous lecture we therefore computed the primordial perturbations at the time of horizon exit,  $\tau_*$ . To relate this to a cosmological observable (like the CMB temperature or the density of galaxies) we need to

- i) relate  $\mathcal{R}$  (or  $h$ ) to the quantity  $\mathcal{Q}$  that is actually measured in an experiment and
- ii) take into account the time evolution of  $\mathcal{R}$  (and  $\mathcal{Q}$ ) once it re-enters the horizon.

Schematically, we may write

$$\mathcal{Q}_{\mathbf{k}}(\tau) = T_{\mathcal{Q}}(k, \tau, \tau_*) \mathcal{R}_{\mathbf{k}}(\tau_*), \quad (249)$$

where  $T_{\mathcal{Q}}$  is the *transfer function* between  $\mathcal{R}$  fluctuations at time  $\tau_*$  and  $\mathcal{Q}$  fluctuations at some later time  $\tau$ . As we have indicated the transfer function may depend on scale. The quantity  $\mathcal{Q}$  may be the temperature fluctuations measured by a CMB satellite such as the Wilkinson Microwave Anisotropy



**Figure 14:** Creation and evolution of perturbations in the inflationary universe. Fluctuations are created quantum mechanically on subhorizon scales (see **Lecture 2**). While comoving scales,  $k^{-1}$ , remain constant the comoving Hubble radius during inflation,  $(aH)^{-1}$ , shrinks and the perturbations exit the horizon and freeze until horizon re-entry at late times. After horizon re-entry the fluctuations evolve into anisotropies in the CMB and perturbations in the LSS. This time-evolution has to be accounted for to relate cosmological observations to the primordial perturbations laid down by inflation (see **Lecture 3**).

Probe (WMAP) or the galaxy density inferred in a galaxy survey such as the Sloan Digital Sky Survey (SDSS).

#### CMB anisotropies

The main result of §17 will be the following relation between the inflationary input spectra  $P(k) \equiv \{P_{\mathcal{R}}(k), P_h(k)\}$  and the angular power spectra of CMB temperature fluctuations and polarization

$$C_{\ell}^{XY} = \frac{2}{\pi} \int k^2 dk \underbrace{P(k)}_{\text{Inflation}} \underbrace{\Delta_{X\ell}(k) \Delta_{Y\ell}(k)}_{\text{Anisotropies}}, \quad (250)$$

where

$$\Delta_{X\ell}(k) = \int_0^{\tau_0} d\tau \underbrace{S_X(k, \tau)}_{\text{Sources}} \underbrace{P_{X\ell}(k[\tau_0 - \tau])}_{\text{Projection}}. \quad (251)$$

The labels  $X, Y$  refer to temperature  $T$  and polarization modes  $E$  and  $B$  (see §17). The integral (250) relates the inhomogeneities predicted by inflation,  $P(k)$ , to the anisotropies observed in the CMB,  $C_{\ell}^{XY}$ . The correlations between the different  $X$  and  $Y$  modes are related by the transfer functions  $\Delta_{X\ell}(k)$  and  $\Delta_{Y\ell}(k)$ . The transfer functions may be written as the line-of-sight integral (251) which factorizes into physical source terms  $S_X(k, \tau)$  and geometric projection factors  $P_{X\ell}(k[\tau_0 - \tau])$  (combinations of Bessel functions). A derivation of the source terms and the projection factors is beyond the scope of this lecture, but may be found in Dodelson's book [5]. An intuitive explanation for these results may be found on Wayne Hu's website.

Our interest in this lecture lies in experimental constraints on the primordial power spectra  $P_{\mathcal{R}}(k)$  and  $P_h(k)$ . To measure the primordial spectra the observed CMB anisotropies  $C_{\ell}^{XY}$  need to be *deconvolved* by taking into account the appropriate transfer functions and projection effects, *i.e.* for a given background cosmology we can compute the evolution and projection effects in (250) and therefore extract the inflationary initial conditions  $P(k)$ . By this deconvolution procedure, the CMB provides a fascinating probe of the early universe.

### *Large-scale structure*

To study fluctuations in the matter distribution (as measured *e.g.* by the distribution of galaxies on the sky) we define the density contrast  $\delta \equiv \delta\rho/\bar{\rho}$ . We distinguish between fluctuations in the density of galaxies  $\delta_g$  and the dark matter density  $\delta$ . A common assumption is that galaxies are (biased) tracers of the underlying dark matter distribution,  $\delta_g = b\delta$ . If we have an independent way of determining the *bias* parameter  $b$ , we can use observations of the galaxy density contrast  $\delta_g$  to infer the underlying dark matter distribution  $\delta$ . The late-time power spectrum of dark matter density fluctuations is related to the primordial spectrum of curvature fluctuations as follows

$$P_{\delta}(k, \tau) = \frac{4}{25} \left( \frac{k}{aH} \right)^4 T_{\delta}^2(k, \tau) P_{\mathcal{R}}(k). \quad (252)$$

The numerical factor and the  $k$ -scaling that have been factored out from the transfer function is conventional. The transfer function  $T_{\delta}$  reflects the relative growth of fluctuations during matter domination,  $\delta \sim a$ , and radiation domination,  $\delta \sim \ln a$ . It usually has to be computed numerically using codes such as **CMBFAST** or **CAMB**, however, in §18.1 we will cite useful fitting functions. Again, since for a fixed background cosmology the transfer function can be assumed as given, observations of the matter power spectrum can be a probe of the initial fluctuations from the early universe.

## 17 Review: The Cosmic Microwave Background

We give a very brief review of the physics and the statistical interpretation of CMB fluctuations. More details may be found in Dodelson's book [5] or Prof. Pierpaoli's lectures at this school.

### 17.1 Temperature Anisotropies

#### 17.1.1 Harmonic Expansion

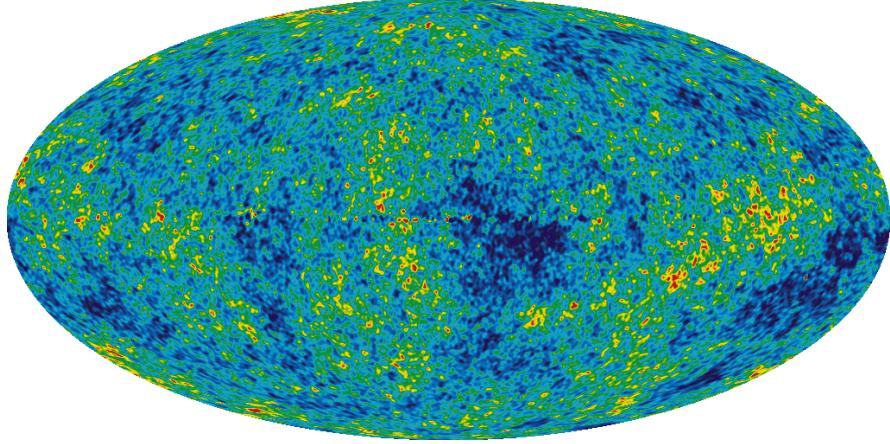
Figure 15 shows a map of the measured CMB temperature fluctuations  $\Delta T(\hat{n})$  relative to the background temperature  $T_0 = 2.7$  K. Here the unit vector  $\hat{n}$  denotes the direction in sky. The harmonic expansion of this map is

$$\Theta(\hat{n}) \equiv \frac{\Delta T(\hat{n})}{T_0} = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{n}), \quad (253)$$

where

$$a_{\ell m} = \int d\Omega Y_{\ell m}^*(\hat{n}) \Theta(\hat{n}). \quad (254)$$

Here,  $Y_{\ell m}(\hat{n})$  are the standard spherical harmonics on a 2-sphere with  $\ell = 0$ ,  $\ell = 1$  and  $\ell = 2$  corresponding to the monopole, dipole and quadrupole, respectively. The magnetic quantum



**Figure 15:** Temperature fluctuations in the CMB. Blue spots represent directions on the sky where the CMB temperature is  $\sim 10^{-5}$  below the mean,  $T_0 = 2.7$  K. This corresponds to photons losing energy while climbing out of the gravitational potentials of overdense regions in the early universe. Yellow and red indicate hot (underdense) regions. The statistical properties of these fluctuations contain important information about both the background evolution and the initial conditions of the universe.

numbers satisfy  $m = -\ell, \dots, +\ell$ . The multipole moments  $a_{\ell m}$  may be combined into the rotationally-invariant angular power spectrum

$$C_\ell^{TT} = \frac{1}{2\ell + 1} \sum_m \langle a_{\ell m}^* a_{\ell m} \rangle, \quad \text{or} \quad \langle a_{\ell m}^* a_{\ell' m'} \rangle = C_\ell^{TT} \delta_{\ell\ell'} \delta_{mm'} . \quad (255)$$

The angular power spectrum is an important tool in the statistical analysis of the CMB. It describes the cosmological information contained in the millions of pixels of a CMB map in terms of a much more compact data representation. Figure 16 shows the most recent measurements of the CMB angular power spectrum. The figure also shows a fit of the theoretical predictions for the CMB spectrum to the data. The theoretical curve depends both on the background cosmological parameters and on the spectrum of initial fluctuations. We hence can use the CMB as a probe of both.

CMB temperature fluctuations are dominated by the *scalar* modes  $\mathcal{R}$  (at least for the values of the tensor-to-scalar ratio now under consideration,  $r < 0.3$ ). The linear evolution which relates  $\mathcal{R}$  and  $\Delta T$  is mediated by the transfer function  $\Delta_{T\ell}(k)$  through the  $k$ -space integral [5]

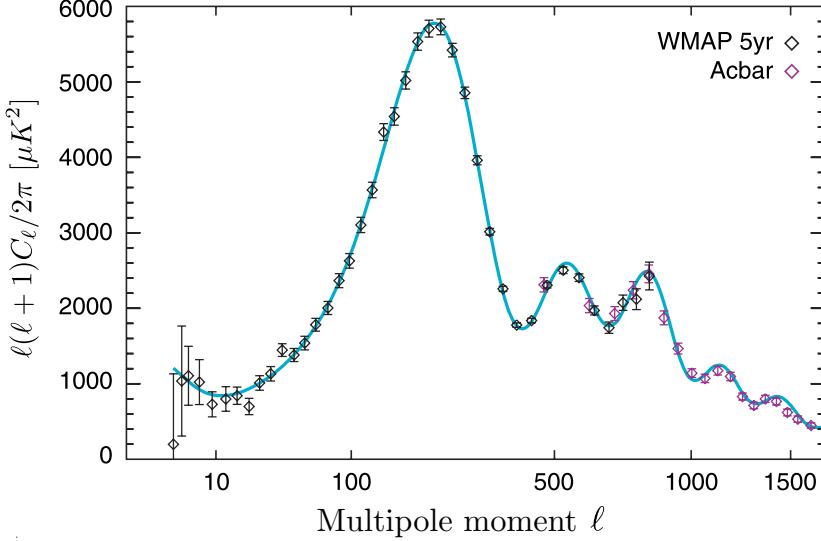
$$a_{\ell m} = 4\pi(-i)^\ell \int \frac{d^3 k}{(2\pi)^3} \Delta_{T\ell}(k) \mathcal{R}_{\mathbf{k}} Y_{\ell m}(\hat{\mathbf{k}}) . \quad (256)$$

Substituting (256) into (255) and using the identity

$$\sum_{m=-\ell}^{\ell} Y_{\ell m}(\hat{\mathbf{k}}) Y_{\ell m}(\hat{\mathbf{k}}') = \frac{2\ell + 1}{4\pi} P_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}') , \quad (257)$$

we find

$$C_\ell^{TT} = \frac{2}{\pi} \int k^2 dk \underbrace{P_{\mathcal{R}}(k)}_{\text{Inflation}} \underbrace{\Delta_{T\ell}(k) \Delta_{T\ell}(k)}_{\text{Anisotropies}} . \quad (258)$$



**Figure 16:** Angular power spectrum of CMB temperature fluctuations.

The transfer functions  $\Delta_{T\ell}(k)$  generally have to be computed numerically using Boltzmann-codes such as **CMBFAST** or **CAMB**. They depend on the parameters of the background cosmology. Assuming a fixed background cosmology the shape of the power spectrum  $C_\ell^{TT}$  contains information about the initial conditions as described by the primordial power spectrum  $P_{\mathcal{R}}(k)$ .<sup>10</sup> Of course, learning from observations about  $P_{\mathcal{R}}(k)$  and hence about inflation is the primary objective of this lecture.

### 17.1.2 Large Scales

On large scales, modes were still outside of the horizon at recombination. The large-scale CMB spectrum has therefore not been affected by subhorizon evolution and is simply the geometric projection of the primordial spectrum from recombination to us today. In this *Sachs-Wolfe regime* the transfer function  $\Delta_{T\ell}(k)$  is simply a Bessel function [5]

$$\Delta_{T\ell}(k) = \frac{1}{3} j_\ell(k[\tau_0 - \tau_{\text{rec}}]). \quad (259)$$

The angular power spectrum on large scales (small  $\ell$ ) therefore is

$$C_\ell^{TT} = \frac{2}{9\pi} \int k^2 dk P_{\mathcal{R}}(k) j_\ell^2(k[\tau_0 - \tau_{\text{rec}}]). \quad (260)$$

The Bessel projection function is peaked at  $k[\tau_0 - \tau_{\text{rec}}] \approx \ell$  and so effectively acts like a  $\delta$ -function mapping between  $k$  and  $\ell$ . Given that modes with wavenumber  $k \approx \ell/(\tau_0 - \tau_{\text{rec}})$  dominate the integral in (260), we can write

$$C_\ell^{TT} \propto k^3 P_{\mathcal{R}}(k) \Big|_{k \approx \ell/(\tau_0 - \tau_{\text{rec}})} \underbrace{\int d \ln x j_\ell^2(x)}_{\propto \ell(\ell+1)} \quad (261)$$

---

<sup>10</sup>In practice, the CMB data is fit simultaneously to the background cosmology and a spectrum of fluctuations.

Hence,

$$\ell(\ell+1)C_\ell^{TT} \propto \Delta_s^2(k)|_{k \approx \ell/(\tau_0 - \tau_{\text{rec}})} \propto \ell^{n_s-1}. \quad (262)$$

For a scale-invariant input spectrum,  $n_s = 1$ , the quantity

$$\mathcal{C}_\ell \equiv \frac{\ell(\ell+1)}{2\pi} C_\ell^{TT} \quad (263)$$

is independent of  $\ell$  (except for a rise at very low  $\ell$  due to the integrated Sachs-Wolfe effect arising from the late-time evolution of the gravitational potential in a dark energy dominated universe). This explains why the CMB power spectrum is often plotted for  $\mathcal{C}_\ell$ .

### 17.1.3 Non-Gaussianity

So far we have shown that the angular power spectrum of CMB fluctuations essentially is a measure of the primordial power spectrum  $P_{\mathcal{R}}(k)$  if we take into account subhorizon evolution and geometric projection effects

$$\langle a_{\ell m}^* a_{\ell' m'} \rangle = C_\ell^{TT} \delta_{\ell\ell'} \delta_{mm'} \Leftrightarrow \langle \mathcal{R}_{\mathbf{k}} \mathcal{R}_{\mathbf{k}'} \rangle = (2\pi)^3 P_{\mathcal{R}}(k) \delta(\mathbf{k} + \mathbf{k}'). \quad (264)$$

If the primordial fluctuations are Gaussian then  $P_{\mathcal{R}}(k)$  contains all the information. Single-field slow-roll inflation in fact predicts that  $\mathcal{R}$  should be Gaussian to a very high degree [16]. However, as we explain below, even a small amount of non-Gaussianity would provide crucial information about the inflationary action as it would require to go beyond the simplest single-field slow-roll models.

The primary measure of non-Gaussianity is the *three-point function* or equivalently the *bispectrum*

$$\langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k}_2} \mathcal{R}_{\mathbf{k}_3} \rangle = (2\pi)^3 B_{\mathcal{R}}(k_1, k_2, k_3) \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3). \quad (265)$$

In the CMB a non-zero bispectrum  $B_{\mathcal{R}}(k_1, k_2, k_3)$  leaves a signature in the angular bispectrum

$$B_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} = \langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle. \quad (266)$$

Substituting (256) into (266) we may relate the primordial bispectrum to the observed CMB bispectrum [19].<sup>11</sup>

Note that the primordial bispectrum  $B_{\mathcal{R}}(k_1, k_2, k_3)$  is a function of three momenta subject only to momentum conservation (*i.e.* the three vectors  $\mathbf{k}_i$  form a closed triangle). This makes observational constraints on non-Gaussianity challenging (there are many different forms of non-Gaussianity to consider), but also means that if detected non-Gaussianity contains a lot of information about the physics of the early universe.

A simple model of primordial non-Gaussianity is *local non-Gaussianity* defined by a Taylor expansion of the curvature perturbation around the Gaussian part  $\mathcal{R}_g$

$$\mathcal{R}(\mathbf{x}) = \mathcal{R}_g(\mathbf{x}) + \frac{3}{5} f_{\text{NL}}^{\mathcal{R}} \star \mathcal{R}_g^2(\mathbf{x}). \quad (267)$$

This is local in real space and the parameter  $f_{\text{NL}}^{\mathcal{R}}$  characterizes the level of non-Gaussianity. The reader is invited to show that (303) implies the following simple form for the bispectrum

$$B_{\mathcal{R}}(k_1, k_2, k_3) = \frac{6}{5} f_{\text{NL}}^{\mathcal{R}} \left[ P_{\mathcal{R}}(k_1)P_{\mathcal{R}}(k_2) + P_{\mathcal{R}}(k_2)P_{\mathcal{R}}(k_3) + P_{\mathcal{R}}(k_3)P_{\mathcal{R}}(k_1) \right]. \quad (268)$$

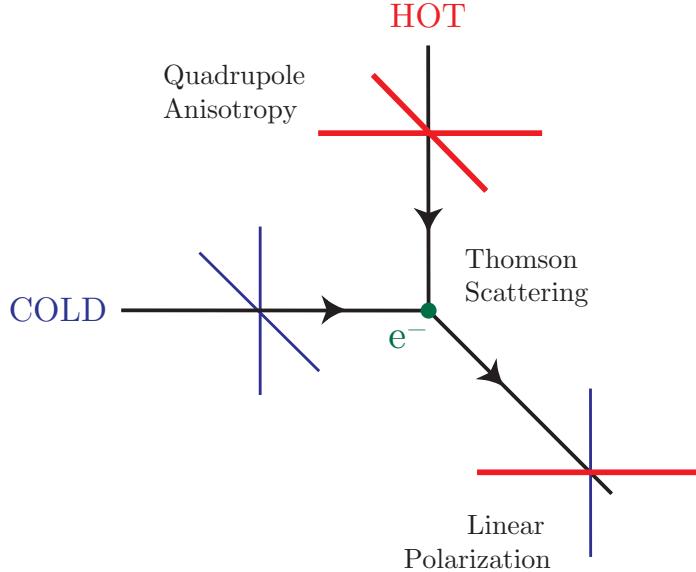
Present observational constraints on non-Gaussianity are therefore often phrased as constraints on the parameter  $f_{\text{NL}}^{\mathcal{R}}$  (see §29.2).

---

<sup>11</sup>Non-linear evolution can lead to additional non-Gaussianity (see **Lecture 4**).

## 17.2 Polarization

CMB polarization is likely to become one of the most important tools to probe the physics governing the early universe. In addition to anisotropies in the CMB temperature, we expect the CMB to become polarized via Thomson scattering [5]. As we now explain, this polarization contains crucial information about the primordial fluctuations and hence about inflation [4].



**Figure 17:** Thomson scattering of radiation with a quadrupole anisotropy generates linear polarization. Red colors (thick lines) represent hot radiation, and blue colors (thin lines) cold radiation.

### 17.2.1 Polarization via Thomson Scattering

Thomson scattering between electrons and photons produces a simple relationship between temperature anisotropy and polarization. If a free electron ‘sees’ an incident radiation pattern that is isotropic, then the outgoing radiation remains unpolarized because orthogonal polarization directions cancel out. However, if the incoming radiation field has a quadrupole component, a net linear polarization is generated via Thomson scattering (see Figure 17). A quadrupole moment in the radiation field is generated when photons decouple from the electrons and protons just before recombination. Hence linear polarization results from the velocities of electrons and protons on scales smaller than the photon diffusion length scale. Since both the velocity field and the temperature anisotropies are created by primordial density fluctuations, a component of the polarization should be correlated with the temperature anisotropy.

### 17.2.2 Characterization of the Radiation Field

The mathematical characterization of CMB polarization anisotropies is slightly more involved than that the description of temperature fluctuations because polarization is not a scalar field so the standard expansion in terms of spherical harmonics is not applicable.

The anisotropy field is defined in terms of a  $2 \times 2$  intensity tensor  $I_{ij}(\hat{n})$ , where as before  $\hat{n}$  denotes the direction on the sky. The components of  $I_{ij}$  are defined relative to two orthogonal basis vectors  $\hat{\mathbf{e}}_1$  and  $\hat{\mathbf{e}}_2$  perpendicular to  $\hat{n}$ . Linear polarization is then described by the Stokes parameters  $Q = \frac{1}{4}(I_{11} - I_{22})$  and  $U = \frac{1}{2}I_{12}$ , while the temperature anisotropy is  $T = \frac{1}{4}(I_{11} + I_{22})$ . The polarization magnitude and angle are  $P = \sqrt{Q^2 + U^2}$  and  $\alpha = \frac{1}{2}\tan^{-1}(U/Q)$ . The quantity  $T$  is invariant under a rotation in the plane perpendicular to  $\hat{n}$  and hence may be expanded in terms of scalar (spin-0) spherical harmonics (253). The quantities  $Q$  and  $U$ , however, transform under rotation by an angle  $\psi$  as a spin-2 field  $(Q \pm iU)(\hat{n}) \rightarrow e^{\mp 2i\psi}(Q \pm iU)(\hat{n})$ . The harmonic analysis of  $Q \pm iU$  therefore requires expansion on the sphere in terms of tensor (spin-2) spherical harmonics [20–22]

$$(Q \pm iU)(\hat{n}) = \sum_{\ell,m} a_{\pm 2,\ell m} \pm_2 Y_{\ell m}(\hat{n}). \quad (269)$$

A description of the mathematical properties of these tensor spherical harmonics would take us to far off the main track of this lecture, so we refer the reader to the classic papers [20, 21] or Dodelson’s book [5].

### 17.2.3 $E$ and $B$ -modes

Instead of the moments  $a_{\pm 2,\ell m}$  it is convenient to introduce the linear combinations

$$a_{E,\ell m} \equiv -\frac{1}{2}(a_{2,\ell m} + a_{-2,\ell m}), \quad a_{B,\ell m} \equiv -\frac{1}{2i}(a_{2,\ell m} - a_{-2,\ell m}). \quad (270)$$

Then one can define two scalar (spin-0) fields instead of the spin-2 quantities  $Q$  and  $U$

$$E(\hat{n}) = \sum_{\ell,m} a_{E,\ell m} Y_{\ell m}(\hat{n}), \quad B(\hat{n}) = \sum_{\ell,m} a_{B,\ell m} Y_{\ell m}(\hat{n}). \quad (271)$$

$E$  and  $B$  completely specify the linear polarization field.  $E$ -polarization is often also characterized as a *curl-free* mode with polarization vectors that are radial around cold spots and tangential around hot spots on the sky. In contrast,  $B$ -polarization is *divergence-free* but has a *curl*: its polarization vectors have vorticity around any given point on the sky.<sup>12</sup> Fig. 18 gives examples of  $E$ - and  $B$ -mode patterns. Although  $E$  and  $B$  are both invariant under rotations, they behave differently under parity transformations. Note that when reflected about a line going through the center, the  $E$ -patterns remain unchanged, while the  $B$ -patterns change sign.

The symmetries of temperature and polarization ( $E$ - and  $B$ -mode) anisotropies allow four types of correlations: the autocorrelations of temperature fluctuations and of  $E$ - and  $B$ -modes denoted by  $TT$ ,  $EE$ , and  $BB$ , respectively, as well as the cross-correlation between temperature fluctuations and  $E$ -modes:  $TE$ . All other correlations ( $TB$  and  $EB$ ) vanish for symmetry reasons.

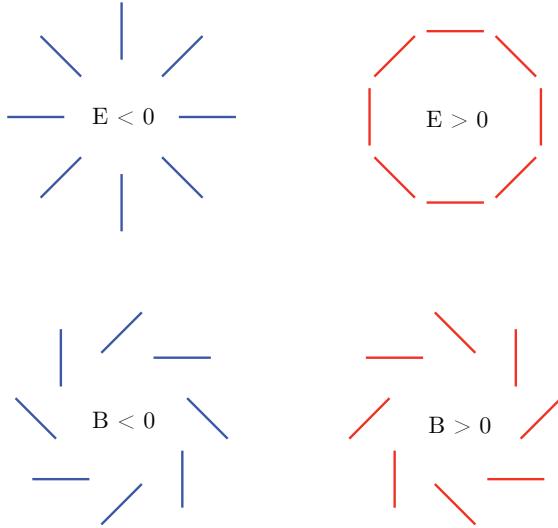
The angular power spectra are defined as before

$$C_\ell^{XY} \equiv \frac{1}{2\ell+1} \sum_m \langle a_{X,\ell m}^* a_{Y,\ell m} \rangle, \quad X, Y = T, E, B. \quad (272)$$

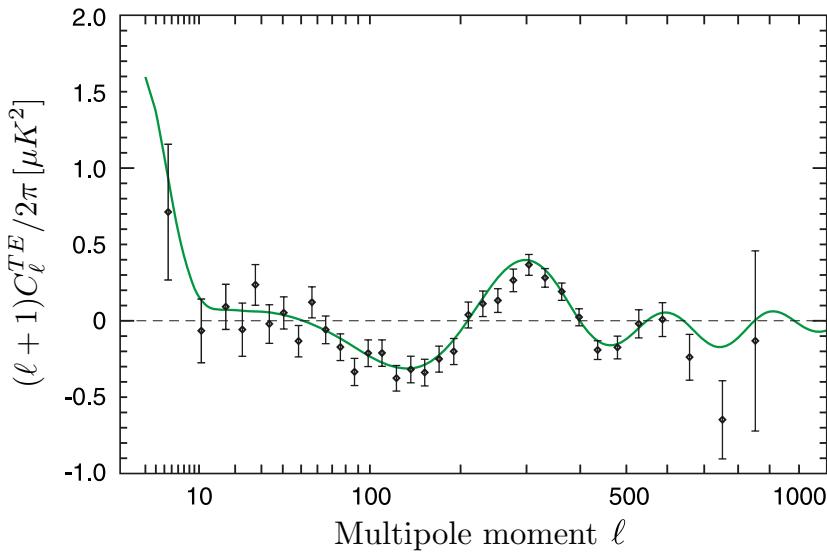
In Fig. 19 we show the latest measurement of the  $TE$  cross-correlation [23]. The  $EE$  spectrum has now begun to be measured, but the errors are still large. So far there are only upper limits on the  $BB$  spectrum, but no detection.

---

<sup>12</sup>Evidently the  $E$  and  $B$  nomenclature reflects the properties familiar from electrostatics,  $\nabla \times \mathbf{E} = 0$  and  $\nabla \cdot \mathbf{B} = 0$ .



**Figure 18:** Examples of  $E$ -mode and  $B$ -mode patterns of polarization. Note that if reflected across a line going through the center the  $E$ -patterns are unchanged, while the positive and negative  $B$ -patterns get interchanged.



**Figure 19:** Power spectrum of the cross-correlation between temperature and  $E$ -mode polarization anisotropies [23]. The anti-correlation for  $\ell = 50 - 200$  (corresponding to angular separations  $5^\circ > \theta > 1^\circ$ ) is a distinctive signature of adiabatic fluctuations on superhorizon scales at the epoch of decoupling [24, 25], confirming a fundamental prediction of the inflationary paradigm.

The cosmological significance of the  $E/B$  decomposition of CMB polarization was realized by the authors of Refs. [20, 21], who proved the following remarkable facts:

- i) scalar (density) perturbations create only  $E$ -modes and *no*  $B$ -modes.
- ii) vector (vorticity) perturbations create mainly  $B$ -modes.<sup>13</sup>
- iii) tensor (gravitational wave) perturbations create both  $E$ -modes and  $B$ -modes.

The fact that *scalars do not produce  $B$ -modes* while *tensors do* is the basis for the often-quoted statement that detection of  $B$ -modes is a smoking gun of tensor modes, and therefore of inflation.

#### 17.2.4 $E$ -modes and Scalars

The power spectrum of  $E$ -modes and the  $TE$  cross-correlation is dominated by inflationary *scalar* modes, *i.e.*

$$C_\ell^{EE} \approx (4\pi)^2 \int k^2 dk \underbrace{P_{\mathcal{R}}(k)}_{\text{Inflation}} \Delta_{E\ell}^2(k), \quad (273)$$

$$C_\ell^{TE} \approx (4\pi)^2 \int k^2 dk \underbrace{P_{\mathcal{R}}(k)}_{\text{Inflation}} \Delta_{T\ell}(k) \Delta_{E\ell}(k). \quad (274)$$

Like  $C_\ell^{TT}$ , the spectra  $C_\ell^{EE}$  and  $C_\ell^{TE}$  provide information about  $P_{\mathcal{R}}(k)$ . However, since the primordial spectrum is convolved with different transfer functions in each case (polarization is generated only by scattering from free electrons), the signals are usefully complementary.

#### 17.2.5 $B$ -modes and Tensors

$B$ -modes are only generated by *tensor* modes, *i.e.*

$$C_\ell^{BB} = (4\pi)^2 \int k^2 dk \underbrace{P_h(k)}_{\text{Inflation}} \Delta_{B\ell}^2(k). \quad (275)$$

Measuring  $C_\ell^{BB}$  is therefore a unique opportunity to access information about primordial tensor fluctuations.

## 18 Review: Large-Scale Structure

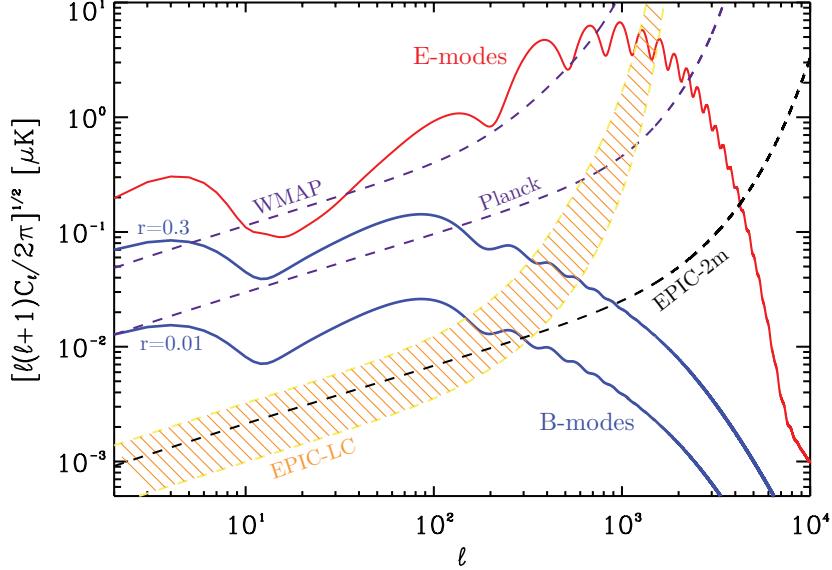
The galaxy (or dark matter) power spectrum is a measure of the spectrum of primordial curvature fluctuations

$$P_\delta(k, z) \Rightarrow P_{\mathcal{R}}(k) \quad (276)$$

if the effects of subhorizon evolution are accounted for. This is done by the dark matter transfer function.

---

<sup>13</sup>However, vectors decay with the expansion of the universe and are therefore believed to be subdominant at recombination. We therefore do not consider them here.



**Figure 20:**  $E$ - and  $B$ -mode power spectra for a tensor-to-scalar ratio saturating current bounds,  $r = 0.3$ , and for  $r = 0.01$ . Shown are also the experimental sensitivities for *WMAP*, *Planck* and two different realizations of a future CMB satellite (*CMBPol*) (EPIC-LC and EPIC-2m).

## 18.1 Dark Matter Transfer Functions

Density fluctuations evolve under the competing influence of pressure and gravity. During radiation domination the large radiation pressure prevents the rapid growth of fluctuations; the density contrast only grows logarithmically,  $\delta \sim \ln a$ . During matter domination the background pressure is negligible and gravitational collapse operates more effectively,  $\delta \sim a$ .

Under the simplifying assumption that there is no significant growth of perturbations between the time of horizon entry and matter domination one may derive the following approximate transfer function

$$T_\delta(k) \approx \begin{cases} 1 & k < k_{\text{eq}} \\ (k_{\text{eq}}/k)^2 & k > k_{\text{eq}} \end{cases}. \quad (277)$$

Although (277) is intuitively appealing for understanding the qualitative shape of the spectrum (*i.e.* the break in the spectrum at  $k \approx k_{\text{eq}}$ ), it is not accurate enough for most applications. A famous fitting function for the matter transfer function was given by Bardeen et al. (BBKS) [26]

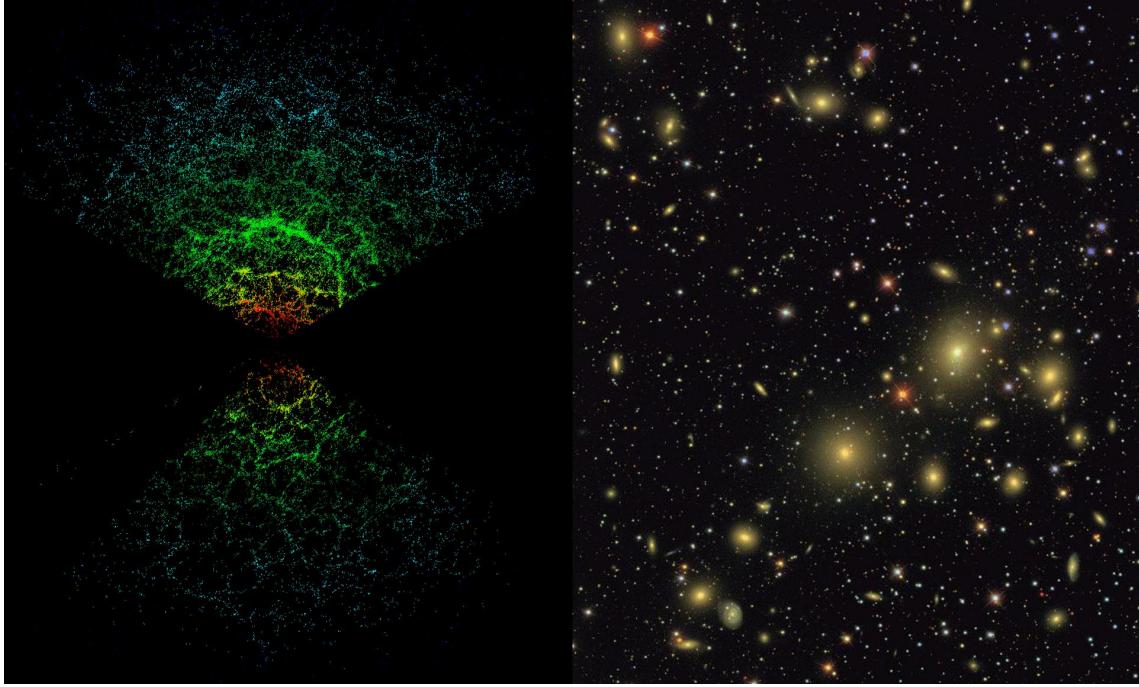
$$T_\delta(q) = \frac{\ln(1 + 2.34q)}{2.34q} [1 + 3.89q + (1.61q)^2 + (5.46q)^3 + (6.71q)^4]^{-1/4}, \quad (278)$$

where  $q = k/\Gamma h \text{ Mpc}^{-1}$  and we defined the shape parameter

$$\Gamma \equiv \Omega h \exp(-\Omega_b - \sqrt{2h}\Omega_b/\Omega). \quad (279)$$

More accurate transfer functions may be found in Eisenstein and Hu [27]. Finally, exact transfer functions may be computed numerically with **CMBFAST**.

For our purposes it is only important to note that (given the background cosmological parameters) the dark matter transfer function can be computed and used to relate the dark matter power spectrum  $P_\delta(k, z)$  to the inflationary spectrum  $P_R(k)$ .



**Figure 21:** Distribution of galaxies. The Sloan Digital Sky Survey (SDSS) has measured the positions and distances (redshifts) of nearly a million galaxies. Galaxies first identified on 2d images, like the one shown above on the right, have their distances measured to create the 3d map. The left image shows a slice of such a 3d map. The statistical properties of the measured distribution of galaxies reveal important information about the structure and evolution of the late time universe.

## 18.2 Galaxy Bias

With the exception of gravitational lensing we unfortunately can't observe the dark matter directly. What we observe (*e.g.* in galaxy surveys like the Sloan Digital Sky Survey (SDSS)) is luminous or baryonic matter. On large scales the following phenomenological ansatz for relating the galaxy distribution and the dark matter has proven useful

$$\delta_g = b \delta, \quad (280)$$

or

$$P_{\delta_g} = b^2 P_\delta. \quad (281)$$

Here,  $b$  is called the (linear) bias parameter. It may be viewed as a parameter describing the ill-understood physics of galaxy formation. The bias parameter  $b$  can be obtained by measuring the galaxy bispectrum  $B_{\delta_g}$ .

Modulo these complications the galaxy power spectrum  $P_{\delta_g}(k)$  is an additional probe of inflationary scalar fluctuations  $P_{\mathcal{R}}(k)$ . As it probes smaller scales it is complementary to observations of the CMB fluctuations.

# 19 Current Evidence for Inflation

Inflation is a hypothesis. In order to increase our confidence that inflation describes the physical reality of the early universe, we should compare the predictions of inflation to cosmological observations. In this section we describe the current observational evidence for inflation, before discussing future tests in the next section.

## 19.1 Flatness

The universe is filled with baryons, dark matter, photons, neutrinos and dark energy

$$\Omega_{\text{tot}} = \Omega_b + \Omega_{\text{cdm}} + \Omega_\gamma + \Omega_\nu + \Omega_\Lambda . \quad (282)$$

The value of  $\Omega_{\text{tot}}$  determines the spatial geometry of the universe with  $\Omega_{\text{tot}} = 1$  corresponding to a flat universe,  $\Omega_{\text{tot}} < 1$  to a negatively curved universe and  $\Omega_{\text{tot}} > 1$  to a positively curved universe. Inflation predicts

$$\Omega_{\text{tot}} = 1 \pm 10^{-5} , \quad (283)$$

while the data shows [23]

$$\Omega_{\text{tot}} = 1 \pm 0.02 . \quad (284)$$

Although this agreement between theory and data is impressive, one could argue that inflation achieves the flatness of the universe somewhat ‘by design’.<sup>14</sup> We should therefore search for additional tests of the inflationary idea.

## 19.2 Coherent Phases and Superhorizon Fluctuations

As we have repeatedly emphasized in these lectures, the observations of the inhomogeneous universe allow detailed test of the inflationary dynamics. In this subsection, we discuss non-trivial qualitative features of the observations that inflation explains naturally, before giving quantitative results in the next subsection.

The following is a trivialization of arguments that have been explained beautifully by Dodelson in [25].

### 19.2.1 The Peaks of the TT Spectrum

Inflation produces a nearly scale-invariant spectrum of perturbations, *i.e.* a particular Fourier mode is drawn from a distribution with zero mean and variance given by

$$\langle \mathcal{R}_{\mathbf{k}} \mathcal{R}_{\mathbf{k}'} \rangle = (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}') P_{\mathcal{R}}(k) \quad (285)$$

where  $k^3 P_{\mathcal{R}}(k) \propto k^{n_s - 1}$  with  $n_s \approx 1$ .

---

<sup>14</sup>However, it is worth pointing out that when Guth introduced inflation in 1980, the flatness of the universe was a non-trivial prediction that at the time was inconsistent with observations!

You might think then that the shape of the power spectrum can be measured in observations, and this is what convinces us that inflation is right. Well, it is true that we can measure the power spectrum, both of the matter and of the radiation, and it is true that the observations agree with the theory. But this is not what tingles our spines when we look at the data. Rather, the truly striking aspect of perturbations generated during inflation is that *all Fourier modes have the same phase* [25].

Consider a Fourier mode with physical wavelength  $\lambda$ . While the mode is inside the horizon during inflation it oscillates with a frequency given by  $1/\lambda$ . However, before inflation ends, the mode exits the horizon, *i.e.* its physical wavelength gets stretched to a length greater than the instantaneous Hubble radius,  $\lambda > H^{-1}$ . After that its amplitude remains constant. Only at a much later time when the mode re-enters the horizon can causal physics affect it and lead to a time-evolution. Since the fluctuation amplitude was constant outside the horizon,  $\dot{\mathcal{R}}$  is very small at horizon re-entry. If we think of each Fourier mode as a linear combination of a sin and a cos mode, inflation excited only the cos modes (defining horizon re-entry as  $t \equiv 0$ ).

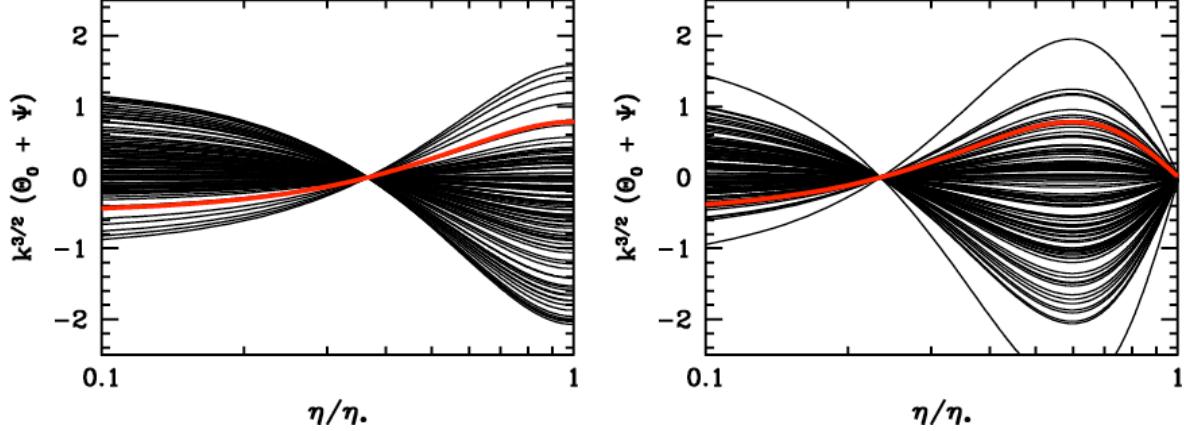
Once inside the horizon the curvature perturbation  $\mathcal{R}$  sources density fluctuations  $\delta$  which evolve under the influence of gravity and pressure

$$\ddot{\delta} - c_s^2 \nabla^2 \delta = F_g[\mathcal{R}], \quad (286)$$

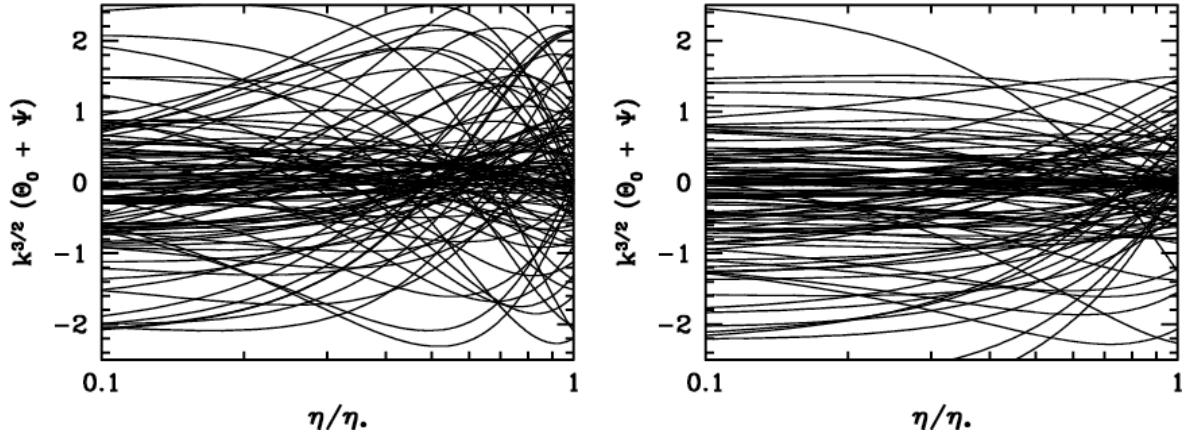
where  $c_s$  is the sound speed and  $F_g$  is the gravitational source term. This leads to oscillations in the density field. In the plasma of the early universe, fluctuations in the matter density were strongly coupled to fluctuations in the radiation. The CMB fluctuation therefore provide a direct snapshot of the conditions of the underlying density field at the time of recombination. Imagine that recombination happens instantaneously (not a terrible approximation). Fluctuations with different wavelengths would be captured at different phases in their oscillations. Modes of a certain wavelength would be captured at maximum or minimum amplitude, while others would be captured at zero amplitude. If all Fourier modes of a given wavelength had the same phases they would interfere coherently and the spectrum of all Fourier would produce a series of peaks and troughs in the CMB power spectrum as seen on the last-scattering surface. This is of course what we see in Fig. 15. However, in order for the theory of initial fluctuations to explain this it needs to involve a mechanism that produces coherent initial phases for all Fourier modes. Inflation does precisely that! Because fluctuations freeze when the exit the horizon the phases for the Fourier modes were set well before the modes of interest entered the horizon. When we are admiring the peak structure of the CMB power spectrum we are really admiring the ability of the primordial mechanism for generating fluctuations to coordinate the phases of all Fourier modes. Without this coherence, the CMB power spectrum would simply be white noise with no peaks and troughs (in fact, this is precisely why cosmic strings or topological defects are ruled out as the primary source for the primordial fluctuations.).

### 19.2.2 $\ell < 100$ in the TE Spectrum

The skeptic might not be convinced by the above argument. The peaks and troughs of the CMB temperature fluctuation spectrum are at  $\ell > 200$  corresponding to angular scales  $\theta < 1^\circ$ . All of these scales were within the horizon at the time of recombination. So it is in principle possible (and people have tried in the 90s) to engineer a theory of structure formation which obeys causality and still



**Figure 22:** Evolution of an infinite number of modes all with the same wavelength. Recombination is at  $\eta = \eta_*$ . (Left) Wavelength corresponding to the first peak in the CMB angular power spectrum. (Right) Wavelength corresponding to the first trough. Although the amplitudes of all these different modes differ from one another, since they start with the same phase, the ones on the left all reach maximum amplitude at recombination, the ones on the right all go to zero at recombination. This leads to the acoustic peaks of the CMB power spectrum.



**Figure 23:** Modes corresponding to the same two wavelengths as in Fig. 22, but this time with random initial phases. The anisotropies at the angular scales corresponding to these wavelengths would have identical rms's if the phases were random, *i.e.* the angular peak structure of the CMB would be washed away.

manages to produce only the ‘cosine mode’. Such a theory would explain the CMB peaks without appealing to inflation. This doesn’t sound like the most elegant thing in the world but it can’t be excluded as a logical possibility.

However, we now show that when considering CMB polarization, then even these highly-tuned alternatives to inflation can be ruled out. Looking at Fig. 19 we see that the cross-correlation between CMB temperature fluctuations and the *E*-mode polarization has a negative peak around  $100 < \ell < 200$ . This anti-correlation signal is also the result of phase coherence, but now the scales involved were *not* within the horizon at recombination. Hence, there is *no* causal mechanism that

could have produced this signal. One is almost forced to consider something like inflation with its shrinking comoving horizon leading to horizon exit and re-entry.<sup>15</sup>

As Dodelson explains [25]

At recombination, [the phase difference between the monopole (sourcing  $T$ ) and the dipole (sourcing  $E$ ) of the density field] causes the product of the two to be negative for  $100 < \ell < 200$  and positive on smaller scales until  $\ell \sim 400$ . But this is precisely what WMAP has observed! We have clear evidence that the monopole and the dipole were out of phase with each other at recombination.

This evidence is exciting for the small scale modes ( $\ell > 200$ ). Just as the acoustic peaks bear testimony to coherent phases, the cross-correlation of polarization and temperatures speaks to the coherence of the dipole as well. It solidifies our picture of the plasma at recombination. The evidence from the larger scale modes ( $\ell < 200$ ) though is positively stupendous. For, these modes were not within the horizon at recombination. So the *only* way they could have their phases aligned is if some primordial mechanism did the job, when they were in causal contact. Inflation is just such a mechanism.

### 19.3 Scale-Invariant, Gaussian and Adiabatic

We now describe quantitative constraints on the primordial fluctuations. The simplest versions of inflation predict that the scalar perturbations are nearly *scale-invariant*, *Gaussian* and *adiabatic*. In this section we give the latest quantitative constraints on these fundamental predictions of the theory.

#### 19.3.1 Spectral Index

As we explained in detail above, observations of the CMB relate to the inflationary spectrum of curvature perturbations  $\mathcal{R}$

$$\boxed{C_\ell^{TT}, C_\ell^{TE}, C_\ell^{EE} \Rightarrow P_{\mathcal{R}}(k)} \quad (287)$$

Here, we present the latest quantitative constraints on  $P_{\mathcal{R}}(k)$  in the standard power-law parameterization

$$\Delta_s^2(k) \equiv \frac{k^3}{2\pi^2} P_{\mathcal{R}}(k) = A_s \left(\frac{k}{k_*}\right)^{n_s - 1}. \quad (288)$$

Measurements of  $n_s$  are degenerate with the tensor-to-scalar ratio  $r$  so constraints on  $n_s$  are often shown as confidence contours in the  $n_s$ - $r$  plane. The latest WMAP 5-year constraints on the scalar spectral index are shown in Fig. 24 and Table 3. Two facts may be noted: i) the spectrum is nearly scale-invariant,  $n_s \approx 1$ , just as inflation predicts and ii) there are already interesting indications

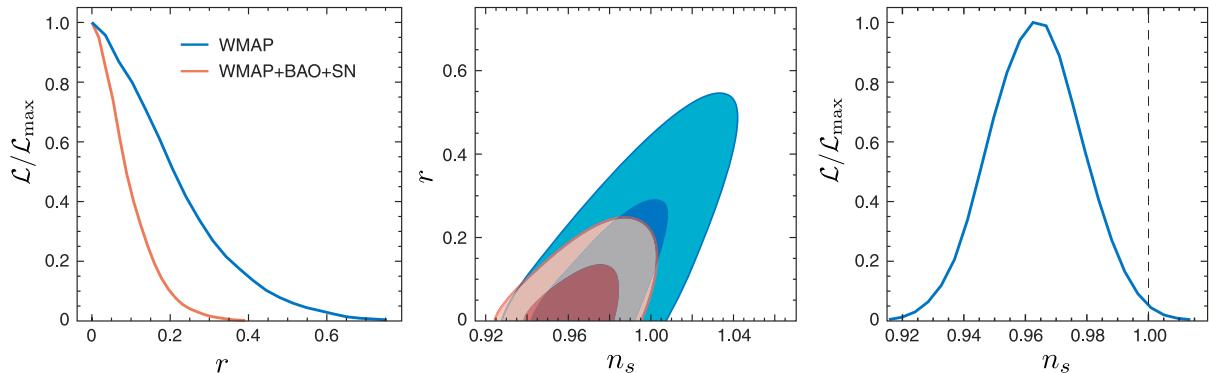
---

<sup>15</sup>It should be mentioned here that there are two ways to get a shrinking comoving Hubble radius,  $1/(aH)$ . During inflation  $H$  is nearly constant and the scale factor  $a$  grows exponentially. However, in a *contracting* spacetime a shrinking horizon can be achieved if  $H$  grows with time. This is the mechanism employed by ekpyrotic/cyclic cosmology. When viewed in terms of the evolution of the comoving Hubble scale inflation and ekpyrosis appear very similar, but there are important differences, *e.g.* in ekpyrosis it is a challenge to match the contracting phase to our conventional big bang expansion.

Parameter	5-year WMAP	WMAP+BAO+SN
$n_s$	$0.963^{+0.014}_{-0.015}$	$0.960^{+0.013}_{-0.013}$
$n_s$	$0.986 \pm 0.022$	$0.970 \pm 0.015$
$r$	$< 0.43$	$< 0.22$
$n_s$	$1.031^{+0.054}_{-0.055}$	$1.017^{+0.042}_{-0.043}$
$\alpha_s$	$-0.037 \pm 0.028$	$-0.028^{+0.020}_{-0.020}$
$n_s$	$1.087^{+0.072}_{-0.073}$	$1.089^{+0.070}_{-0.068}$
$r$	$< 0.58$	$< 0.55$
$\alpha_s$	$-0.050 \pm 0.034$	$-0.058 \pm 0.028$

**Table 3:** 5-year WMAP constraints on the primordial power spectra in the power law parameterization [23]. We present results for  $(n_s)$ ,  $(n_s, r)$ ,  $(n_s, \alpha_s)$  and  $(n_s, r, \alpha)$  marginalized over all other parameters of a flat  $\Lambda$ CDM model.

that the spectrum is not perfectly scale-invariant, but slightly red,  $n_s < 1$ . This deviation from scale-invariance provides the first test of the detailed time-dependence of the inflationary expansion. In fact, as we have seen in **Lecture 2**, inflation predicts this percent level deviation from scale-invariance.<sup>16</sup>



**Figure 24:** WMAP 5-year constraints on the inflationary parameters  $n_s$  and  $r$  [23]. The WMAP-only results are shown in blue, while constraints from WMAP plus other cosmological observations are in red. The third plot assumes that  $r$  is negligible.

### 19.3.2 Gaussianity

If  $\mathcal{R}$  is Gaussian then the power spectrum  $P_{\mathcal{R}}(k)$  (two-point correlations in real space) is the end of the story. However, if  $\mathcal{R}$  is *non-Gaussian* then the fluctuations have a non-zero bispectrum  $B_{\mathcal{R}}(k_1, k_2, k_3)$  (corresponding to three-point correlations in real space). There is only one way to be

<sup>16</sup>For inflation to end, the Hubble parameter  $H$  has to change in time. This time-dependence changes the conditions at the time when each fluctuation mode exits the horizon and therefore gets translated into a scale-dependence of the fluctuations.

Gaussian but many ways to be non-Gaussian, so constraints on non-Gaussianity are a bit hard to describe. One of the simplest forms of non-Gaussianity is described by the parameterization

$$\mathcal{R}(\mathbf{x}) = \mathcal{R}_g(\mathbf{x}) + \frac{3}{5} f_{\text{NL}} \star \mathcal{R}_g^2(\mathbf{x}). \quad (289)$$

where  $\mathcal{R}_g$  is Gaussian. In this *local model* for non-Gaussianity, the information is reduced to a single number  $f_{\text{NL}}$ . The latest constraint on  $f_{\text{NL}}$  by Smith, Senatore, and Zaldarriaga [28] is

$$-4 < f_{\text{NL}} < 80 \quad \text{at } 95\% \text{ CL}. \quad (290)$$

Notice that an  $f_{\text{NL}}$  value of order 100 corresponds to a 0.1% correction to  $\mathcal{R}_g \sim 10^{-5}$  in Eq. (289). The constraint (290) therefore implies that the CMB is Gaussian to 0.1%! This is better than our constraint on the curvature of the universe which is usually celebrated as the triumph of inflation. The CMB is highly Gaussian and it didn't have to be that way. However, if inflation is correct then the observed Gaussianity is a rather natural consequence.<sup>17</sup>

### 19.3.3 Adiabaticity

In single-field inflation, the fluctuations of the inflaton field on large scales (where spatial gradients can be neglected) can be identified with a local shift backwards or forwards along the trajectory of the homogeneous background field. These shifts along the inflaton trajectory affect the total density in different parts of the universe after inflation, but cannot give rise to variations in the relative density between different components. Hence, single field inflation produces purely *adiabatic* primordial density perturbations characterized by an overall curvature perturbation,  $\mathcal{R}$ . This means that all perturbations of the cosmological fluid (photons, neutrinos, baryons and cold dark matter (CDM) particles) originate from the same curvature perturbation  $\mathcal{R}$  and satisfy the adiabaticity property,  $\delta(n_m/n_r) = 0$ , or

$$\frac{\delta\rho_m}{\rho_m} = \frac{3}{4} \frac{\delta\rho_r}{\rho_r}, \quad (291)$$

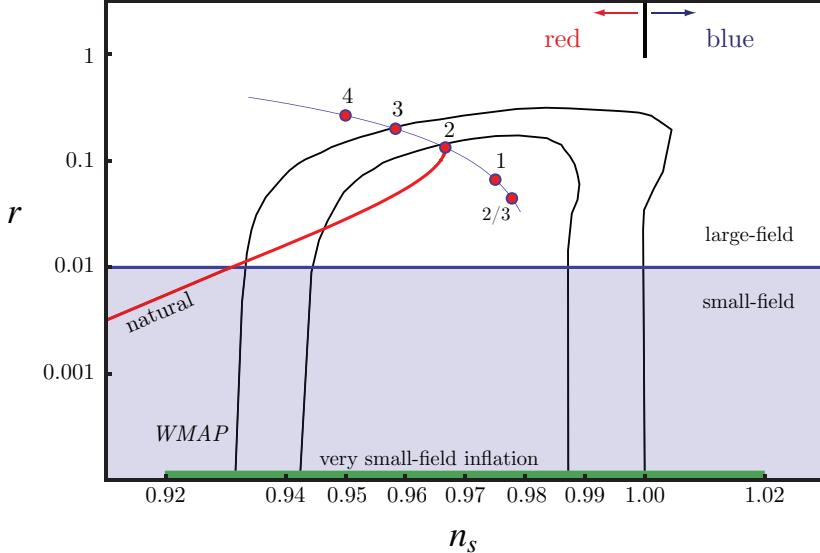
where the index  $m$  collectively stands for non-relativistic species (baryons or CDM) and  $r$  for relativistic species (photons or neutrinos). The latest data shows no violation of the condition (291) [23]. If such a violation were to be found this would be a clear signature of multi-field inflation (see §20.5).

## 19.4 Testing Slow-Roll Models

In Figure 25 we present current observational constraints on some of the simplest single-field slow-roll models of inflation (see **Lecture 1**). Future measurements will significantly cut into the parameter space of allowed models.

---

<sup>17</sup>Non-Gaussianity is a measure of *interactions* of the inflaton field. However, for slow-roll dynamics to occur, the inflaton has to be very weakly self-interacting (the potential is very flat) and the non-Gaussianity is necessarily small,  $f_{\text{NL}} \sim \mathcal{O}(0.01)$ .



**Figure 25:** Constraints on single-field slow-roll models in the  $n_s$ - $r$  plane. The value of  $r$  determines whether the models involve large or small field variations. The value of  $n_s$  classifies the scalar spectrum as red or blue. Combinations of the values of  $r$  and  $n_s$  determine whether the curvature of the potential was positive ( $\eta > 0$ ) or negative ( $\eta < 0$ ) when the observable universe exited the horizon. Also shown are the WMAP 5-year constraints on  $n_s$  and  $r$  [23] as well as the predictions of a few representative models of single-field slow-roll inflation: *chaotic inflation*:  $\lambda_p \phi^p$ , for general  $p$  (thin solid line) and for  $p = 4, 3, 2, 1, \frac{2}{3}$  (●); models with  $p = 2$  [29],  $p = 1$  [30] and  $p = \frac{2}{3}$  [31] have recently been obtained in string theory; *natural inflation*:  $V_0[1 - \cos(\phi/\mu)]$  (solid line); *very small-field inflation*: models of inflation with a very small tensor amplitude,  $r \ll 10^{-4}$  (green bar); examples of such models in string theory include warped D-brane inflation [32–34], Kähler inflation [35], and racetrack inflation [36].

## 20 Future Tests of Inflation

We are only at the beginning of really testing the inflationary paradigm. The flatness of the universe and the near scale-invariance, Gaussianity and adiabaticity of the density fluctuations are encouraging evidence for inflation<sup>18</sup>, but they are not proof that inflation really occurred.<sup>19</sup> Let us therefore look into the future and describe how future experiments can provide further tests of inflationary physics.

### 20.1 Amplitude of Tensor Modes

Probably the single most important piece of evidence for inflation would come from a measurement of a primordial tensor amplitude. We showed above that a detection of primordial CMB  $B$ -modes

<sup>18</sup>Note that at any stage we could have made measurements that would have falsified the whole idea of inflation.

<sup>19</sup>I know, I used the word ‘proof’ when Karl Popper taught us that we can never prove a theory.

would be virtually impossible to explain by anything other inflationary gravitational waves [37]

$$\boxed{C_\ell^{BB} \Rightarrow P_h(k)} \quad (292)$$

where

$$\Delta_t^2(k) \equiv \frac{k^3}{2\pi^2} P_h(k) = A_t \left( \frac{k}{k_*} \right)^{n_t}. \quad (293)$$

We also explained in **Lecture 2** that the tensor amplitude  $A_t$  is directly linked with the energy scale of inflation. As a single clue about the physics of inflation, what could be more important and higher on the wish-list of inflationary theorists? In addition, a detection of tensor modes would imply that the inflaton field moved over a super-Planckian distance in field space, making string theorists and quantum gravity affectionatos think hard about Planck-suppressed corrections to the inflaton potential (see **Lecture 5**).

## 20.2 Scale Dependence of Scalar Modes

The variation of the spectral index  $n_s$  with scale (also called the “running” of the spectral index) arises only at second-order in slow-roll and is therefore expected to be small

$$\alpha_s \equiv \frac{dn_s}{d \ln k} \sim \mathcal{O}(\epsilon^2). \quad (294)$$

On the other hand, an unexpected large positive or negative running would force us to rethink some of our simplest notions about inflation and the generation of perturbations.

## 20.3 Scale Dependence of Tensor Modes

Measuring the amplitude of primordial tensor fluctuations from inflation will be a significant observational challenge. Hoping to measure its dependence on scale seems unrealistic unless the tensor amplitude is near its current upper limit. In single-field slow-roll models the tensor-to-scalar ratio  $r$  and the tensor spectral index  $n_t$  are related by the *consistency relation*

$$r = -8n_t. \quad (295)$$

Measuring (295) would offer another way to falsify single-field slow-roll inflation.

## 20.4 Non-Gaussianity

The primordial fluctuations are to a high degree Gaussian. However, as we now describe, even a small non-Gaussianity would encode a tremendous amount of information about the inflationary action. We mentioned that the three-point function of inflationary fluctuations is the prime diagnostic of non-Gaussian statistics. In momentum space, the three-point correlation function can be written generically as

$$\langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k}_2} \mathcal{R}_{\mathbf{k}_3} \rangle = (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) f_{NL} F(k_1, k_2, k_3). \quad (296)$$

Here,  $f_{NL}$  is a dimensionless parameter defining the amplitude of non-Gaussianity, while the function  $F(k_1, k_2, k_3)$  captures the momentum dependence. The amplitude and sign of  $f_{NL}$ , as well as the

shape and scale dependence of  $F(k_1, k_2, k_3)$ , depend on the details of the interaction generating the non-Gaussianity, making the three-point function a powerful discriminating tool for probing models of the early universe [19].

Two simple and distinct shapes  $F(k_1, k_2, k_3)$  are generated by two very different mechanisms [38]: The *local shape* is a characteristic of multi-field models and takes its name from the expression for the primordial curvature perturbation  $\mathcal{R}$  in real space,

$$\mathcal{R}(\mathbf{x}) = \mathcal{R}_g(\mathbf{x}) + \frac{3}{5} f_{\text{NL}}^{\text{local}} \mathcal{R}_g(\mathbf{x})^2, \quad (297)$$

where  $\mathcal{R}_g(\mathbf{x})$  is a Gaussian random field. Fourier transforming this expression shows that the signal is concentrated in “squeezed” triangles where  $k_1 \ll k_2, k_3$ . Local non-Gaussianity arises in multi-field models where the fluctuations of an isocurvature field (see below) are converted into curvature perturbations. As this conversion happens outside of the horizon, when gradients are irrelevant, one generates non-linearities of the form (297). Specific models of this type include multi-field inflation [39–51], the curvaton scenario [52, 53], inhomogeneous reheating [54, 55], and New Ekpyrotic models [56–62].

The second important shape is called *equilateral* as it is largest for configurations with  $k_1 \sim k_2 \sim k_3$ . The equilateral form is generated by single-field models with non-canonical kinetic terms such as DBI inflation [63], ghost inflation [64, 65] and more general models with small sound speed [66, 67].

## 20.5 Isocurvature Fluctuations

In inflationary models with more than one field the perturbations are not necessarily adiabatic. With more than one field, fluctuations orthogonal to the background trajectory can affect the relative density between different matter components even if the total density and therefore the spatial curvature is unperturbed [68]. There are various different possibilities for such *isocurvature* perturbations (also called non-adiabatic or entropic perturbations), *e.g.* we may define relative perturbations between CDM and photons

$$\mathcal{S}_m \equiv \frac{\delta\rho_m}{\rho_m} - \frac{3}{4} \frac{\delta\rho_\gamma}{\rho_\gamma}. \quad (298)$$

Adiabatic and isocurvature perturbations lead to a different peak structure in the CMB fluctuations. CMB measurements can therefore distinguish between the different types of fluctuations and in fact already show that isocurvature perturbations have to be a subdominant component (if at all present).

Isocurvature perturbations could be correlated with the adiabatic perturbations. To capture this we define the following correlation parameter

$$\beta \equiv \frac{P_{S\mathcal{R}}}{\sqrt{P_S P_{\mathcal{R}}}}, \quad (299)$$

where  $P_{\mathcal{R}}$  and  $P_S$  are the power spectra of adiabatic and isocurvature fluctuations and  $P_{S\mathcal{R}}$  is their cross-correlation. Parameterizing the relative amplitude between the two types of perturbations by a coefficient  $\alpha$

$$\frac{P_S}{P_{\mathcal{R}}} \equiv \frac{\alpha}{1 - \alpha}, \quad (300)$$

the present constraints on the isocurvature contribution are  $\alpha_0 < 0.067$  (96% CL) in the uncorrelated case ( $\beta = 0$ ) and  $\alpha_{-1} < 0.0037$  (95% CL) in the totally anti-correlated case ( $\beta = -1$ ).

Theoretical predictions for the amplitude of isocurvature perturbations are complicated by the fact that they are strongly model-dependent: it does not depend entirely on the multi-field inflationary dynamics, but also on the post-inflationary evolution. If all particle species are in thermal equilibrium after inflation and their local densities are uniquely given by their temperature (with vanishing chemical potential) then the primordial perturbations are adiabatic [69, 70]. Thus, it is important to note that the existence of primordial isocurvature modes requires at least one field to decay into some species whose abundance is not determined by thermal equilibrium (*e.g.* CDM after decoupling) or respects some conserved quantum numbers, like baryon or lepton numbers.

## 21 Summary: Lecture 3

Observations of the cosmic microwave background (CMB) and the large-scale structure (LSS) may be used to constrain the spectrum of primordial seed fluctuations. This makes CMB and LSS experiments probes of the early universe. To extract this information about the inflationary era the late time evolution of fluctuations has to be accounted for. This is done with numerical codes such as **CMBFAST** and **CAMB**.

Current observations are in beautiful agreement with the basic inflationary predictions: The universe is flat with a spectrum of nearly scale-invariant, Gaussian and adiabatic density fluctuations. The fluctuations show non-zero correlations on scales that were bigger than the horizon at recombination. Furthermore, the peak structure of the CMB spectrum is evidence that the fluctuations were created with coherent phases.

Future tests of inflation will mainly come from measurements of CMB polarization.  $B$ -modes of CMB polarization are a unique signature of inflationary gravitational waves. The  $B$ -mode amplitude is a direct measure of the energy scale of inflation. In addition, non-Gaussianity measurements potentially carry a wealth of information about the physics of inflation by constraining interactions of the inflaton field.

Finally, the following measurements would falsify single-field slow-roll inflation:

- Large non-Gaussianity,  $f_{\text{NL}} > 1$ .
- Non-zero isocurvature perturbations,  $\alpha \neq 0$ .
- Large running of the scalar spectrum,  $|\alpha_s| > 0.001$ .
- Violation of the tensor consistency relation,  $r \neq -8n_t$ .

## Part V

# Lecture 4: Primordial Non-Gaussianity

---

### Abstract

In this lecture we summarize key theoretical results in the study of primordial non-Gaussianity. Most results are stated without proof, but their significance for constraining the fundamental physical origin of inflation is explained. After introducing the bispectrum as a basic diagnostic of non-Gaussian statistics, we show that its momentum dependence is a powerful probe of the inflationary action. Large non-Gaussianity can only arise if inflaton interactions are significant during inflation. In single-field slow-roll inflation non-Gaussianity is therefore predicted to be unobservably small, while it can be significant in models with multiple fields, higher-derivative interactions or non-standard initial states. Finally, we end the lecture with a discussion of the observational prospects for detecting or constraining primordial non-Gaussianity.

---

## 22 Preliminaries

In [Lecture 2](#) we expanded the inflaton action to second order in the comoving curvature perturbation  $\mathcal{R}$ . This free field action allowed us to compute the power spectrum  $P_{\mathcal{R}}(k)$ . As we mentioned in [Lecture 3](#), if the fluctuations  $\mathcal{R}$  are drawn from a Gaussian distribution, then the power spectrum (or two-point correlation function) contains all the information.<sup>20</sup> In this lecture, we discuss what further information about inflation can be obtained by studying higher-order correlation functions beyond the two-point function.

### 22.1 The Bispectrum and Local Non-Gaussianity

The Fourier transform of the two-point function is the power spectrum

$$\langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k}_2} \rangle = (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2) P_{\mathcal{R}}(k_1). \quad (301)$$

The Fourier transform of the three-point function is the bispectrum

$$\boxed{\langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k}_2} \mathcal{R}_{\mathbf{k}_3} \rangle = (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) B_{\mathcal{R}}(k_1, k_2, k_3)}. \quad (302)$$

One of the first ways to parameterize non-Gaussianity was via a non-linear correction to a Gaussian perturbation  $\mathcal{R}_g$ ,

$$\boxed{\mathcal{R}(\mathbf{x}) = \mathcal{R}_g(\mathbf{x}) + \frac{3}{5} f_{\text{NL}}^{\text{local}} [\mathcal{R}_g(\mathbf{x})]^2}. \quad (303)$$

---

<sup>20</sup>The tree-point function and all odd higher-point correlation functions vanish for Gaussian fluctuations, while all even higher-point functions can be expressed in terms of the two-point function.

This definition is local in real space and therefore called *local non-Gaussianity*. Experimental constraints on non-Gaussianity (see **Lecture 3**) are often set on the parameter  $f_{\text{NL}}^{\text{local}}$  defined via Eqn. (303).<sup>21</sup> Using Eqn. (303) the bispectrum of local non-Gaussianity may be derived

$$B_{\mathcal{R}}(k_1, k_2, k_3) = \frac{6}{5} f_{\text{NL}}^{\text{local}} \times [P_{\mathcal{R}}(k_1)P_{\mathcal{R}}(k_2) + P_{\mathcal{R}}(k_2)P_{\mathcal{R}}(k_3) + P_{\mathcal{R}}(k_3)P_{\mathcal{R}}(k_1)] . \quad (304)$$

For a scale-invariant spectrum,  $P_{\mathcal{R}}(k) = Ak^{-3}$ , this is

$$B_{\mathcal{R}}(k_1, k_2, k_3) = \frac{6}{5} f_{\text{NL}}^{\text{local}} \times A^2 \left[ \frac{1}{(k_1 k_2)^3} + \frac{1}{(k_2 k_3)^3} + \frac{1}{(k_3 k_1)^3} \right] . \quad (305)$$

Without loss of generality, let us order the momenta such that  $k_3 \leq k_2 \leq k_1$ . The bispectrum for local non-Gaussianity is then largest when the smallest  $k$  (*i.e.*  $k_3$ ) is very small,  $k_3 \ll k_1 \sim k_2$ . In this *squeezed* limit, the bispectrum for local non-Gaussianity becomes

$$\lim_{k_3 \ll k_1 \sim k_2} B_{\mathcal{R}}(k_1, k_2, k_3) = \frac{12}{5} f_{\text{NL}}^{\text{local}} \times P_{\mathcal{R}}(k_1)P_{\mathcal{R}}(k_3) . \quad (306)$$

## 22.2 Shapes of Non-Gaussianity

The delta function in Eqn. (302) enforces that the three Fourier modes of the bispectrum form a closed triangle. Different inflationary models predict maximal signal for different triangle configurations. This *shape of non-Gaussianity* is potentially a powerful probe of the mechanism that laid down the primordial perturbations.

It will be convenient to define the shape function

$$\mathcal{S}(k_1, k_2, k_3) \equiv N(k_1 k_2 k_3)^2 B_{\mathcal{R}}(k_1, k_2, k_3) , \quad (307)$$

where  $N$  is an appropriate normalization factor. Two commonly discussed shapes are the local model,

$$\mathcal{S}^{\text{local}}(k_1, k_2, k_3) \propto \frac{K_3}{K_{111}} , \quad (308)$$

and the equilateral model,

$$\mathcal{S}^{\text{equil}}(k_1, k_2, k_3) \propto \frac{\tilde{k}_1 \tilde{k}_2 \tilde{k}_3}{K_{111}} . \quad (309)$$

Here, we have introduced a notation first defined by Fergusson and Shellard [71],

$$K_p = \sum_i (k_i)^p \quad \text{with} \quad K = K_1 \quad (310)$$

$$K_{pq} = \frac{1}{\Delta_{pq}} \sum_{i \neq j} (k_i)^p (k_j)^q \quad (311)$$

$$K_{pqr} = \frac{1}{\Delta_{pqr}} \sum_{i \neq j \neq l} (k_i)^p (k_j)^q (k_l)^r \quad (312)$$

$$\tilde{k}_{ip} = K_p - 2(k_i)^p \quad \text{with} \quad \tilde{k}_i = \tilde{k}_{i1} , \quad (313)$$

---

<sup>21</sup>The factor of  $3/5$  Eqn. (303) is conventional since non-Gaussianity was first defined in terms of the Newtonian potential,  $\Phi(\mathbf{x}) = \Phi_g(\mathbf{x}) + f_{\text{NL}}[\Phi_g(\mathbf{x})]^2$ , which during the matter era is related to  $\mathcal{R}$  by a factor of  $3/5$ .

where  $\Delta_{pq} = 1 + \delta_{pq}$  and  $\Delta_{pqr} = \Delta_{pq}(\Delta_{qr} + \delta_{pr})$  (no summation). This notation significantly compresses the increasingly complex expressions for the bispectra discussed in the literature. Physically motivated models for producing non-Gaussian perturbations often produce signals that peak at special triangle configurations. Three special cases are:

- i) *squeezed triangle* ( $k_1 \approx k_2 \gg k_3$ )

This is dominant mode of models with multiple light fields during inflation.

- ii) *equilateral triangle* ( $k_1 = k_2 = k_3$ )

Signals that peak at equilateral triangles arise in models with higher-derivative interactions.

- iii) *folded triangle* ( $k_1 = 2k_2 = 2k_3$ )

Folded triangles arise in models with non-standard initial states.

In addition, there are the intermediate cases: *elongated triangle*  $k_1 = k_2 + k_3$ , *isosceles triangle*  $k_1 > k_2 = k_3$ .

## 22.3 $f_{\text{NL}}$ : The Amplitude of Non-Gaussianity

For arbitrary shape functions we measure the magnitude of non-Gaussianity by defining the generalized  $f_{\text{NL}}$  parameter

$$f_{\text{NL}} = \boxed{\frac{5}{18} \frac{B_{\mathcal{R}}(k, k, k)}{P_{\mathcal{R}}(k)^2}}. \quad (314)$$

# 23 Theoretical Expectations

## 23.1 Single-Field Slow-Roll Inflation

Successful slow-roll inflation demands that the interactions of the inflaton field are weak. Since the wave function of free fields in the ground state is Gaussian, the fluctuations created during slow-roll inflation are expected to be Gaussian. Maldacena [16] first derived the bispectrum for slow-roll inflation

$$\mathcal{S}^{\text{SR}}(k_1, k_2, k_3) \propto (\varepsilon - 2\eta) \frac{K_3}{K_{111}} + \varepsilon \left( K_{12} + 8 \frac{K_{22}}{K} \right) \quad (315)$$

$$\approx (4\varepsilon - 2\eta) \mathcal{S}^{\text{local}}(k_1, k_2, k_3) + \frac{5}{3} \varepsilon \mathcal{S}^{\text{equil}}(k_1, k_2, k_3), \quad (316)$$

where  $\mathcal{S}^{\text{local}}$  and  $\mathcal{S}^{\text{equil}}$  are normalized so that  $\mathcal{S}^{\text{local}}(k, k, k) = \mathcal{S}^{\text{equil}}(k, k, k)$ . It peaks at squeezed triangles and has an amplitude that is suppressed by slow-roll parameters. This makes intuitive sense since the slow-roll parameters characterize deviations of the inflaton from a free field.

## 23.2 The Maldacena Theorem

Under the assumption of single-field inflation, but *no* other assumption about the inflationary action Creminelli and Zaldarriaga [72] were able to prove a powerful theorem first considered for the special

case of slow-roll inflation by Maldacena:

$$\lim_{k_1 \rightarrow 0} \langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k}_2} \mathcal{R}_{\mathbf{k}_3} \rangle = (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \underline{(1 - n_s)} P_{\mathcal{R}}(k_1) P_{\mathcal{R}}(k_3) , \quad (317)$$

where

$$\langle \mathcal{R}_{\mathbf{k}_i} \mathcal{R}_{\mathbf{k}_j} \rangle = (2\pi)^3 \delta(\mathbf{k}_i + \mathbf{k}_j) P_{\mathcal{R}}(k_i) . \quad (318)$$

Eqn. (317) says that for single-field inflation, the squeezed limit of the three-point function is suppressed by  $(1 - n_s)$  and vanishes for perfectly scale-invariant perturbations.

### Proof:

The squeezed triangle correlates one long-wavelength mode,  $k_L = k_3$  to two short-wavelength modes,  $k_S = k_1 \approx k_2$ ,

$$\langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k}_2} \mathcal{R}_{\mathbf{k}_3} \rangle \approx \langle (\mathcal{R}_{\mathbf{k}_S})^2 \mathcal{R}_{\mathbf{k}_L} \rangle \quad (319)$$

Modes with longer wavelength freeze earlier. Therefore,  $k_L$  will be already frozen outside the horizon when the two smaller modes freeze and act as a background field for the two short-wavelength modes.

Why should  $(\mathcal{R}_{\mathbf{k}_S})^2$  be correlated with  $\mathcal{R}_{\mathbf{k}_L}$ ? The theorem says that “it isn’t if  $\mathcal{R}_{\mathbf{k}}$  is precisely scale-invariant”. The proof is simplest in real-space (see Creminelli and Zaldarriaga [72]): The long-wavelength curvature perturbation  $\mathcal{R}_{\mathbf{k}_L}$  rescales the spatial coordinates (or changes the effective scale factor) within a given Hubble patch

$$ds^2 = -dt^2 + a(t)^2 e^{2\mathcal{R}} d\mathbf{x}^2 . \quad (320)$$

The two-point function  $\langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k}_2} \rangle$  will depend on the value of the background fluctuations  $\mathcal{R}_{\mathbf{k}_L}$  already frozen outside the horizon. In position space the variation of the two-point function given by the long-wavelength fluctuations  $\mathcal{R}_{\mathbf{k}_L}$  is at linear order

$$\frac{\partial}{\partial \mathcal{R}_L} \langle \mathcal{R}(x) \mathcal{R}(0) \rangle \cdot \mathcal{R}_L = x \frac{d}{dx} \langle \mathcal{R}(x) \mathcal{R}(0) \rangle \cdot \mathcal{R}_L \quad (321)$$

To get the three-point function Creminelli and Zaldarriaga multiply (321) by  $\mathcal{R}_L$  and average over it. Going to Fourier space gives (317).<sup>22</sup> QED.

A detection of non-Gaussianity in the squeezed limit can therefore rule out single-field inflation! In particular, this statement is independent of: the form of the potential, the form of the kinetic term (or sound speed) and the initial vacuum state.

## 23.3 Models with Large Non-Gaussianity

### 23.3.1 Higher-Derivative Interactions

Although Maldacena proved that for single-field *slow-roll* inflation non-Gaussianity is always small, single-field models can still give large non-Gaussianity if higher-derivative terms are important during

---

<sup>22</sup>For more details see Cheung et al. [73].

inflation (as opposed to assuming a canonical kinetic term and no higher-derivative corrections as in slow-roll inflation). Consider the following action

$$S = \frac{1}{2} \int d^4x \sqrt{-g} [R - P(X, \phi)] , \quad \text{where } X \equiv (\partial_\mu \phi)^2 . \quad (322)$$

Here,  $P(X, \phi)$  is an arbitrary function of the kinetic term  $X = (\partial_\mu \phi)^2$ . These models in general have a non-trivial sound speed for the propagation of fluctuations

$$c_s^2 = \frac{P_{,X}}{P_{,X} + 2XP_{,XX}} . \quad (323)$$

The second-order action for  $\mathcal{R}$  (giving  $P_{\mathcal{R}}$ ) is

$$S_{(2)} = \int d^4x \varepsilon \left[ a^3 (\dot{\mathcal{R}})^2 / \underline{\underline{c}_s^2} - a(\partial_i \mathcal{R})^2 \right] + \mathcal{O}(\varepsilon^2) \quad (324)$$

The third-order action for  $\mathcal{R}$  (giving  $B_{\mathcal{R}}$ ) is

$$S_{(3)} = \int d^4x \underline{\underline{\varepsilon}}^2 \left[ \dots a^3 (\dot{\mathcal{R}})^2 \mathcal{R} / \underline{\underline{c}_s^2} + \dots a(\partial_i \mathcal{R})^2 \mathcal{R} + \dots a^3 (\dot{\mathcal{R}})^3 / \underline{\underline{c}_s^2} \right] + \mathcal{O}(\varepsilon^3) \quad (325)$$

We notice that the third order action is suppressed by an extra factor of  $\varepsilon$  relative to the second order action. This is a reflection of the fact that non-Gaussianity is small in the slow-roll limit:  $P(X, \phi) = X - V(\phi)$ ,  $c_s^2 = 1$ . However, away from the slow-roll limit, for small sound speeds,  $c_s^2 \ll 1$ , a few interaction terms in (325) get boosted and non-Gaussianity can become significant. The signal is peaked at equilateral triangles, with

$$f_{\text{NL}}^{\text{equil}} = -\frac{35}{108} \left( \frac{1}{c_s^2} - 1 \right) + \frac{5}{81} \left( \frac{1}{c_s^2} - 1 - 2\Lambda \right) \quad (326)$$

where

$$\Lambda \equiv \frac{X^2 P_{,XX} + \frac{2}{3} X^3 P_{,XXX}}{XP_{,X} + 2X^2 P_{,XX}} . \quad (327)$$

Whether actions with arbitrary  $P(X, \phi)$  exist in consistent high-energy theories is an important challenge for these models. It is encouraging that one of the most interaction models of inflation in string theory, DBI inflation (see **Lecture 5**), has precisely such a structure with

$$P_{\text{DBI}}(X, \phi) = -f^{-1}(\phi) \sqrt{1 - 2f(\phi)X} + f^{-1}(\phi) - V(\phi) . \quad (328)$$

In this case, the second term in (329) is identically zero and we find

$$f_{\text{NL}}^{\text{DBI}} = -\frac{35}{108} \left( \frac{1}{c_s^2} - 1 \right) . \quad (329)$$

The shape function for DBI inflation is

$$\mathcal{S}^{\text{DBI}}(k_1, k_2, k_3) \propto \frac{1}{K_{111} K^2} (K_5 + 2K_{14} - 3K_{23} + 2K_{113} - 8K_{122}) . \quad (330)$$

### 23.3.2 Multiple Fields

In single-field slow-roll inflation interactions of the inflaton are constrained by the requirement that inflation should occur. However, if more than one field was relevant during inflation this constraint may be circumvented. Models like the curvaton mechanism or inhomogeneous reheating exploit this to create non-Gaussian fluctuations via fluctuations in a second field that is not the inflaton. The signal is peaked at squeezed triangles.

### 23.3.3 Non-Standard Vacuum

If inflation started in an excited state rather than the Bunch-Davies vacuum, remnant non-Gaussianity may be observable [74] (unless inflation lasted much more than the minimal number of  $e$ -folds, in which case the effect is exponentially diluted). The signal is peaked at folded triangles with a shape function

$$\mathcal{S}^{\text{folded}}(k_1, k_2, k_3) \propto \frac{1}{K_{111}}(K_{12} - K_3) + 4 \frac{K_2}{(\tilde{k}_1 \tilde{k}_2 \tilde{k}_3)^2}. \quad (331)$$

## 24 Observational Prospects

### 24.1 Cosmic Microwave Background

The non-linear evolution of fluctuation can induce its own non-Gaussianity. The effects haven't been fully computed. Often only their order of magnitude is estimated.

### 24.2 Large-Scale Structure

Scale-dependent bias may become a promising probe of primordial non-Gaussianity.

## **25 Summary: Lecture 4**

## Part VI

# Lecture 5: Inflation in String Theory

---

### Abstract

We end this lecture series with a discussion of a slightly more advanced topic: inflation in string theory. We provide a pedagogical overview of the subject.<sup>23</sup> Our central theme is the sensitivity of inflation to Planck-scale physics, which we argue provides both the primary motivation and the central theoretical challenge for realizing inflation in string theory. We illustrate these issues through two case studies of inflationary scenarios in string theory: warped D-brane inflation and axion monodromy inflation. Finally, we indicate opportunities for future progress both theoretically and observationally.

---

## 26 Why Combine Two Speculative Ideas?

In the previous lectures we have seen that inflation is remarkably successful as a phenomenological model for the dynamics of the very early universe. However, a detailed understanding of the physical origin of the inflationary expansion has remained elusive. Inflation and string theory are both ambitious attempts to understand the physical universe at the highest energies. Both inflation and string theory are speculative theories that still await experimental confirmation. One may therefore wonder why it is timely to address the problem of inflation in string theory.

In this lecture we will highlight specific aspects of inflation that depend sensitively on the ultraviolet (UV) completion of quantum field theory and gravity, *i.e.* on the field content and interactions at energies approaching the Planck scale. Such issues are most naturally addressed in a theory of Planck-scale physics, for which string theory is arguably the best-developed candidate. This motivates understanding the physics of inflation in string theory.

## 27 UV Sensitivity of Inflation

### 27.1 Effective Field Theory and Inflation

As a phenomenon in quantum field theory coupled to general relativity, inflation does not appear to be natural. In particular, the set of Lagrangians suitable for inflation is a minute subset of the set of all possible Lagrangians. Moreover, in wide classes of models, inflation emerges only for rather special initial conditions, *e.g.* initial configurations with tiny kinetic energy, in the case of small-field scenarios. Although one would hope to explore and quantify the naturalness both of inflationary Lagrangians and of inflationary initial conditions, the question of initial conditions appears inextricable from the active yet incomplete program of understanding measures in eternal

---

<sup>23</sup>The presentation of this lecture is based on a recent review with Liam McAllister [11].

inflation (see §31 for a critical evaluation). In this lecture we will focus on the question of how (un)natural it is to have a Lagrangian suitable for inflation.

For a single inflaton field with a canonical kinetic term, the necessary conditions for inflation can be stated in terms of the inflaton potential (see **Lecture 1**). Inflation requires a potential that is quite flat in Planck units

$$\epsilon = M_{\text{pl}}^2 \left( \frac{V_{,\phi}}{V} \right)^2 \ll 1, \quad \eta = \frac{M_{\text{pl}}^2}{2} \frac{V_{,\phi\phi}}{V} \ll 1. \quad (332)$$

As we now argue, this condition is sensitive to Planck-scale physics.

Let us recall that the presence of some form of new physics at the Planck scale is required in order to render graviton-graviton scattering sensible, just as unitarity of  $W$ - $W$  scattering requires new physics at the TeV scale. Although we know that new degrees of freedom must emerge, we cannot say whether the physics of the Planck scale is a finite theory of quantum gravity, such as string theory, or is instead simply an effective theory for some unimagined physics at yet higher scales. However, the structure of the Planck-scale theory has meaningful – and, in very favorable cases, testable – consequences for the form of the inflaton potential.

As usual, the effects of high-scale physics above some cutoff  $\Lambda$  are efficiently described by the coefficients of operators in the low-energy effective theory. Integrating out particles of mass  $M \geq \Lambda$  gives rise to operators of the form

$$\frac{\mathcal{O}_\delta}{M^{\delta-4}}, \quad (333)$$

where  $\delta$  denotes the mass dimension of the operator.

Sensitivity to such operators is commonplace in particle physics: for example, bounds on flavor-changing processes place limits on physics above the TeV scale, and lower bounds on the proton lifetime even allow us to constrain GUT-scale operators that would mediate proton decay. However, particle physics considerations alone do not often reach beyond operators of dimension  $\delta = 6$ , nor go beyond  $M \sim M_{\text{GUT}}$ . (Scenarios of gravity-mediated supersymmetry breaking are one exception.) Equivalently, Planck-scale processes, and operators of very high dimension, are irrelevant for most of particle physics: they decouple from low-energy phenomena.

In inflation, however, the flatness of the potential in Planck units introduces sensitivity to  $\delta \leq 6$  *Planck-suppressed* operators, such as

$$\frac{\mathcal{O}_6}{M_{\text{pl}}^2}. \quad (334)$$

An understanding of such operators is required to address the smallness of the eta parameter, *i.e.* to ensure that the theory supports at least 60  $e$ -folds of inflationary expansion. This sensitivity to dimension-six Planck-suppressed operators is therefore common to all models of inflation.

For large-field models of inflation the UV sensitivity of the inflaton action is dramatically enhanced. In this important class of inflationary models the potential becomes sensitive to an *infinite* series of operators of arbitrary dimension (see §27.3).

## 27.2 The Eta Problem

In the absence of any specific symmetries protecting the inflaton potential, contributions to the Lagrangian of the general form

$$\frac{\mathcal{O}_6}{M_{\text{pl}}^2} = \frac{\mathcal{O}_4}{M_{\text{pl}}^2} \phi^2 \quad (335)$$

are allowed. If the dimension-four operator  $\mathcal{O}_4$  has a vacuum expectation value (vev) comparable to the inflationary energy density,  $\langle \mathcal{O}_4 \rangle \sim V$ , then this term corrects the inflaton mass by order  $H$ , or equivalently corrects the eta parameter by order one, leading to an important problem for inflationary model-building. Let us reiterate that contributions of this form may be thought of as arising from integrating out Planck-scale degrees of freedom. In this section we discuss this so-called *eta problem* first in effective field theory, §27.2.1, and then illustrate the problem in a supergravity example, §27.2.2.

### 27.2.1 Radiative Instability of the Inflaton Mass

In a generic effective theory with cutoff  $\Lambda$ , the mass of a scalar field runs to the cutoff scale unless it is protected by some symmetry. Since the cutoff for an effective theory of inflation is at least the Hubble scale,  $\Lambda \geq H$ , this implies that a small inflaton mass ( $m_\phi \ll H$ ) is radiatively unstable. Equivalently, the eta parameter receives radiative corrections,

$$\Delta\eta = \frac{\Delta m_\phi^2}{3H^2} \geq 1, \quad (336)$$

preventing prolonged inflation.

The difficulty here is analogous to the Higgs hierarchy problem, but supersymmetry does not suffice to stabilize the inflaton mass: the inflationary energy necessarily breaks supersymmetry, and the resulting splittings in supermultiplets are of order  $H$ , so that supersymmetry does not protect a small inflaton mass  $m_\phi \ll H$ .

In §28.3 we discuss the natural proposal to protect the inflaton potential via a shift symmetry  $\phi \rightarrow \phi + \text{const.}$ , which is equivalent to identifying the inflaton with a pseudo-Nambu-Goldstone-boson. In the absence of such a symmetry the eta problem seems to imply the necessity of fine-tuning the inflationary action in order to get inflation.

### 27.2.2 Supergravity Example

An important instance of the eta problem arises in locally-supersymmetric theories, *i.e.* in supergravity [75]. This case is relevant for many string theory models of inflation because four-dimensional supergravity is the low-energy effective theory of supersymmetric string compactifications.

In  $\mathcal{N} = 1$  supergravity, a key term in the scalar potential is the F-term potential,

$$V_F = e^{K/M_{\text{pl}}^2} \left[ K^{\varphi\bar{\varphi}} D_\varphi W \overline{D_\varphi W} - \frac{3}{M_{\text{pl}}^2} |W|^2 \right], \quad (337)$$

where  $K(\varphi, \bar{\varphi})$  and  $W(\varphi)$  are the Kähler potential and the superpotential, respectively;  $\varphi$  is a complex scalar field which is taken to be the inflaton; and we have defined  $D_\varphi W \equiv \partial_\varphi W + M_{\text{pl}}^{-2} (\partial_\varphi K) W$ .

For simplicity of presentation, we have assumed that there are no other light degrees of freedom, but generalizing our expressions to include other fields is straightforward.

The Kähler potential determines the inflaton kinetic term,  $-K_{,\varphi\bar{\varphi}} \partial\varphi\partial\bar{\varphi}$ , while the superpotential determines the interactions. To derive the inflaton mass, we expand  $K$  around some chosen origin, which we denote by  $\varphi \equiv 0$  without loss of generality, *i.e.*  $K(\varphi, \bar{\varphi}) = K_0 + K_{,\varphi\bar{\varphi}}|_0 \varphi\bar{\varphi} + \dots$ . The inflationary Lagrangian then becomes

$$\mathcal{L} \approx -K_{,\varphi\bar{\varphi}} \partial\varphi\partial\bar{\varphi} - V_0 \left( 1 + K_{,\varphi\bar{\varphi}}|_0 \frac{\varphi\bar{\varphi}}{M_{\text{pl}}^2} + \dots \right) \quad (338)$$

$$\equiv -\partial\phi\partial\bar{\phi} - V_0 \left( 1 + \frac{\phi\bar{\phi}}{M_{\text{pl}}^2} \right) + \dots, \quad (339)$$

where we have defined the canonical inflaton field  $\phi\bar{\phi} \approx K_{\varphi\bar{\varphi}}|_0 \varphi\bar{\varphi}$  and  $V_0 \equiv V_F|_{\varphi=0}$ . We have retained the leading correction to the potential originating in the expansion of  $e^{K/M_{\text{pl}}^2}$  in Eqn. (337), which could plausibly be called a universal correction in F-term scenarios. The omitted terms, some of which can be of the same order as the terms we keep, arise from expanding

$$\left[ K^{\varphi\bar{\varphi}} D_\varphi W \overline{D_\varphi W} - \frac{3}{M_{\text{pl}}^2} |W|^2 \right] \quad (340)$$

in Eqn. (337) and clearly depend on the model-dependent structure of the Kähler potential and the superpotential.

The result is of the form of Eqn. (334) with

$$\mathcal{O}_6 = V_0 \phi\bar{\phi} \quad (341)$$

and implies a large model-independent contribution to the eta parameter

$$\Delta\eta = 1, \quad (342)$$

as well as a model-dependent contribution which is typically of the same order. It is therefore clear that in an inflationary scenario driven by an F-term potential, eta will generically be of order unity.

Under what circumstances can inflation still occur, in a model based on a supersymmetric Lagrangian? One obvious possibility is that the model-dependent contributions to eta approximately cancel the model-independent contribution, so that the smallness of the inflaton mass is a result of fine-tuning. In the case study of §28.2 we will provide a concrete example in which the structure of all relevant contributions to eta can be computed, so that one can sensibly pursue such a fine-tuning argument.

Clearly, it would be far more satisfying to exhibit a mechanism that *removes* the eta problem by ensuring that  $\Delta\eta \ll 1$ . This requires either that the F-term potential is negligible, or that the inflaton does not appear in the F-term potential. The first case does not often arise, because F-term potentials play an important role in presently-understood models for stabilization of the compact dimensions of string theory [76]. However, in §28.3 we will present a scenario in which the inflaton is an axion and does not appear in the Kähler potential, or in the F-term potential, to any order in perturbation theory. This evades the particular incarnation of the eta problem that we have described above.

## 27.3 The Lyth Bound

In **Lecture 2** we derived the Lyth bound [77]:

$$\frac{\Delta\phi}{M_{\text{pl}}} \simeq \mathcal{O}(1) \left( \frac{r_\star}{0.01} \right)^{1/2}, \quad (343)$$

where  $r_\star$  is the value of the tensor-to-scalar ratio on CMB scales. In any model with  $r_\star > 0.01$  one must therefore ensure that  $\epsilon, |\eta| \ll 1$  over a super-Planckian range  $\Delta\phi > M_{\text{pl}}$ . This result implies two necessary conditions for *large-field inflation*:

- i) an obvious requirement is that large field ranges are *kinematically* allowed, *i.e.* that the scalar field space (in canonical units) has diameter  $> M_{\text{pl}}$ . This is nontrivial, as in typical string compactifications many fields are not permitted such large excursions.
- ii) the flatness of the inflaton potential needs to be controlled *dynamically* over a super-Planckian field range. We discuss this challenge in effective field theory in §27.4 and in string theory in §28.3.

## 27.4 Super-Planckian Fields and Flat Potentials

To begin, let us consider super-Planckian field excursions in the context of Wilsonian effective field theory.

### 27.4.1 No Shift Symmetry

In the absence of any special symmetries, the potential in large-field inflation becomes sensitive to an infinite series of Planck-suppressed operators. The physical interpretation of these terms is as follows: as the inflaton expectation value changes, any other fields  $\chi$  to which the inflaton couples experience changes in mass, self-coupling, etc. In particular, any field coupled with at least gravitational strength to the inflaton experiences significant changes when the inflaton undergoes a super-Planckian excursion. These variations of the  $\chi$  masses and couplings in turn feed back into changes of the inflaton potential and therefore threaten to spoil the delicate flatness required for inflation. Note that this applies not just to the light ( $m \ll H$ ) degrees of freedom, but even to fields with masses near the Planck scale: integrating out Planck-scale degrees of freedom generically (*i.e.*, for couplings of order unity) introduces Planck-suppressed operators in the effective action. For nearly all questions in particle physics, such operators are negligible, but in inflation they play an important role.

The particular operators which appear are determined, as always, by the symmetries of the low-energy action. As an example, imposing only the symmetry  $\phi \rightarrow -\phi$  on the inflaton leads to the following effective action:

$$\mathcal{L}_{\text{eff}}(\phi) = -\frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{1}{4}\lambda\phi^4 - \sum_{p=1}^{\infty} [\lambda_p\phi^4 + \nu_p(\partial\phi)^2] \left( \frac{\phi}{M_{\text{pl}}} \right)^{2p} + \dots \quad (344)$$

Unless the UV theory enjoys further symmetries, one expects that the coefficients  $\lambda_p$  and  $\nu_p$  are of order unity. Thus, whenever  $\phi$  traverses a distance of order  $M_{\text{pl}}$  in a direction that is not protected

by a suitably powerful symmetry, the effective Lagrangian receives substantial corrections from an infinite series of higher-dimension operators. In order to have inflation, the potential should of course be approximately flat over a super-Planckian range. If this is to arise by accident or by fine-tuning, it requires a conspiracy among infinitely many coefficients, which has been termed ‘functional fine-tuning’ (compare this to the eta problem which only requires tuning of one mass parameter).

### 27.4.2 Shift Symmetry

There is a sensible way to control this infinite series of corrections: one can invoke an approximate symmetry that forbids the inflaton from coupling to other fields in any way that would spoil the structure of the inflaton potential. Such a shift symmetry,

$$\phi \rightarrow \phi + const \tag{345}$$

protects the inflaton potential in a natural way.

In the case with a shift symmetry, the action of chaotic inflation [78]

$$\mathcal{L}_{\text{eff}}(\phi) = -\frac{1}{2}(\partial\phi)^2 - \lambda_p \phi^p, \tag{346}$$

with small coefficient  $\lambda_p$  is technically natural. However, because we require that this symmetry protects the inflaton even from couplings to Planck-scale degrees of freedom, it is essential that the symmetry should be approximately respected by the Planck-scale theory – in other words, the proposed symmetry of the low-energy effective action should admit a UV-completion. Hence, large-field inflation should be formulated in a theory that has access to information about approximate symmetries at the Planck scale. Let us remark that in effective field theory in general, UV-completion of an assumed low-energy symmetry is rarely an urgent question. The present situation is different because we do *not* know whether all reasonable effective actions can in fact arise as low-energy limits of string theory, and indeed it has been conjectured that many effective theories do not admit UV-completion in string theory [79–81]. Therefore, it is important to verify that any proposed symmetry of Planck-scale physics can be realized in string theory.

To construct an inflationary model with detectable gravitational waves, we are therefore interested in finding, in string theory, a configuration that has both a large kinematic range, and a potential protected by a shift symmetry that is approximately preserved by the full string theory.

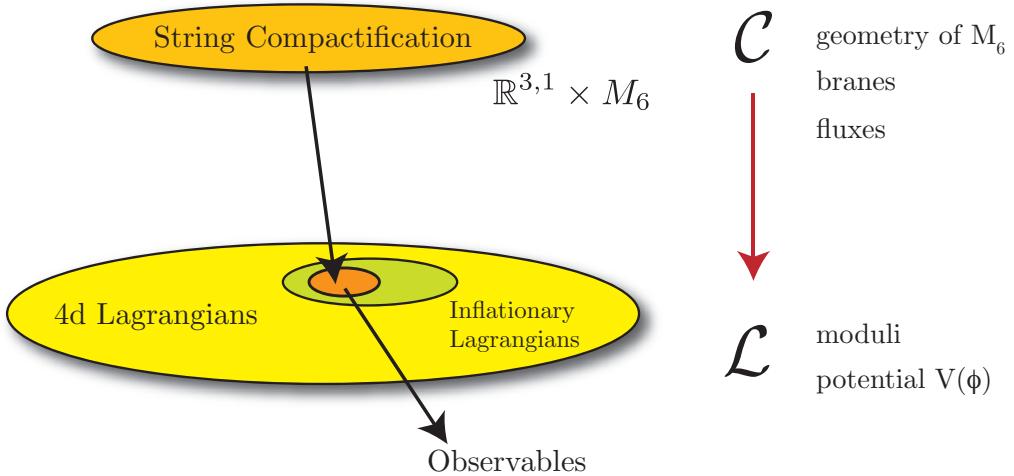
## 28 Inflation in String Theory

### 28.1 From String Compactifications to the Inflaton Action

#### 28.1.1 Elements of String Compactifications

It is a famous fact that the quantum theory of strings is naturally defined in more than four spacetime dimensions, with four-dimensional physics emerging upon compactification of the additional spatial dimensions. For concreteness, we will focus on compactifications of the critical ten-dimensional type IIB string theory on six-dimensional Calabi-Yau spaces.

The vast number of distinct compactifications in this class are distinguished by their topology, geometry, and discrete data such as quantized fluxes and wrapped D-branes. A central task in



**Figure 26:** From 10d Compactification Data to 4d Action.

string theory model-building is to understand in detail how the ten-dimensional sources determine the four-dimensional effective theory. If we denote the ten-dimensional compactification data by  $\mathcal{C}$ , the procedure in question may be written schematically as

$$\mathcal{S}_{10}[\mathcal{C}] \rightarrow \mathcal{S}_4. \quad (347)$$

Distinct compactification data  $\mathcal{C}$  give rise to a multitude of four-dimensional effective theories  $\mathcal{S}_4$  with varied field content, kinetic terms, scalar potentials, and symmetry properties (this is the *landscape* of solutions to string theory). By understanding the space of possible data  $\mathcal{C}$  and the nature of the map in Eqn. (347), we can hope to identify, and perhaps even classify, compactifications that give rise to interesting four-dimensional physics.



**Figure 27:** Elements of Flux Compactifications: Fluxes and Wrapped Branes.

### 28.1.2 The Effective Inflaton Action

For our purposes, the most important degrees of freedom of the effective theory are four-dimensional scalar fields. Scalar fields known as *moduli* arise from deformations of the compactification manifold, typically numbering in the hundreds for the Calabi-Yau spaces under consideration, and from the positions, orientations, and gauge field configurations of any D-branes. From given compactification data one can compute the kinetic terms and scalar potentials of the moduli; in turn, the expectation values of the moduli determine the parameters of the four-dimensional effective theory. In the presence of generic ten-dimensional sources of stress-energy, such as D-branes and quantized fluxes, there is an energy cost for deforming the compactification, and many (though not always all) of the moduli fields become massive [82].

It is useful to divide the scalar fields arising in  $\mathcal{S}_4$  into a set of light fields  $\phi, \psi$  with masses below the Hubble scale ( $m_\phi, m_\psi \ll H$ ) and a set of heavy fields  $\chi$  with masses much greater than the Hubble scale ( $m_\chi \gg H$ ). Here one of the light fields, denoted  $\phi$ , has been identified as the inflaton candidate.

To understand whether successful inflation can occur, one must understand all the scalar fields, both heavy and light. First, sufficiently massive moduli fields are effectively frozen during inflation, and one should integrate them out to obtain an effective action for the light fields only,

$$\mathcal{S}_4(\phi, \psi, \chi) \rightarrow \mathcal{S}_{4,\text{eff}}(\phi, \psi). \quad (348)$$

Integrating out these heavy modes generically induces contributions to the potential of the putative inflaton: that is, *moduli stabilization contributes to the eta problem*. This is completely analogous to the appearance of corrections from higher-dimension operators in our discussion of effective field theory in §27.1.

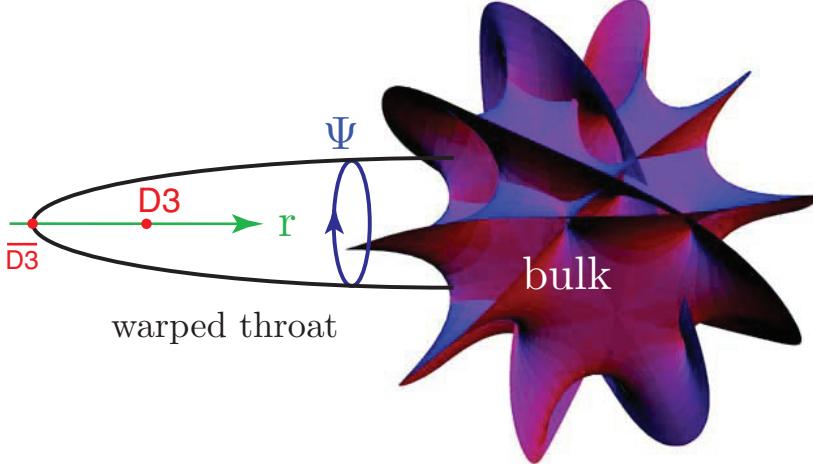
Next, if scalar fields in addition to the inflaton are light during inflation, they typically have important effects on the dynamics, and one should study the evolution of all fields  $\psi$  with masses  $m_\psi \ll H$ . Moreover, even if the resulting multi-field inflationary dynamics is suitable, light degrees of freedom can create problems for late-time cosmology. Light scalars absorb energy during inflation and, if they persist after inflation, they can release this energy during or after Big Bang nucleosynthesis, spoiling the successful predictions of the light element abundances. Moreover, light moduli would be problematic in the present universe, as they mediate fifth forces of gravitational strength. To avoid these late-time problems, it suffices to ensure that  $m_\psi \gg 30$  TeV, as in this case the moduli decay before Big Bang nucleosynthesis. A simplifying assumption that is occasionally invoked is that all fields aside from the inflaton should have  $m \gg H$ , but this is not required on physical grounds: it serves only to arrange that the effective theory during inflation has only a single degree of freedom.

## 28.2 Case Study: Warped D-brane Inflation

In string theory models of inflation the operators contributing to the inflaton potential can be enumerated, and in principle even their coefficients can be computed in terms of given compactification data. To illustrate these issues, it is useful to examine a concrete model in detail. In the following we therefore present a case study of a comparatively well-understood model of small-field inflation, *warped D-brane inflation*.

### 28.2.1 D3-branes in Warped Throat Geometries

In this scenario inflation is driven by the motion of a D3-brane in a warped throat region of a stabilized compact space [83]. To preserve four-dimensional Lorentz (or de Sitter) invariance, the D3-brane fills our four-dimensional spacetime and is pointlike in the extra dimensions (see Figure 28). The global compactification is assumed to be a warped product of four-dimensional spacetime



**Figure 28:** D3-brane inflation in a warped throat geometry. The D3-branes are spacetime-filling in four dimensions and therefore pointlike in the extra dimensions. The circle stands for the base manifold  $X_5$  with angular coordinates  $\Psi$ . The brane moves in the radial direction  $r$ . At  $r_{UV}$  the throat attaches to a compact Calabi-Yau space. Anti-D3-branes minimize their energy at the tip of the throat,  $r_{IR}$ .

(with metric  $g_{\mu\nu}$ ) and a conformally-Calabi-Yau space,

$$ds^2 = e^{2A(y)} g_{\mu\nu} dx^\mu dx^\nu + e^{-2A(y)} g_{mn} dy^m dy^n, \quad (349)$$

with  $g_{mn}$  a Calabi-Yau metric that can be approximated in some region by a cone over a five-dimensional Einstein manifold  $X_5$ ,

$$g_{mn} dy^m dy^n = dr^2 + r^2 ds_{X_5}^2. \quad (350)$$

A canonical example of such a throat region is the Klebanov-Strassler (KS) geometry [84], for which  $X_5$  is the  $(SU(2) \times SU(2))/U(1)$  coset space  $T^{1,1}$ , and the would-be conical singularity at the tip of the throat,  $r = 0$ , is smoothed by the presence of appropriate fluxes. The tip of the throat is therefore located at a finite radial coordinate  $r_{IR}$ , while at  $r = r_{UV}$  the throat is glued into an unwarped bulk geometry. In the relevant regime  $r_{IR} \ll r < r_{UV}$  the warp factor may be written as [85]

$$e^{-4A(r)} = \frac{R^4}{r^4} \ln \frac{r}{r_{IR}}, \quad R^4 \equiv \frac{81}{8} (g_s M \alpha')^2, \quad (351)$$

where

$$\ln \frac{r_{UV}}{r_{IR}} \approx \frac{2\pi K}{3g_s M}. \quad (352)$$

Here,  $M$  and  $K$  are integers specifying the flux background [84, 86].

Warping sourced by fluxes is commonplace in modern compactifications, and there has been much progress in understanding the stabilization of the moduli of such a compactification [82]. Positing a stabilized compactification containing a KS throat therefore seems reasonable given present knowledge.

### 28.2.2 The Field Range Bound

### 28.2.3 The D3-brane Potential

Inflation proceeds as a D3-brane moves radially inward in the throat region, towards an anti-D3-brane that is naturally situated at the tip of the throat. The inflaton kinetic term is determined by the Dirac-Born-Infeld (DBI) action for a probe D3-brane, and leads to an identification of the canonical inflaton field with a multiple of the radial coordinate,  $\phi^2 \equiv T_3 r^2$ . Here,  $T_3 \equiv [(2\pi)^3 g_s \alpha'^2]^{-1}$  is the D3-brane tension, with  $g_s$  the string coupling and  $2\pi\alpha'$  the inverse string tension. The exit from inflation occurs when open strings stretched between the approaching pair become tachyonic and condense, annihilating the branes.

In this simplified picture, inflation is driven by the extremely weak (warping-suppressed) Coulomb interaction of the brane-antibrane pair [83]. The true story, however, is more complex, as moduli stabilization introduces new terms in the inflaton potential which typically overwhelm the Coulomb term and drive more complicated dynamics [83, 87–91]. This pattern is precisely what we anticipated in our effective field theory discussion: integrating out moduli fields can be expected to induce important corrections to the potential.

An important correction induced by moduli stabilization is the inflaton mass term arising from the supergravity F-term potential, §27.2.2. In a vacuum stabilized by an F-term potential, *i.e.* by superpotential terms involving the moduli, one finds the mass term  $H_0^2 \phi^2 = \frac{1}{3} V_0(\phi_\star) \frac{\phi^2}{M_{\text{pl}}^2}$  [83], where  $\phi_\star$  is an arbitrary reference value for the inflaton field and the parameter  $H_0$  should not be confused with the present-day Hubble constant.

However, one expects additional contributions to the potential from a variety of other sources, such as additional effects in the compactification that break supersymmetry [91]. Let us define  $\Delta V(\phi)$  to encapsulate all contributions to the potential aside from the Coulomb interaction  $V_0(\phi)$  and the mass term  $H_0^2 \phi^2$ ; then the total potential and the associated contributions to the eta parameter may be written as

$$V(\phi) = V_0(\phi) + H_0^2 \phi^2 + \Delta V(\phi) \quad (353)$$

$$\eta(\phi) = \eta_0 + \frac{2}{3} + \Delta\eta(\phi) = ? \quad (354)$$

where  $\eta_0 \ll 1$  because the Coulomb interaction is very weak. (More generally,  $V_0(\phi)$  can be *defined* to be all terms in  $V(\phi)$  with negligible contributions to  $\eta$ . Besides the brane-antibrane Coulomb interaction, this can include any other sources of nearly-constant energy, *e.g.* bulk contributions to the cosmological constant.)

Clearly,  $\eta$  can only be small if  $\Delta V$  can cancel the mass term in Eqn. (353). We must therefore enumerate all relevant contributions to  $\Delta V$ , and attempt to understand the circumstances under which an approximate cancellation can occur. Note that identifying a subset of contributions to  $\Delta V$  while remaining ignorant of others is insufficient.

Warped D-brane inflation has received a significant amount of theoretical attention in part because of its high degree of computability. Quite generally, if we had access to the full data of an explicit, stabilized compactification with small curvatures and weak string coupling, we would in principle be able to compute the potential of a D-brane inflaton to any desired accuracy, by performing a careful dimensional reduction. This is not possible at present for a generic compact Calabi-Yau, for two reasons: for general Calabi-Yau spaces hardly any metric data is available, and examples with entirely explicit moduli stabilization are rare.

However, a sufficiently long throat is well-approximated by a *noncompact* throat geometry (*i.e.*, a throat of infinite length), for which the Calabi-Yau metric can often be found, as in the important example of the Klebanov-Strassler solution [84], which is entirely explicit and everywhere smooth. Having complete metric data greatly facilitates the study of probe D-brane dynamics, at least at the level of an unstabilized compactification. Furthermore, we will now explain how the effects of moduli stabilization and of the finite length of the throat can be incorporated systematically. The method involves examining perturbations to the supergravity solution that describes the throat in which the D3-brane moves. For concreteness we will work with the example of a KS throat, but the method is far more general. Our treatment will allow us to give explicit expressions for the correction terms  $\Delta V$  in Eqn. (353), and hence to extract the characteristics of inflation in the presence of moduli stabilization.

#### 28.2.4 Sketch of the Supergravity Analysis

In the following we describe the computation of the inflaton potential for warped D3-brane inflation. This is only meant to give a flavor of the challenges involved in understanding the full potential. For the full details we refer the reader to [32, 87–89, 91].

Type IIB string theory contains a good dozen of fields (going by names such as dilaton,  $p$ -form fluxes, warp factors, etc.). In principle we would have to worry that all those fields could couple to the inflaton degree of freedom and hence have to be considered when computing the inflaton potential to the desired accuracy. However, D3-branes are special in that they only couple to a specific combination of the warp factor and the five-form flux and are blind to perturbations in all other fields

$$V_{D3}(\phi) = T_3(e^{4A} - \alpha) \equiv T_3\Phi_- \quad (355)$$

where the scalar function  $\alpha(y)$  is related to the five-form flux  $F_5$ . We are therefore interested in perturbations of the object  $\Phi_- = e^{4A} - \alpha$ . In the KS background  $\Phi_-$  vanishes, but coupling of the throat to the bulk geometry and interaction with moduli-stabilizing degrees of freedom like wrapped D7-branes, induces a non-zero  $\Phi_-$ . To study the induced  $\Phi_-$  perturbations, we investigate the equation of motion

$$\nabla^2\Phi_-(\phi, \Psi) = \frac{1}{24}|\mathcal{G}_-|^2 + R, \quad (356)$$

where  $\mathcal{G}_-$  is a special (imaginary anti-self-dual) combination of 3-form fluxes and  $R$  is the 4-dimensional Ricci scalar. Solving Eqn. (356) (in combination with an equation for  $\mathcal{G}_-$ ) one finds [91]

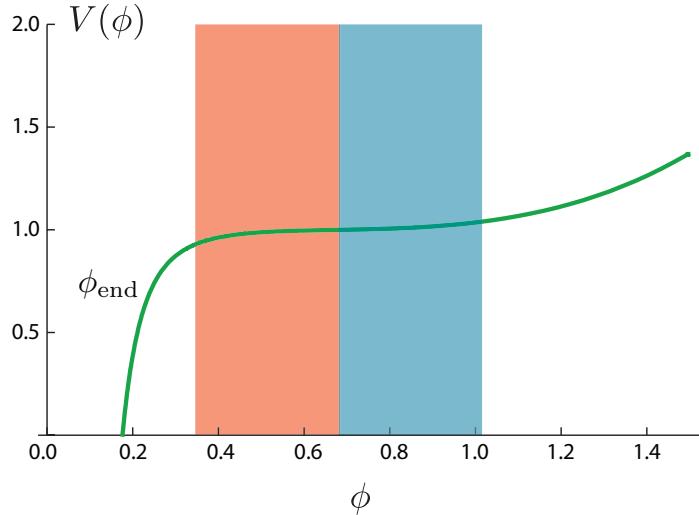
$$V_{D3} = \sum_{\Delta} \phi^{\Delta} f_{\Delta}(\Psi), \quad (357)$$

where

$$\Delta = 1, \frac{3}{2}, 2, \dots \quad (358)$$

The discrete *spectrum* (358) of corrections to the inflaton potential determines the phenomenology of the model.

### 28.2.5 Phenomenological Implications



**Figure 29:** Inflection Point Inflation

### 28.2.6 Summary and Perspective

In §27.1 we explained how the eta problem is sensitive to dimension-six Planck suppressed operators. In effective field theory models of inflation one can of course always *assume* a solution to the eta problem by a cancellation of the contributing correction terms; in other words, one can postulate that a flat potential  $V(\phi)$  arises after an approximate cancellation among dimension-six Planck-suppressed corrections. In string theory models of inflation, to follow this path would be to abdicate the opportunity to use Planck-suppressed contributions as a (limited) window onto string theory. Moreover, once  $\phi$  is identified with a physical degree of freedom of a string compactification, the precise form of the potential is in principle fully specified by the remaining data of the compactification. (Mixing conjecture into the analysis at this stage would effectively transform a ‘string-derived’ scenario into a ‘string-inspired’ scenario; the latter may be interesting as a cosmological model, but will not contribute to our understanding of string theory.) Thus, overcoming the eta problem becomes a detailed computational question. One can in principle compute the full potential from first principles, and in practice one can often classify corrections to the leading-order potential.

In this section, we have enumerated the leading corrections for warped D-brane inflation and showed that an accidental cancellation (or fine-tuning) allows small eta over a limited range of inflaton values. This gives a non-trivial existence proof for inflationary solutions in warped throat models with D3-branes.

## 28.3 Case Study: Axion Monodromy Inflation

We now turn to our second case study, an example of large-field inflation in string theory. As we have just discussed, the particular challenge in these models is the need to control an infinite series of contributions to the inflaton potential, arising from couplings of the inflaton to degrees of freedom with masses near the Planck scale. Direct enumeration and fine-tuning of such terms (as in the small-field example in §28.2) is manifestly impractical, and it appears essential to develop a symmetry argument controlling or forbidding these terms.

An influential proposal in this direction is Natural Inflation [92] (see [Lecture 1](#)), in which a pseudo-Nambu-Goldstone boson (*i.e.*, an axion) is the inflaton. At the perturbative level, the axion field  $a$  enjoys a continuous shift symmetry  $a \rightarrow a + \text{const}$  which is broken by nonperturbative effects to a discrete symmetry  $a \rightarrow a + 2\pi$ . The nonperturbative effects generate a periodic potential

$$V(\phi) = \frac{\Lambda^4}{2} \left[ 1 - \cos \left( \frac{\phi}{f} \right) \right] + \dots, \quad (359)$$

where  $\Lambda$  is a dynamically-generated scale,  $f$  is known as the axion decay constant,  $\phi \equiv af$ , and the omitted terms are higher harmonics.

As explained above, an important question, in any proposed effective theory in which a super-Planckian field range is protected by a shift symmetry, is whether this structure can be UV-completed. We should therefore search in string theory for an axion with decay constant  $f > M_{\text{pl}}$ .

### 28.3.1 Axions in String Theory

#### Axions from $p$ -Forms

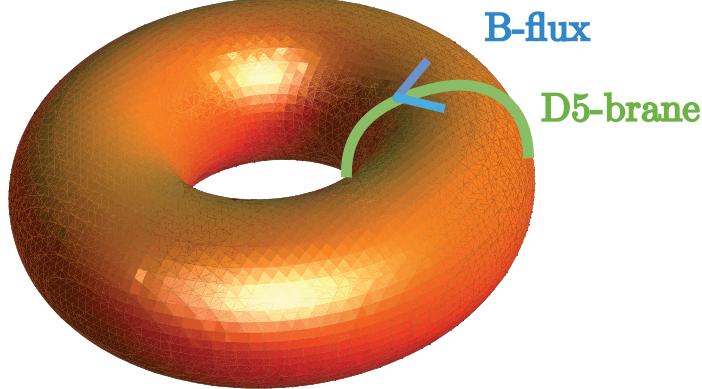
Axions are plentiful in string compactifications, arising from  $p$ -form gauge potentials integrated on  $p$ -cycles of the compact space. For example, in type IIB string theory, there are axions  $b_i = 2\pi \int_{\Sigma_i} B$  arising from integrating the Neveu-Schwarz (NS) two-form  $B$  over two-cycles  $\Sigma_i$ , as well as axions  $c_i = 2\pi \int_{\Sigma_i} C$  arising from the Ramond-Ramond (RR) two-form  $C$ . In the absence of additional ingredients such as fluxes and space-filling wrapped branes, the potential for these axions is classically flat and has a continuous shift symmetry which originates in the gauge invariance of the ten-dimensional action. Instanton effects break this symmetry to a discrete subgroup,  $b_i \rightarrow b_i + 2\pi$  ( $c_i \rightarrow c_i + 2\pi$ ). This leads to a periodic contribution to the axion potential whose periodicity we will now estimate. We will find that the axion decay constants are smaller than  $M_{\text{pl}}$  in known, computable limits of string theory [93, 94]. Readers less familiar with string compactifications can accept this assertion and skip to §28.3.2 without loss of continuity.

#### Axion Decay Constants in String Theory

Let  $\omega^i$  be a basis for  $H^2(X, \mathbb{Z})$ , the space of two-forms on the compact space  $X$ , with  $\int_{\Sigma_i} \omega^j = \alpha' \delta_i^j$ . The NS two-form potential  $B$  may be expanded as

$$B = \frac{1}{2\pi} \sum_i b_i(x) \omega^i, \quad (360)$$

with  $x$  the four-dimensional spacetime coordinate. The axion decay constant can be inferred from the normalization of the axion kinetic term, which in this case descends from the ten-dimensional



**Figure 30:** Axion Monodromy

term

$$\frac{1}{(2\pi)^7 g_s^2 \alpha'^4} \int d^{10}x \frac{1}{2} |dB|^2 \supset \frac{1}{2} \int d^4x \sqrt{-g} \gamma^{ij} (\partial^\mu b_i \partial_\mu b_j) , \quad (361)$$

where

$$\gamma^{ij} \equiv \frac{1}{6(2\pi)^9 g_s^2 \alpha'^4} \int_X \omega^i \wedge \star \omega^j \quad (362)$$

and  $\star$  is the six-dimensional Hodge star. By performing the integral over the internal space  $X$  and diagonalizing the field space metric as  $\gamma^{ij} \rightarrow f_i^2 \delta_{ij}$ , one can extract the axion decay constant  $f_i$ .

It is too early to draw universal conclusions, but a body of evidence suggests that the resulting axion periodicities are always smaller than  $M_{\text{pl}}$  in computable limits of string theory [93, 94]. As this will be essential for our arguments, we will illustrate this result in a simple example. Suppose that the compactification is isotropic, with typical length-scale  $L$  and volume  $L^6$ . Then using

$$\alpha' M_{\text{pl}}^2 = \frac{2}{(2\pi)^7} \frac{L^6}{g_s^2 \alpha'^3} \quad (363)$$

we find from Eq. (362) that

$$f^2 \approx M_{\text{pl}}^2 \frac{\alpha'^2}{6(2\pi)^2 L^4} . \quad (364)$$

In controlled compactifications we require  $L \gg \sqrt{\alpha'}$ , so that  $f \ll M_{\text{pl}}$ . Qualitatively similar conclusions apply in much more general configurations [93, 94].

### 28.3.2 Axion Inflation in String Theory

The above result would seem to imply that Natural Inflation from a single axion field cannot be realized in known string compactifications: string theory provides many axions, but none of these has a sufficiently large field range. However, there are at least two reasonable proposals to circumvent this obstacle.

#### N-inflation

The first suggestion was that a collective excitation of many hundreds of axions could have an effective field range large enough for inflation [95, 96]. The role of the inflaton is played by the

collective field

$$\phi^2 = \sum_{i=1}^N \phi_i^2. \quad (365)$$

Even if each individual field has a sub-Planckian field range,  $\phi_i < M_{\text{pl}}$ , for sufficiently large number of fields  $N$ , the effective field  $\phi$  can have a super-Planck excursion. This ‘N-flation’ proposal is a specific example of assisted inflation [97], but, importantly, one in which symmetry helps to protect the axion potential from corrections that would impede inflation. Although promising, this scenario still awaits a proof of principle demonstration, as the presence of a large number of light fields leads to a problematic renormalization of the Newton constant, and hence to an effectively reduced field range. For recent studies of N-flation see [98, 99].

## Axion Monodromy

We will instead describe an elementary mechanism, *monodromy*, which allows inflation to persist through multiple circuits of a single periodic axion field. A system is said to undergo monodromy if, upon transport around a closed loop in the (naive) configuration space, the system reaches a new configuration. A spiral staircase is a canonical example: the naive configuration space is described by the angular coordinate, but the system changes upon transport by  $2\pi$ . (In fact, we will find that this simple model gives an excellent description of the potential in axion monodromy inflation.) The idea of using monodromy to achieve controlled large-field inflation in string theory was first proposed by Silverstein and Westphal [31], who discussed a model involving a D4-brane wound inside a nilmanifold. In this section we will focus instead on the subsequent axion monodromy proposal of Ref. [100], where a monodromy arises in the four-dimensional potential energy upon transport around a circle in the field space parameterized by an axion.

Monodromies of this sort are possible in a variety of compactifications, but we will focus on a single concrete example. Consider type IIB string theory on a Calabi-Yau orientifold, *i.e.* a quotient of a Calabi-Yau manifold by a discrete symmetry that includes worldsheet orientation reversal and a geometric involution. Specifically, we will suppose that the involution has fixed points and fixed four-cycles, known as O3-planes and O7-planes, respectively. If in addition the compactification includes a D5-brane that wraps a suitable two-cycle  $\Sigma$  and fills spacetime, then the axion  $b = 2\pi \int_{\Sigma} B$  can exhibit monodromy in the potential energy. (Similarly, a wrapped NS5-brane produces monodromy for the axion  $c = 2\pi \int_{\Sigma} C$ .) In other words, a D5-brane wrapping  $\Sigma$  carries a potential energy that is *not* a periodic function of the axion, as the shift symmetry of the axion action is broken by the presence of the wrapped brane; in fact, the potential energy increases without bound as  $b$  increases.

In the D5-brane case, the relevant potential comes from the Dirac-Born-Infeld action for the wrapped D-brane,

$$S_{\text{DBI}} = \frac{1}{(2\pi)^5 g_s \alpha'^3} \int_{\mathcal{M}_4 \times \Sigma} d^6 \xi \sqrt{\det(G + B)} \quad (366)$$

$$= \frac{1}{(2\pi)^6 g_s \alpha'^2} \int_{\mathcal{M}_4} d^4 x \sqrt{-g} \sqrt{(2\pi)^2 \ell_{\Sigma}^4 + b^2}, \quad (367)$$

where  $\ell_{\Sigma}$  is the size of the two-cycle  $\Sigma$  in string units. The brane energy, Eq. (367), is clearly not invariant under the shift symmetry  $b \rightarrow b + 2\pi$ , although this is a symmetry of the corresponding compactification without the wrapped D5-brane. Thus, the DBI action leads directly to monodromy

for  $b$ . Moreover, when  $b \gg \ell_\Sigma^2$ , the potential is asymptotically *linear* in the canonically-normalized field  $\varphi_b \propto b$ .

The qualitative inflationary dynamics in this model is as follows: One begins with a D5-brane wrapping a curve  $\Sigma$ , upon which  $\int_\Sigma B$  is taken to be large. In other words, the axion  $b$  has a large initial vev. Inflation proceeds by the reduction of this vev, until finally  $\int_\Sigma B = 0$  and the D5-brane is nearly ‘empty’, *i.e.* has little worldvolume flux. During this process the D5-brane does not move, nor do any of the closed-string moduli shift appreciably. For small axion vevs, the asymptotically linear potential we have described is inaccurate, and the curvature of the potential becomes non-negligible; see Eqn. (367). At this stage, the axion begins to oscillate around its origin. Couplings between the axion and other degrees of freedom, either closed string modes or open string modes, drain energy from the inflaton oscillations. If a sufficient fraction of this energy is eventually transmitted to visible-sector degrees of freedom – which may reside, for example, on a stack of D-branes elsewhere in the compactification – then the hot Big Bang begins. The details of reheating depend strongly on the form of the couplings between the Standard Model degrees of freedom and the inflaton, and this is an important open question, both in this model and in string inflation more generally.

### 28.3.3 Compactification Considerations

Having explained the essential idea of axion monodromy inflation, we must still ensure that the proposed inflationary mechanism is compatible with moduli stabilization and can be realized in a consistent compactification. An immediate concern is whether there are additional contributions to the potential, beyond the linear term identified above, that could have important effects during inflation. As we have emphasized throughout this review, one expects that in the absence of a symmetry protecting the inflaton potential, generic corrections due to moduli stabilization will contribute  $\Delta\eta \sim \mathcal{O}(1)$ . It is therefore essential to verify that the continuous shift symmetry which protects the inflaton potential is preserved to an appropriate degree by the stabilized compactification. For the special case of moduli stabilization in which nonperturbative effects play a role, ensuring that the shift symmetry is not spoiled can be quite subtle. This is described in detail in [100].

### 28.3.4 Summary and Perspective

The Lyth bound shows that an observable gravitational wave signal correlates with the inflaton field moving over a super-Planckian distance during inflation. Effective field theory models of large-field inflation then require a shift symmetry to protect the flatness of the potential over a super-Planckian range. It has therefore become an important question whether such shift symmetries arise in string theory and can be used to realize large-field inflation.

In this section, we argued that the first examples of shift symmetries in string theory that protect the potential over a super-Planckian range are becoming available. We explained the dual role of the monodromy: i) it results in a large kinematic field range  $\Delta\phi > M_{\text{pl}}$  by allowing a small fundamental domain to be traversed repeatedly, and ii) in combination with the shift symmetry it controls corrections to the potential over a super-Planckian range. The shift symmetry, only weakly broken by  $V$ , controls corrections  $\Delta V$  within a fundamental domain, and furthermore relates corrections in one fundamental domain to those in any other. Monodromy therefore effectively

reduces a large-field problem to a small-field problem [31].

Although more work is required to understand these models and the compactifications in which they arise, monodromy appears to be a robust and rather promising mechanism for realizing large-field inflation, and hence an observable gravitational wave signal, in string theory.

## 29 Outlook

### 29.1 Theoretical Prospects

As we hope this lecture has illustrated, theoretical progress in recent years has been dramatic. A decade ago, only a few proposals for connecting string theory to cosmology were available, and the problem of stabilizing the moduli had not been addressed. We now have a wide array of inflationary models motivated by string theory, and the best-studied examples among these incorporate some information about moduli stabilization. Moreover, a few mechanisms for inflation in string theory have been shown to be robust, persisting after full moduli stabilization with all relevant corrections included.

Aside from demonstrating that inflation is possible in string theory, what has been accomplished? In our view the primary use of explicit models of inflation in string theory is as test cases, or toy models, for the sensitivity of inflation to quantum gravity. On the theoretical front, these models have underlined the importance of the eta problem in general field theory realizations of inflation; they have led to mechanisms for inflation that might seem unnatural in field theory, but are apparently natural in string theory; and they have sharpened our understanding of the implications of a detection of primordial tensor modes.

It is of course difficult to predict the direction of future theoretical progress, not least because unforeseen fundamental advances in string theory can be expected to enlarge the toolkit of inflationary model-builders. However, it is safe to anticipate further gradual progress in moduli stabilization, including the appearance of additional explicit examples with all moduli stabilized; entirely explicit models of inflation in such compactifications will undoubtedly follow. At present, few successful models exist in M-theory or in heterotic string theory, and under mild assumptions, inflation can be shown to be impossible in certain classes of type IIA compactifications [101–103]. It would be surprising if it turned out that inflation is much more natural in one weakly-coupled limit of string theory than in the rest, and the present disparity can be attributed in part to the differences among the moduli-stabilizing tools presently available in the various limits. Clearly, it would be useful to understand how inflation can arise in more diverse string vacua.

The inflationary models now available in string theory are subject to stringent theoretical constraints arising from consistency requirements (*e.g.*, tadpole cancellation) and from the need for some degree of computability. In turn, these limitations lead to correlations among the cosmological observables, *i.e.* to predictions. Some of these constraints will undoubtedly disappear as we learn to explore more general string compactifications. However, one can hope that some constraints may remain, so that the set of inflationary effective actions derived from string theory would be a proper subset of the set of inflationary effective actions in a general quantum field theory. Establishing such a proposition would require a far more comprehensive understanding of string compactifications than is available at present.

## 29.2 Observational Signatures?

The theoretical aspects of inflation described in this lecture are interesting largely because they can be *tested* experimentally using present and future cosmological data (see [Lecture 3](#)).

As we have repeatedly emphasized throughout these lectures, the most dramatic confirmation of inflation would come from a detection of  $B$ -mode polarization, which would establish the energy scale of inflation and would indicate that the inflaton traversed a super-Planckian distance. As we have argued in this lecture, super-Planckian displacements are a key instance in which the inflaton effective action is particularly sensitive to the physics of the Planck scale. As a concrete example of the discriminatory power of tensor perturbations, any detection of primordial gravitational waves would exclude the warped D3-brane inflation scenario of §28.2 [104], while an upper bound  $r < 0.07$  (or a detection with  $r \gg 0.07$ ) would exclude the axion monodromy scenario of §28.3 [100].

A further opportunity arises because single-field slow-roll inflation predicts null results for many cosmological observables, as the primordial scalar fluctuations are predicted to be scale-invariant, Gaussian and adiabatic to a high degree. A detection of non-Gaussianity, isocurvature fluctuations or a large scale-dependence (running) would therefore rule out single-field slow-roll inflation. Inflationary effective actions that do allow for a significant non-Gaussianity, non-adiabaticity or scale-dependence often require higher-derivative interactions and/or more than one light field, and such actions arise rather naturally in string theory. Although we have focused in this lecture on the sensitivity of the inflaton potential to Planck-scale physics, the inflaton kinetic term is equally UV-sensitive, and string theory provides a promising framework for understanding the higher-derivative interactions that can produce significant non-Gaussianity [105, 106].

Finally, CMB temperature and polarization anisotropies induced by relic cosmic strings or other topological defects provide probes of the physics of the end of inflation or of the post-inflationary era. Cosmic strings are automatically produced at the end of brane-antibrane inflation [107, 108], and the stability and phenomenological properties of the resulting cosmic string network are determined by the properties of the warped geometry. Detecting cosmic superstrings via lensing or through their characteristic bursts of gravitational waves is an exciting prospect.

## **30 Summary: Lecture 5**

Recent work by many authors has led to the emergence of robust mechanisms for inflation in string theory. The primary motivations for these works are the sensitivity of inflationary effective actions to the ultraviolet completion of gravity, and the prospect of empirical tests using precision cosmological data. String theory models of inflation have now achieved a reasonable level of theoretical control and are genuinely falsifiable by observational data. Indeed, many string inflation models are already significantly constrained by the current data. A more difficult question is how cosmological observations might possibly provide evidence in favor of a string theory model of inflation. Present observations and present theoretical considerations do not oblige us to expect an eventual positive result. However, if we are fortunate enough to detect evidence for string theory in the sky, this will most plausibly arise through a distinctive signature that is unnatural, albeit presumably possible, in field theory models. Perhaps the best hope would be a striking correlation of many observables.

# **Part VII**

# **Conclusions**

## **31 Future Prospects and Open Problems**

coming soon ...

## 32 Guide to Further Reading

The following textbooks, reviews and papers have been useful to me in the preparation of these lectures. The student will find valuable further details about inflation in those works.

### Textbooks

- Mukhanov, *Physical Foundations of Cosmology*  
A nice treatment of early universe cosmology and the theory of cosmological perturbations.
- Dodelson, *Modern Cosmology*  
An excellent book about cosmology with a strong focus on the cosmic microwave background.  
Very readable, *i.e.* you can read it while lying down.
- Weinberg, *Cosmology*  
It is by Steven Weinberg!
- Liddle and Lyth, *Cosmological Inflation and Large-Scale Structure*
- Longair, *Galaxy Formation*
- Birell and Davies, *Quantum Field Theory in Curved Spacetime*

### Reviews

- Baumann et al., *Probing Inflation with CMB Polarization*  
White paper of the Inflation Working Group of the CMBPol Mission Concept Study. More than 60 experts on inflation combined to write this very comprehensive review.
- Baumann and Peiris, *Cosmological Inflation: Theory and Observations*  
In this review Hiranya Peiris and I summarize the basics of inflation and CMB observations for a non-expert audience. This might be too elementary for the readers of these lectures, but could be of interest if you are looking for some bedtime reading.
- Baumann and McAllister, *Advances in String Inflation*  
In this review Liam McAllister and I describe the challenge of realizing inflation in string theory.
- Lyth and Riotto, *Particle Physics Models of Inflation*  
What these lectures lack on inflationary model-building may be found here.
- Bassett et al., *Inflation Dynamics and Reheating*  
What these lectures lack on reheating may be found here.

- Kinney, *TASI Lectures on Inflation*

Will Kinney's lectures at TASI 2008. Perfect as a first read on inflation. It is hoped that these TASI 2009 lectures make a good second read. I tried to complement Will's lectures by giving more technical details.

- Malik and Wands, *Cosmological Perturbations*

A nice review of first and second-order perturbation theory. Many useful formulas.

## Papers

Some of the original papers on inflation are very accessible and well worth reading:

- Guth, *Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems*

This classic is of course a must-read. It provides a very clear explanation of the big bang puzzles.

- Maldacena, *Non-Gaussian Features of Primordial Fluctuations in Single Field Inflationary Models*

This paper provided the first rigorous computation of the three-point function for slow-roll inflation. It also gives one of the clearest and most elegant expositions of the calculation of the power spectra of inflationary fluctuations. My treatment in these lectures was heavily inspired by Maldacena's paper.

# Part VIII

# Appendix

## A Cosmological Perturbation Theory

In this Appendix we summarize basic facts of cosmological perturbation theory. This is based on unpublished lecture notes of a course at Princeton University by Uros Seljak and Chris Hirata as well as a review by Malik and Wands [109].

### A.1 The Perturbed Universe

We consider perturbations to the homogeneous background spacetime and the stress-energy of the universe.

#### A.1.1 Metric Perturbations

The most general first-order perturbation to a spatially flat FRW metric is

$$ds^2 = -(1 + 2\Phi)dt^2 + 2a(t)B_i dx^i dt + a^2(t)[(1 - 2\Psi)\delta_{ij} + 2E_{ij}]dx^i dx^j \quad (\text{A.1})$$

where  $\Phi$  is a 3-scalar called the *lapse*,  $B_i$  is a 3-vector called the *shift*,  $\Psi$  is a 3-scalar called the spatial *curvature* perturbation, and  $E_{ij}$  is a spatial *shear* 3-tensor which is symmetric and traceless,  $E_i^i = \delta^{ij}E_{ij} = 0$ . 3-surfaces of constant time  $t$  are called *slices* and curves of constant spatial coordinates  $x^i$  but varying time  $t$  are called *threads*.

#### A.1.2 Stress-Energy Perturbations

The stress-energy tensor may be described by a density  $\rho$ , a pressure  $p$ , a 4-velocity  $u^\mu$  (of the frame in which the 3-momentum density vanishes), and an anisotropic stress  $\Sigma^{\mu\nu}$ .

Density and pressure perturbations are defined in an obvious way

$$\delta\rho(t, x^i) \equiv \rho(t, x^i) - \bar{\rho}(t), \quad \text{and} \quad \delta p(t, x^i) \equiv p(t, x^i) - \bar{p}(t). \quad (\text{A.2})$$

Here, the background values have been denoted by overbars. The 4-velocity has only three independent components (after the metric is fixed) since it has to satisfy the constraint  $g_{\mu\nu}u^\mu u^\nu = -1$ . In the perturbed metric (A.1) the perturbed 4-velocity is

$$u_\mu \equiv (-1 - \Phi, v_i), \quad \text{or} \quad u^\mu \equiv (1 - \Phi, v^i + B^i). \quad (\text{A.3})$$

Here,  $u_0$  is chosen so that the constraint  $u_\mu u^\mu = -1$  is satisfied to first order in all perturbations. Anisotropic stress vanishes in the unperturbed FRW universe, so  $\Sigma^{\mu\nu}$  is a first-order perturbation. Furthermore,  $\Sigma^{\mu\nu}$  is constrained by

$$\Sigma^{\mu\nu}u_\nu = \Sigma_\mu^\mu = 0. \quad (\text{A.4})$$

The orthogonality with  $u_\mu$  implies  $\Sigma^{00} = \Sigma^{0j} = 0$ , *i.e.* only the spatial components  $\Sigma^{ij}$  are non-zero. The trace condition then implies  $\Sigma_i^i = 0$ . Anisotropic stress is therefore a traceless symmetric 3-tensor.

Finally, with these definitions the perturbed stress-tensor is

$$T_0^0 = -(\bar{\rho} + \delta\rho) \quad (\text{A.5})$$

$$T_i^0 = (\bar{\rho} + \bar{p})v_i \quad (\text{A.6})$$

$$T_0^i = -(\bar{\rho} + \bar{p})(v^i + B^i) \quad (\text{A.7})$$

$$T_j^i = \delta_j^i(\bar{p} + \delta p) + \Sigma_j^i. \quad (\text{A.8})$$

If there are several contributions to the stress-energy tensor (*e.g.* photons, baryons, dark matter, etc.), they are added:  $T_{\mu\nu} = \sum_I T_{\mu\nu}^I$ . This implies

$$\delta\rho = \sum_I \delta\rho_I \quad (\text{A.9})$$

$$\delta p = \sum_I \delta p_I \quad (\text{A.10})$$

$$(\bar{\rho} + \bar{p})v^i = \sum_I (\bar{\rho}_I + \bar{p}_I)v_I^i \quad (\text{A.11})$$

$$\Sigma^{ij} = \sum_I \Sigma_I^{ij}. \quad (\text{A.12})$$

Density, pressure and anisotropic stress perturbations simply add. However, velocities do not add, which motivates defining the 3-momentum density

$$\delta q^i \equiv (\bar{\rho} + \bar{p})v^i, \quad (\text{A.13})$$

such that

$$\delta q^i = \sum_I \delta q_I^i. \quad (\text{A.14})$$

## A.2 Scalars, Vectors and Tensors

The Einstein equations relate metric perturbations to the stress-energy perturbations. Einstein's equations are both complicated (coupled second-order partial differential equations) and non-linear. Fortunately, the symmetries of the flat FRW background spacetime allow perturbations to be decomposed into independent scalar, vector and tensor components. This reduces the Einstein equations to a set of uncoupled ordinary differential equations.

### A.2.1 Helicity and SVT-Decomposition in Fourier Space

The decomposition into scalar, vector and tensor perturbations is most elegantly explained in Fourier space. We define the Fourier components of a general perturbation  $\delta Q(\mathbf{x})$  as follows

$$\delta Q(t, \mathbf{k}) = \int d^3x \delta Q(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}. \quad (\text{A.15})$$

First note that as a consequence of translation invariance different Fourier modes (different wavenumbers  $k$ ) evolve independently.<sup>24</sup>

---

<sup>24</sup>The following proof was related to me by Uros Seljak and Chris Hirata.

---

**Proof:**

Consider the linear evolution of  $N$  perturbations  $\delta Q_I$ ,  $I = 1, \dots, N$  from an initial time  $t_1$  to a final time  $t_2$

$$\delta Q_I(t_2, \mathbf{k}) = \sum_{J=1}^N \int d^3\bar{\mathbf{k}} T_{IJ}(t_2, t_1, \mathbf{k}, \bar{\mathbf{k}}) \delta Q_J(t_1, \bar{\mathbf{k}}), \quad (\text{A.16})$$

where the transfer matrix  $T_{IJ}(t_2, t_1, \mathbf{k}, \bar{\mathbf{k}})$  follows from the Einstein equations and we have allowed for the possibility of a mixing of  $k$ -modes. We now show that translation invariance in fact forbids such couplings. Consider the coordinate transformation

$$x^{i'} = x^i + \Delta x^i, \quad \text{where } \Delta x^i = \text{const.} \quad (\text{A.17})$$

You may convince yourself that the Fourier amplitude gets shifted as follows

$$\delta Q'_I(t, \mathbf{k}) = e^{-ik_j \Delta x^j} \delta Q_I(t, \mathbf{k}). \quad (\text{A.18})$$

Thus the evolution equation in the primed coordinate system becomes

$$\delta Q'_I(t_2, \mathbf{k}) = \sum_{J=1}^N \int d^3\bar{\mathbf{k}} e^{-ik_j \Delta x^j} T_{IJ}(t_2, t_1, \mathbf{k}, \bar{\mathbf{k}}) e^{i\bar{k}_j \Delta x^j} \delta Q'_J(t_1, \bar{\mathbf{k}}) \quad (\text{A.19})$$

$$\equiv \sum_{J=1}^N \int d^3\bar{\mathbf{k}} T'_{IJ}(t_2, t_1, \mathbf{k}, \bar{\mathbf{k}}) \delta Q_J(t_1, \bar{\mathbf{k}}). \quad (\text{A.20})$$

By translation invariance the equations of motion must be the same in both coordinate systems, *i.e.* the transfer matrices  $T_{IJ}$  and  $T'_{IJ}$  must be the same

$$T_{IJ}(t_2, t_1, \mathbf{k}, \bar{\mathbf{k}}) = e^{i(\bar{k}_j - k_j) \Delta x^j} T_{IJ}(t_2, t_1, \mathbf{k}, \bar{\mathbf{k}}). \quad (\text{A.21})$$

This must hold for all  $\Delta x^j$ . Hence, either  $\bar{\mathbf{k}} = \mathbf{k}$  or  $T_{IJ}(t_2, t_1; \mathbf{k}, \bar{\mathbf{k}}) = 0$ , *i.e.* the perturbation  $\delta Q_I(t_2, \mathbf{k})$  of wavevector  $\mathbf{k}$  depends only on the initial perturbations of wavevector  $\mathbf{k}$ . At linear order there is no coupling of different  $k$ -modes. QED.

---

Now consider rotations around the Fourier vector  $\mathbf{k}$  by an angle  $\psi$ . We classify perturbations according to their *helicity*  $m$ : a perturbation of helicity  $m$  has its amplitude multiplied by  $e^{im\psi}$  under the above rotation. We define scalar, vector and tensor perturbations as having helicities 0,  $\pm 1$ ,  $\pm 2$ , respectively.

Consider a Fourier mode with wavevector  $\mathbf{k}$ . Without loss of generality we may assume that  $\mathbf{k} = (0, 0, k)$  (or use rotational invariance of the background). The spatial dependence of any perturbation then is

$$\delta Q \propto e^{ikx^3}. \quad (\text{A.22})$$

To study rotations around  $\mathbf{k}$  it proves convenient to switch to the helicity basis

$$\mathbf{e}_\pm \equiv \frac{\mathbf{e}_1 \pm i\mathbf{e}_2}{\sqrt{2}}, \quad \mathbf{e}_3, \quad (\text{A.23})$$

where  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  is the Cartesian basis. A rotation around the 3-axis by an angle  $\psi$  has the following effect

$$\begin{pmatrix} x^{1'} \\ x^{2'} \end{pmatrix} = \begin{pmatrix} \cos \psi & \sin \psi \\ -\sin \psi & \cos \psi \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}, \quad x^{3'} = x^3, \quad (\text{A.24})$$

and

$$\mathbf{e}'_{\pm} = e^{\pm i\psi} \mathbf{e}_{\pm}, \quad \mathbf{e}'_3 = \mathbf{e}_3. \quad (\text{A.25})$$

The contravariant components of any tensor  $T_{i_1 i_2 \dots i_n}$  transform as

$$T'_{i_1 i_2 \dots i_n} = e^{i(n_+ - n_-)\psi} T_{i_1 i_2 \dots i_n} \equiv e^{im\psi} T_{i_1 i_2 \dots i_n} \quad (\text{A.26})$$

where  $n_+$  and  $n_-$  count the number of plus and minus indices in  $i_1 \dots i_n$ , respectively. Helicity is defined as the difference  $m \equiv n_+ - n_-$ .

In the helicity basis  $\{\mathbf{e}_{\pm}, \mathbf{e}_3\}$ , a 3-scalar  $\alpha$  has a single component with no indices and is therefore obviously of helicity 0; a 3-vector  $\beta_i$  has 3 components  $\beta_+, \beta_-, \beta_3$  of helicity  $\pm 1$  and 0; a symmetric and traceless 3-tensor  $\gamma_{ij}$  has 5 components  $\gamma_{--}, \gamma_{++}, \gamma_{-3}, \gamma_{+3}, \gamma_{33}$  (the tracelessness condition makes  $\gamma_{-+}$  redundant), of helicity  $\pm 2, \pm 1$  and 0.

Rotational invariance of the background implies that helicity scalars, vectors and tensors evolve independently.<sup>25</sup>

### Proof:

Consider  $N$  perturbations  $\delta Q_I$ ,  $I = 1, \dots, N$  of helicity  $m_I$ . The linear evolution is

$$\delta Q_I(t_2, \mathbf{k}) = \sum_{J=1}^N T_{IJ}(t_2, t_1, \mathbf{k}) \delta Q_J(t_1, \mathbf{k}), \quad (\text{A.27})$$

where the transfer matrix  $T_{IJ}(t_2, t_1, \mathbf{k})$  follows from the Einstein equations. Under rotation the perturbations transform as

$$\delta Q'_I(t, \mathbf{k}) = e^{im_I\psi} \delta Q_I(t, \mathbf{k}) \quad (\text{A.28})$$

and

$$\delta Q'_I(t_2, \mathbf{k}) = \sum_{J=1}^N e^{im_I\psi} T_{IJ}(t_2, t_1, \mathbf{k}) e^{-im_J\psi} \delta Q'_J(t_1, \mathbf{k}). \quad (\text{A.29})$$

By rotational invariance of the equations of motion

$$T_{IJ}(t_2, t_1, \mathbf{k}) = e^{im_I\psi} T_{IJ}(t_2, t_1, \mathbf{k}) e^{-im_J\psi} = e^{i(m_I - m_J)\psi} T_{IJ}(t_2, t_1, \mathbf{k}), \quad (\text{A.30})$$

which has to hold for any angle  $\psi$ ; it follows that either  $m_I = m_J$ , *i.e.*  $\delta Q_I$  and  $\delta Q_J$  have the same helicity or  $T_{IJ}(t_2, t_1, \mathbf{k}) = 0$ . This proves that the equations of motion don't mix modes of different helicity. QED.

<sup>25</sup>The following proof was related to me by Uros Seljak and Chris Hirata.

### A.2.2 Real Space SVT-Decomposition

In the last section we have seen that 3-scalars correspond to helicity scalars, 3-vectors decompose into helicity scalars and vectors, and 3-tensors decompose into helicity scalars, vectors and tensors. We now look at this from a different perspective.

A 3-scalar is obviously also a helicity scalar

$$\alpha = \alpha^S. \quad (\text{A.31})$$

Consider a 3-vector  $\beta_i$ . We argue that it can be decomposed as

$$\beta_i = \beta_i^S + \beta_i^V, \quad (\text{A.32})$$

where

$$\beta_i^S = \nabla_i \hat{\beta}, \quad \nabla^i \beta_i^V = 0, \quad (\text{A.33})$$

or, in Fourier space,

$$\beta_i^S = -\frac{ik_i}{k} \beta, \quad k_i \beta_i^V = 0. \quad (\text{A.34})$$

Here, we have defined  $\beta \equiv k \hat{\beta}$ .

**Exercise 9** Show that  $\beta_i^V$  is a helicity vector.

Similarly, a traceless, symmetric 3-tensor can be written as

$$\gamma_{ij} = \gamma_{ij}^S + \gamma_{ij}^V + \gamma_{ij}^T, \quad (\text{A.35})$$

where

$$\gamma_{ij}^S = \left( \nabla_i \nabla_j - \frac{1}{3} \delta_{ij} \nabla^2 \right) \hat{\gamma} \quad (\text{A.36})$$

$$\gamma_{ij}^V = \frac{1}{2} (\nabla_i \hat{\gamma}_j + \nabla_j \hat{\gamma}_i), \quad \nabla_i \hat{\gamma}_i = 0 \quad (\text{A.37})$$

$$\nabla_i \gamma_{ij}^T = 0. \quad (\text{A.38})$$

or

$$\gamma_{ij}^S = \left( -\frac{k_i k_j}{k^2} + \frac{1}{3} \delta_{ij} \right) \gamma \quad (\text{A.39})$$

$$\gamma_{ij}^V = -\frac{i}{2k} (k_i \gamma_j + k_j \gamma_i), \quad k_i \gamma_i = 0 \quad (\text{A.40})$$

$$k_i \gamma_{ij}^T = 0. \quad (\text{A.41})$$

Here, we have defined  $\gamma \equiv k^2 \hat{\gamma}$  and  $\gamma_i \equiv k \hat{\gamma}_i$ .

**Exercise 10** Show that  $\gamma_{ij}^V$  and  $\gamma_{ij}^T$  are a helicity vector and a helicity tensor, respectively.

Choosing  $\mathbf{k}$  along the 3-axis, *i.e.*  $\mathbf{k} = (0, 0, k)$  we find

$$\gamma_{ij}^S = \frac{1}{3} \begin{pmatrix} \gamma & 0 & 0 \\ 0 & \gamma & 0 \\ 0 & 0 & -2\gamma \end{pmatrix} \quad (\text{A.42})$$

$$\gamma_{ij}^V = -\frac{i}{2} \begin{pmatrix} 0 & 0 & \gamma_1 \\ 0 & 0 & \gamma_2 \\ \gamma_1 & \gamma_2 & 0 \end{pmatrix} \quad (\text{A.43})$$

$$\gamma_{ij}^T = \begin{pmatrix} \gamma^\times & \gamma^+ & 0 \\ \gamma^+ & -\gamma^\times & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (\text{A.44})$$

## A.3 Scalars

### A.3.1 Metric Perturbations

Four scalar metric perturbations  $\Phi$ ,  $B_{,i}$ ,  $\Psi\delta_{ij}$  and  $E_{,ij}$  may be constructed from 3-scalars, their derivatives and the background spatial metric, *i.e.*

$$ds^2 = -(1 + 2\Phi)dt^2 + 2a(t)B_{,i}dx^i dt + a^2(t)[(1 - 2\Psi)\delta_{ij} + 2E_{,ij}]dx^i dx^j \quad (\text{A.45})$$

Here, we have absorbed the  $\nabla^2 E \delta_{ij}$  part of the helicity scalar  $E_{ij}^S$  in  $\Psi\delta_{ij}$ .

The intrinsic Ricci scalar curvature of constant time hypersurfaces is

$$R_{(3)} = \frac{4}{a^2} \nabla^2 \Psi. \quad (\text{A.46})$$

This explains why  $\Psi$  is often referred to as the curvature perturbation.

There are two scalar gauge transformations

$$t \rightarrow t + \alpha, \quad (\text{A.47})$$

$$x^i \rightarrow x^i + \delta^{ij}\beta_{,j}. \quad (\text{A.48})$$

Under these coordinate transformations the scalar metric perturbations transform as

$$\Phi \rightarrow \Phi - \dot{\alpha} \quad (\text{A.49})$$

$$B \rightarrow B + a^{-1}\alpha - a\dot{\beta} \quad (\text{A.50})$$

$$E \rightarrow E - \beta \quad (\text{A.51})$$

$$\Psi \rightarrow \Psi + H\alpha. \quad (\text{A.52})$$

Note that the combination  $\dot{E} - B/a$  is independent of the spatial gauge and only depends on the temporal gauge. It is called the scalar potential for the anisotropic shear of world lines orthogonal to constant time hypersurfaces. To extract physical results it is useful to define gauge-invariant combinations of the scalar metric perturbations. Two important gauge-invariant quantities were introduced by Bardeen [110]

$$\Phi_B \equiv \Phi - \frac{d}{dt}[a^2(\dot{E} - B/a)] \quad (\text{A.53})$$

$$\Psi_B \equiv \Psi + a^2H(\dot{E} - B/a). \quad (\text{A.54})$$

### A.3.2 Matter Perturbations

Matter perturbations are also gauge-dependent, *e.g.* density and pressure perturbations transform as follows under temporal gauge transformations

$$\delta\rho \rightarrow \delta\rho - \dot{\bar{\rho}}\alpha, \quad \delta p \rightarrow \delta p - \dot{\bar{p}}\alpha. \quad (\text{A.55})$$

Adiabatic pressure perturbations are defined as

$$\delta p_{ad} \equiv \frac{\dot{\bar{p}}}{\bar{\rho}} \delta\rho. \quad (\text{A.56})$$

The non-adiabatic, or entropic, part of the pressure perturbations is then gauge-invariant

$$\delta p_{en} \equiv \delta p - \frac{\dot{\bar{p}}}{\bar{\rho}} \delta\rho. \quad (\text{A.57})$$

The scalar part of the 3-momentum density,  $(\delta q)_i$ , transforms as

$$\delta q \rightarrow \delta q + (\bar{\rho} + \bar{p})\alpha. \quad (\text{A.58})$$

We may then define the gauge-invariant comoving density perturbation

$$\delta\rho_m \equiv \delta\rho - 3H\delta q. \quad (\text{A.59})$$

Finally, two important gauge-invariant quantities are formed from combinations of matter and metric perturbations. The *curvature perturbation on uniform density hypersurfaces* is

$$-\zeta \equiv \Psi + \frac{H}{\dot{\bar{\rho}}} \delta\rho. \quad (\text{A.60})$$

The *comoving curvature perturbation* is

$$\mathcal{R} = \Psi - \frac{H}{\bar{\rho} + \bar{p}} \delta q. \quad (\text{A.61})$$

We will show that  $\zeta$  and  $\mathcal{R}$  are equal on superhorizon scales, where they become time-independent. The computation of the inflationary perturbation spectrum is most clearly phrased in terms of  $\zeta$  and  $\mathcal{R}$ .

### A.3.3 Einstein Equations

To relate the metric and stress-energy perturbations, we consider the perturbed Einstein equations

$$\delta G_{\mu\nu} = 8\pi G \delta T_{\mu\nu}. \quad (\text{A.62})$$

We work at linear order. This leads to the *energy and momentum constraint equations*

$$3H(\dot{\Psi} + H\Phi) + \frac{k^2}{a^2} [\Psi + H(a^2\dot{E} - aB)] = -4\pi G \delta\rho \quad (\text{A.63})$$

$$\dot{\Psi} + H\Phi = -4\pi G \delta q. \quad (\text{A.64})$$

These can be combined into the gauge-invariant *Poisson equation*

$$\frac{k^2}{a^2} \Psi_B = -4\pi G \delta \rho_m. \quad (\text{A.65})$$

The Einstein equation also yield two *evolution equations*

$$\ddot{\Psi} + 3H\dot{\Psi} + H\dot{\Phi} + (3H^2 + 2\dot{H})\Phi = 4\pi G \left( \delta p - \frac{2}{3}k^2 \delta \Sigma \right) \quad (\text{A.66})$$

$$(\partial_t + 3H)(\dot{E} - B/a) + \frac{\Psi - \Phi}{a^2} = 8\pi G \delta \Sigma. \quad (\text{A.67})$$

The last equation may be written as

$$\Psi_B - \Phi_B = 8\pi G a^2 \delta \Sigma. \quad (\text{A.68})$$

In the absence of anisotropic stress this implies,  $\Psi_B = \Phi_B$ .

Energy-momentum conservation,  $\nabla_\mu T^{\mu\nu} = 0$ , gives the *continuity and Euler equations*

$$\dot{\delta\rho} + 3H(\delta\rho + \delta p) = \frac{k^2}{a^2} \delta q + (\bar{\rho} + \bar{p})[3\dot{\Psi} + k^2(\dot{E} + B/a)] \quad (\text{A.69})$$

$$\dot{\delta q} + 3H\delta q = -\delta p + \frac{2}{3}k^2 \delta \Sigma - (\bar{\rho} + \bar{p})\Phi. \quad (\text{A.70})$$

Expressed in terms of the curvature perturbation on uniform-density hypersurfaces,  $\zeta$ , Eqn. (A.69) reads

$$\dot{\zeta} = -H \frac{\delta p_{en}}{\bar{\rho} + \bar{p}} - \Pi, \quad (\text{A.71})$$

where  $\delta p_{en}$  is the non-adiabatic component of the pressure perturbation, and  $\Pi$  is the scalar shear along comoving world lines

$$\frac{\Pi}{H} \equiv -\frac{k^2}{3H} \left[ \dot{E} - B/a + \frac{\delta q}{a^2(\bar{\rho} + \bar{p})} \right] \quad (\text{A.72})$$

$$= -\frac{k^2}{3a^2 H^2} \left[ \zeta - \Psi_B \left( 1 - \frac{2\bar{\rho}}{9(\bar{\rho} + \bar{p})} \frac{k^2}{a^2 H^2} \right) \right]. \quad (\text{A.73})$$

For adiabative perturbations,  $\delta p_{en} = 0$  on superhorizon scales,  $k/(aH) \ll 1$  (*i.e.*  $\Pi/H \rightarrow 0$  for finite  $\zeta$  and  $\Psi_B$ ), the curvature perturbation  $\zeta$  is constant. This is a crucial result for our computation of the inflationary spectrum of  $\zeta$  in **Lecture 2**. It justifies computing  $\zeta$  at horizon exit and ignoring superhorizon evolution.

### A.3.4 Popular Gauges

For reference we now give the Einstein and conservation equations in various popular gauges:

- **Synchronous gauge**

A popular gauge, especially for numerical implementation of the perturbation equations (*cf.* CMB-FAST or CAMB), is synchronous gauge. It is defined by

$$\Phi = B = 0. \quad (\text{A.74})$$

The Einstein equations become

$$3H\dot{\Psi} + \frac{k^2}{a^2} [\Psi + Ha^2\dot{E}] = -4\pi G \delta\rho \quad (\text{A.75})$$

$$\dot{\Psi} = -4\pi G \delta q \quad (\text{A.76})$$

$$\ddot{\Psi} + 3H\dot{\Psi} = 4\pi G \left( \delta p - \frac{2}{3}k^2\delta\Sigma \right) \quad (\text{A.77})$$

$$(\partial_t + 3H)\dot{E} + \frac{\Psi}{a^2} = 8\pi G \delta\Sigma. \quad (\text{A.78})$$

The conservation equation are

$$\dot{\delta\rho} + 3H(\delta\rho + \delta p) = \frac{k^2}{a^2}\delta q + (\bar{\rho} + \bar{p})[3\dot{\Psi} + k^2\dot{E}] \quad (\text{A.79})$$

$$\dot{\delta q} + 3H\delta q = -\delta p + \frac{2}{3}k^2\delta\Sigma. \quad (\text{A.80})$$

- **Newtonian gauge**

The Newtonian gauge has its name because it reduces to Newtonian gravity in the small-scale limit. It is popular for analytic work since it leads to algebraic relations between metric and stress-energy perturbations.

Newtonian gauge is defined by

$$B = E = 0 \quad (\text{A.81})$$

and

$$ds^2 = (1 + 2\Phi)dt^2 + a^2(t)(1 - 2\Psi)\delta_{ij}dx^i dx^j \quad (\text{A.82})$$

The Einstein equations are

$$3H(\dot{\Psi} + H\Phi) + \frac{k^2}{a^2}\Psi = -4\pi G \delta\rho \quad (\text{A.83})$$

$$\dot{\Psi} + H\Phi = -4\pi G \delta q \quad (\text{A.84})$$

$$\ddot{\Psi} + 3H\dot{\Psi} + H\dot{\Phi} + (3H^2 + 2\dot{H})\Phi = 4\pi G \left( \delta p - \frac{2}{3}k^2\delta\Sigma \right) \quad (\text{A.85})$$

$$\frac{\Psi - \Phi}{a^2} = 8\pi G \delta\Sigma \quad (\text{A.86})$$

The continuity equations are

$$\dot{\delta\rho} + 3H(\delta\rho + \delta p) = \frac{k^2}{a^2}\delta q + 3(\bar{\rho} + \bar{p})\dot{\Psi} \quad (\text{A.87})$$

$$\dot{\delta q} + 3H\delta q = -\delta p + \frac{2}{3}k^2\delta\Sigma - (\bar{\rho} + \bar{p})\Phi \quad (\text{A.88})$$

- **Uniform density gauge**

The uniform density gauge is useful for describing the evolution of perturbations on super-horizon scales. As its name suggests it is defined by

$$\delta\rho = 0 \quad (\text{A.89})$$

In addition, it is convenient to take

$$E = 0, \quad -\Psi \equiv \zeta \quad (\text{A.90})$$

The Einstein equations are

$$3H(-\dot{\zeta} + H\Phi) - \frac{k^2}{a^2} [\zeta + aHB] = 0 \quad (\text{A.91})$$

$$-\dot{\zeta} + H\Phi = -4\pi G \delta q \quad (\text{A.92})$$

$$-\ddot{\zeta} - 3H\dot{\zeta} + H\dot{\Phi} + (3H^2 + 2\dot{H})\Phi = 4\pi G \left( \delta p - \frac{2}{3}k^2\delta\Sigma \right) \quad (\text{A.93})$$

$$(\partial_t + 3H)B/a + \frac{\zeta + \Phi}{a^2} = -8\pi G \delta\Sigma \quad (\text{A.94})$$

The continuity equations are

$$3H\delta p = \frac{k^2}{a^2}\delta q + (\bar{\rho} + \bar{p})[-3\dot{\zeta} + k^2B/a] \quad (\text{A.95})$$

$$\dot{\delta q} + 3H\delta q = -\delta p + \frac{2}{3}k^2\delta\Sigma - (\bar{\rho} + \bar{p})\Phi \quad (\text{A.96})$$

- **Comoving gauge**

Comoving gauge is defined by the vanishing of the scalar momentum density,

$$\delta q = 0, \quad E = 0 \quad (\text{A.97})$$

It is also conventional to set  $-\Psi \equiv \mathcal{R}$  in this gauge.

The Einstein equations are

$$3H(-\dot{\mathcal{R}} + H\Phi) + \frac{k^2}{a^2} [-\mathcal{R} - aHB] = -4\pi G \delta\rho \quad (\text{A.98})$$

$$-\dot{\mathcal{R}} + H\Phi = 0 \quad (\text{A.99})$$

$$-\ddot{\mathcal{R}} - 3H\dot{\mathcal{R}} + H\dot{\Phi} + (3H^2 + 2\dot{H})\Phi = 4\pi G \left( \delta p - \frac{2}{3}k^2\delta\Sigma \right) \quad (\text{A.100})$$

$$(\partial_t + 3H)B/a + \frac{\mathcal{R} + \Phi}{a^2} = -8\pi G \delta\Sigma \quad (\text{A.101})$$

The continuity equations are

$$\dot{\delta\rho} + 3H(\delta\rho + \delta p) = (\bar{\rho} + \bar{p})[-3\dot{\mathcal{R}} + k^2B/a] \quad (\text{A.102})$$

$$0 = -\delta p + \frac{2}{3}k^2\delta\Sigma - (\bar{\rho} + \bar{p})\Phi. \quad (\text{A.103})$$

Equations (A.103) and (A.99) may be combined into

$$\Phi = \frac{-\delta p + \frac{2}{3}\Sigma}{\bar{\rho} + \bar{p}}, \quad kB = \frac{4\pi Ga^2\delta\rho - k^2\mathcal{R}}{aH}. \quad (\text{A.104})$$

- **Spatially-flat gauge**

A convenient gauge for computing inflationary perturbation is spatially-flat gauge

$$\Psi = E = 0. \quad (\text{A.105})$$

During inflation all scalar perturbations are then described by  $\delta\phi$ .

The Einstein equations are

$$3H^2\Phi + \frac{k^2}{a^2}[-aHB)] = -4\pi G \delta\rho \quad (\text{A.106})$$

$$H\Phi = -4\pi G \delta q \quad (\text{A.107})$$

$$H\dot{\Phi} + (3H^2 + 2\dot{H})\Phi = 4\pi G \left( \delta p - \frac{2}{3}k^2\delta\Sigma \right) \quad (\text{A.108})$$

$$(\partial_t + 3H)B/a + \frac{\Phi}{a^2} = -8\pi G \delta\Sigma. \quad (\text{A.109})$$

The continuity equations are

$$\dot{\delta\rho} + 3H(\delta\rho + \delta p) = \frac{k^2}{a^2}\delta q + (\bar{\rho} + \bar{p})[k^2B/a] \quad (\text{A.110})$$

$$\dot{\delta q} + 3H\delta q = -\delta p + \frac{2}{3}k^2\delta\Sigma - (\bar{\rho} + \bar{p})\Phi. \quad (\text{A.111})$$

## A.4 Vectors

### A.4.1 Metric Perturbations

Vector type metric perturbations are defined as

$$ds^2 = -dt^2 + 2a(t)S_i dx^i dt + a^2(t)[\delta_{ij} + 2F_{(i,j)}]dx^i dx^j, \quad (\text{A.112})$$

where  $S_{i,i} = F_{i,i} = 0$ . The vector gauge transformation is

$$x^i \rightarrow x^i + \beta^i, \quad \beta_{i,i} = 0. \quad (\text{A.113})$$

They lead to the transformations

$$S_i \rightarrow S_i + a\dot{\beta}_i \quad (\text{A.114})$$

$$F_i \rightarrow F_i - \beta_i. \quad (\text{A.115})$$

The combination  $\dot{F}_i + S_i/a$  is called the gauge-invariant vector shear perturbation.

### A.4.2 Matter Perturbations

We define the vector part of the anisotropic stress by

$$\delta\Sigma_{ij} = \partial_{(i}\Sigma_{j)}. \quad (\text{A.116})$$

### A.4.3 Einstein Equations

For vector perturbation there are only two Einstein equations,

$$\dot{\delta q_i} + 3H\delta q_i = k^2\delta\Sigma_i, \quad (\text{A.117})$$

$$k^2(\dot{F}_i + S_i/a) = 16\pi G\delta q_i. \quad (\text{A.118})$$

In the absence of anisotropic stress ( $\delta\Sigma_i = 0$ ) the divergence-free momentum  $\delta q_i$  decays with the expansion of the universe; see Eqn. (A.117). The shear perturbation  $\dot{F}_i + S_i/a$  then vanishes by Eqn. (A.118). Under most circumstances vector perturbations are therefore subdominant. They won't play an important role in these lectures. In particular, vector perturbations aren't created by inflation.

## A.5 Tensors

### A.5.1 Metric Perturbations

Tensor metric perturbations are defined as

$$ds^2 = -dt^2 + a^2(t)[\delta_{ij} + h_{ij}]dx^i dx^j, \quad (\text{A.119})$$

where  $h_{ij,i} = h_i^i = 0$ . Tensor perturbations are automatically gauge-invariant (at linear order). It is conventional to decompose tensor perturbations into eigenmodes of the spatial Laplacian,  $\nabla^2 e_{ij} = -k^2 e_{ij}$ , with comoving wavenumber  $k$  and scalar amplitude  $h(t)$

$$h_{ij} = h(t)e_{ij}^{(+,\times)}(x). \quad (\text{A.120})$$

Here, + and  $\times$  denote the two possible polarization states.

### A.5.2 Matter Perturbations

Tensor perturbations are sourced by anisotropic stress  $\Sigma_{ij}$ .

### A.5.3 Einstein Equations

For tensor perturbations there is only one Einstein equation. In the absence of anisotropic this is

$$\ddot{h} + 3H\dot{h} + \frac{k^2}{a^2}h = 0. \quad (\text{A.121})$$

This is a wave equation describing the evolution of gravitational waves in an expanding universe. Gravitational waves are produced by inflation, but then decay with the expansion of the universe. However, at recombination their amplitude may still be large enough to leave distinctive signatures in  $B$ -modes of CMB polarization.

## A.6 Statistics

We recall some basic facts about statistics. More details may be found in Licia Verde's notes [111].

### A.6.1 Fourier Conventions

Different conventions exist for the normalization of Fourier transforms. Defining

$$\mathcal{R}_\mathbf{k} = A \int d^3x \mathcal{R}(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} \quad (\text{A.122})$$

$$\mathcal{R}(\mathbf{x}) = B \int d^3k \mathcal{R}_\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{x}} \quad (\text{A.123})$$

implies that the Dirac delta function is

$$\delta(\mathbf{k}) = BA \int d^3x e^{\pm i\mathbf{k}\cdot\mathbf{x}}, \quad BA = \frac{1}{(2\pi)^3}. \quad (\text{A.124})$$

Except for the constraint  $BA = 1/(2\pi)^3$  different conventions are possible for the values of  $A$  and  $B$ . These conventions can lead to some confusion about factors of  $2\pi$  in the normalization of the power spectrum. In the main text we follow the convention  $A = 1$ ,  $B = 1/(2\pi)^3$  (the other common convention is  $A = B = 1/(2\pi)^{3/2}$ ; it is nice, since it makes the basis function  $e^{i\mathbf{k}\mathbf{x}}$  orthonormal rather than just orthogonal.), but in this Appendix we will keep things general in order to help identifying normalization errors in the literature.

### A.6.2 Two-Point Correlation Function

We make frequent use of the two-point correlation function

$$\xi_{\mathcal{R}}(r) \equiv \langle \mathcal{R}(\mathbf{x}) \mathcal{R}(\mathbf{x} + \mathbf{r}) \rangle. \quad (\text{A.125})$$

Here, we have made the assumption that by isotropy  $\xi$  depends only on  $r \equiv |\mathbf{r}|$  (distance not orientation).

### A.6.3 Power Spectrum

Consider the following ensemble average

$$\langle \mathcal{R}_\mathbf{k} \mathcal{R}_{\mathbf{k}'} \rangle, \quad (\text{A.126})$$

where  $\mathcal{R}_\mathbf{k}^* = \mathcal{R}_{-\mathbf{k}}$  because  $\mathcal{R}(\mathbf{x})$  is real. Substituting (A.122) gives

$$\langle \mathcal{R}_\mathbf{k} \mathcal{R}_{\mathbf{k}'} \rangle = A^2 \int d^3x e^{-i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}} \int d^3r \xi_{\mathcal{R}}(r) e^{-i\mathbf{k}\cdot\mathbf{r}} \quad (\text{A.127})$$

$$= \frac{A}{B} \delta(\mathbf{k} + \mathbf{k}') \int d^3r \xi_{\mathcal{R}}(r) e^{-i\mathbf{k}\cdot\mathbf{r}}. \quad (\text{A.128})$$

If we *define* the power spectrum as the Fourier transform of the two-point correlation function

$$P_{\mathcal{R}}(k) \equiv A \int d^3r \xi_{\mathcal{R}}(r) e^{-i\mathbf{k}\cdot\mathbf{r}}, \quad (\text{A.129})$$

then we get

$$\langle \mathcal{R}_{\mathbf{k}} \mathcal{R}_{\mathbf{k}'} \rangle = \frac{1}{B} P_{\mathcal{R}}(k) \delta(\mathbf{k} + \mathbf{k}') . \quad (\text{A.130})$$

Notice that often the power spectrum is defined as

$$\langle \mathcal{R}_{\mathbf{k}} \mathcal{R}_{\mathbf{k}'} \rangle = (2\pi)^3 P_{\mathcal{R}}(k) \delta(\mathbf{k} + \mathbf{k}') . \quad (\text{A.131})$$

In the present discussion we realize that this implies a fixed Fourier convention,  $B = 1/(2\pi)^3$ , if we mean by the power spectrum really the Fourier transform of the two-point function; this is often not done correctly in the literature.

Consider the variance

$$\sigma_{\mathcal{R}}^2 \equiv \langle \mathcal{R}^2(x) \rangle = \xi_{\mathcal{R}}(0) = B \int d^3 k P_{\mathcal{R}}(k) . \quad (\text{A.132})$$

This is often defined as

$$\sigma_{\mathcal{R}}^2 \equiv \int d \ln k \Delta_{\mathcal{R}}^2(k) , \quad (\text{A.133})$$

where

$$\Delta_{\mathcal{R}}^2(k) \equiv 4\pi B k^3 P_{\mathcal{R}}(k) . \quad (\text{A.134})$$

In the common Fourier convention  $B = 1/(2\pi)^3$  this becomes

$$\Delta_{\mathcal{R}}^2(k) \equiv \frac{k^3}{2\pi^2} P_{\mathcal{R}}(k) . \quad (\text{A.135})$$

For other Fourier conventions the relation between  $\Delta_{\mathcal{R}}^2(k)$  and  $P_{\mathcal{R}}(k)$  will differ by a numerical factor.

#### A.6.4 Bispectrum

For Gaussian perturbations the power spectrum contains all the information (all higher-order correlation functions can be expressed in terms of the two-point function). Non-Gaussianity is measured by a non-zero three-point function, or equivalently in Fourier space the bispectrum

$$\langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k}_2} \mathcal{R}_{\mathbf{k}_3} \rangle = (2\pi)^3 B_{\mathcal{R}}(k_1, k_2, k_3) \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) . \quad (\text{A.136})$$

## B Free Field Action for $\mathcal{R}$

In this appendix we compute the second order action for the comoving curvature perturbation  $\mathcal{R}$ . This is a basic element for the quantization of cosmological scalar perturbations in **Lecture 2**.

We consider slow-roll models of inflation which are described by a canonical scalar field  $\phi$  minimally coupled to gravity

$$S = \frac{1}{2} \int d^4x \sqrt{-g} [R - (\nabla\phi)^2 - 2V(\phi)], \quad (\text{A.137})$$

in units where  $M_{\text{pl}}^{-2} \equiv 8\pi G = 1$ . We will study perturbations of this action due to fluctuations in the scalar field  $\delta\phi(t, x^i) \equiv \phi(t, x^i) - \bar{\phi}(t)$  and the metric. We will treat metric fluctuations in the ADM formalism [112].

### B.1 Slow-Roll Background

We consider a flat background metric

$$ds^2 = -dt^2 + a(t)^2 \delta_{ij} dx^i dx^j = a^2(\tau) (-d\tau^2 + \delta_{ij} dx^i dx^j), \quad (\text{A.138})$$

with scale factor  $a(t)$  and Hubble parameter  $H(t) \equiv \partial_t \ln a$  satisfying the Friedmann equations

$$3H^2 = \frac{1}{2}\dot{\phi}^2 + V(\phi), \quad \dot{H} = -\frac{1}{2}\dot{\phi}^2. \quad (\text{A.139})$$

The scalar field satisfies the Klein-Gordon equation

$$\ddot{\phi} + 3H\dot{\phi} + V_{,\phi} = 0. \quad (\text{A.140})$$

The standard slow-roll parameters are

$$\epsilon = \frac{1}{2} \left( \frac{V_{,\phi}}{V} \right)^2 \approx \frac{1}{2} \frac{\dot{\phi}^2}{H^2}, \quad \eta = \frac{V_{,\phi\phi}}{V} \approx -\frac{\ddot{\phi}}{H\dot{\phi}} + \frac{1}{2} \frac{\dot{\phi}^2}{H^2}. \quad (\text{A.141})$$

### B.2 ADM Formalism

We treat fluctuations in the ADM formalism [112] where spacetime is sliced into three-dimensional hypersurfaces

$$ds^2 = -N^2 dt^2 + g_{ij}(dx^i + N^i dt)(dx^j + N^j dt). \quad (\text{A.142})$$

Here,  $g_{ij}$  is the three-dimensional metric on slices of constant  $t$ . The lapse function  $N(\mathbf{x})$  and the shift function  $N_i(\mathbf{x})$  contain the same information as the metric perturbations  $\Phi$  and  $B$  in (A.45). However, they were chosen in such a way that they appear as non-dynamical Lagrange multipliers in the action, *i.e.* their equations of motion are purely algebraic. The action (A.137) becomes

$$S = \frac{1}{2} \int d^4x \sqrt{-g} \left[ NR^{(3)} - 2NV + N^{-1}(E_{ij}E^{ij} - E^2) + N^{-1}(\dot{\phi} - N^i \partial_i \phi)^2 - Ng^{ij} \partial_i \phi \partial_j \phi - 2V \right], \quad (\text{A.143})$$

where

$$E_{ij} \equiv \frac{1}{2}(\dot{g}_{ij} - \nabla_i N_j - \nabla_j N_i), \quad E = E_i^i. \quad (\text{A.144})$$

$E_{ij}$  is related to the extrinsic curvature of the three-dimensional spatial slices  $K_{ij} = N^{-1}E_{ij}$ .

**Exercise 11 (ADM action)** *Confirm (A.143).*

### B.2.1 Comoving Gauge

To fix time and spatial reparameterizations we choose the following gauge for the dynamical fields  $g_{ij}$  and  $\phi$

$$\delta\phi = 0, \quad g_{ij} = a^2[(1 - 2\mathcal{R})\delta_{ij} + h_{ij}], \quad \partial_i h_{ij} = h_i^j = 0. \quad (\text{A.145})$$

In this gauge the inflaton field is unperturbed and all scalar degrees of freedom are parameterized by the metric fluctuation  $\mathcal{R}(t, \mathbf{x})$ . Geometrically,  $\mathcal{R}$  measures the spatial curvature of constant- $\phi$  hypersurfaces,  $R^{(3)} = 4\nabla^2\mathcal{R}/a^2$ . An important property of  $\mathcal{R}$  is that it remains constant outside the horizon. This allows us in **Lecture 2** to restrict our computation to correlation functions at horizon crossing.

### B.2.2 Constraint Equations

The ADM action (A.143) implies the following constraint equations for the Lagrange multipliers  $N$  and  $N^i$

$$\nabla_i[N^{-1}(E_j^i - \delta_j^i E)] = 0, \quad (\text{A.146})$$

$$R^{(3)} - 2V - N^{-2}(E_{ij}E^{ij} - E^2) - N^{-2}\dot{\phi}^2 = 0. \quad (\text{A.147})$$

**Exercise 12 (Constraint equations)** *Derive the constraint equations (A.146) and (A.147) from the ADM action (A.143).*

To solve the constraints, we split the shift vector  $N_i$  into irrotational (scalar) and incompressible (vector) parts

$$N_i \equiv \psi_{,i} + \tilde{N}_i, \quad \text{where} \quad \tilde{N}_{i,i} = 0, \quad (\text{A.148})$$

and define the lapse perturbation as

$$N \equiv 1 + \alpha. \quad (\text{A.149})$$

The quantities  $\alpha$ ,  $\psi$  and  $\tilde{N}_i$  then admit expansions in powers of  $\mathcal{R}$ ,

$$\begin{aligned} \alpha &= \alpha_1 + \alpha_2 + \dots, \\ \psi &= \psi_1 + \psi_2 + \dots, \\ \tilde{N}_i &= \tilde{N}_i^{(1)} + \tilde{N}_i^{(2)} + \dots, \end{aligned} \quad (\text{A.150})$$

where, e.g.  $\alpha_n = \mathcal{O}(\mathcal{R}^n)$ . The constraint equations may then be set to zero order-by-order.

**Exercise 13 (First-order solution of constraint equations)** *Show that at first order (A.147) implies*

$$\alpha_1 = \frac{\dot{\mathcal{R}}}{H}, \quad \partial^2 \tilde{N}_i^{(1)} = 0. \quad (\text{A.151})$$

*With an appropriate choice of boundary conditions one may set  $\tilde{N}_i^{(1)} \equiv 0$ . Show that at first order (A.146) implies*

$$\psi_1 = -\frac{\mathcal{R}}{H} + \frac{a^2}{H}\epsilon \partial^{-2}\dot{\mathcal{R}}, \quad (\text{A.152})$$

*where  $\partial^{-2}$  is defined via  $\partial^{-2}(\partial^2\phi) = \phi$ .*

### B.2.3 The Free Field Action

Substituting the first-order solutions for  $N$  and  $N_i$  back into the action, one finds the following second order action [16]

$$S_2 = \frac{1}{2} \int d^4x a^3 \frac{\dot{\phi}^2}{H^2} \left[ \dot{\mathcal{R}}^2 - a^{-2} (\partial_i \mathcal{R})^2 \right]. \quad (\text{A.153})$$

**Exercise 14 (Second-order action)** *Confirm (A.153). Hint: use integration by parts and the background equations of motion.*

The quadratic action (A.153) for  $\mathcal{R}$  is the main result of this appendix and forms the basis for the quantization of cosmological perturbations in **Lecture 2**.

## C Slow-Roll Inflation in the Hamilton-Jacobi Approach

### C.1 Hamilton-Jacobi Formalism

The Hamilton-Jacobi approach treats the Hubble expansion rate  $H(\phi) = \mathcal{H}/a$  as the fundamental quantity, considered as a function of time. Consider

$$H_{,\phi} = \frac{H'}{\phi'} = \frac{-(\mathcal{H}^2 - \mathcal{H}')/a}{\phi'} = -\frac{\phi'}{2a}, \quad (\text{B.1})$$

where we used  $\mathcal{H}^2 - \mathcal{H}' = a^2(\rho + p)/2 = (\phi')^2/2$  and primes are derivatives with respect to conformal time. This gives the master equation

$$\boxed{\frac{d\phi}{dt} = \frac{\phi'}{a} = -2H_{,\phi}}. \quad (\text{B.2})$$

This allows us to rewrite the Friedmann equation

$$H^2 = \frac{1}{3} \left[ \frac{1}{2} \left( \frac{d\phi}{dt} \right)^2 + V(\phi) \right] \quad (\text{B.3})$$

in the following way

$$\boxed{[H_{,\phi}]^2 - \frac{3}{2}H^2 = -\frac{1}{2}V(\phi)}. \quad (\text{B.4})$$

Notice the following important consequence of the Hamilton-Jacobi equation (B.4): For any specified function  $H(\phi)$ , it produces a potential  $V(\phi)$  which admits the given  $H(\phi)$  as an exact inflationary solution. Integrating (B.2)

$$\int dt = -\frac{1}{2} \int \frac{d\phi}{H'(\phi)} \quad (\text{B.5})$$

relates  $\phi$  to proper time  $t$ . This enables us to obtain  $H(t)$ , which can be integrated to give  $a(t)$ . The Hamilton-Jacobi formalism can be used to generate infinitely many inflationary models with exactly known analytic solutions for the background expansion. However, here we are more concerned with the fact that it allows an elegant and intuitive definition of the slow-roll parameters.

### C.2 Hubble Slow-Roll Parameters

During slow-roll inflation the background spacetime is approximately de Sitter. Any deviation of the background equation of state

$$w = \frac{p}{\rho} = \frac{(\phi')^2/2a^2 - V}{(\phi')^2/2a^2 + V}$$

from the perfect de Sitter limit  $w = -1$  may be defined by the parameter

$$\boxed{\varepsilon \equiv \frac{3}{2}(1+w)}. \quad (\text{B.6})$$

We can express the Friedmann equations

$$\mathcal{H}^2 = \frac{1}{3}a^2\rho \quad (\text{B.7})$$

$$\mathcal{H}' = -\frac{1}{6}a^2(\rho + 3p) \quad (\text{B.8})$$

in terms of  $\varepsilon$

$$\mathcal{H}^2 = \frac{1}{3} \frac{(\phi')^2}{\varepsilon} \quad (\text{B.9})$$

$$\mathcal{H}' = \mathcal{H}^2(1 - \varepsilon). \quad (\text{B.10})$$

Hence,

$$\varepsilon = 1 - \frac{\mathcal{H}'}{\mathcal{H}^2} = \frac{d(H^{-1})}{dt}. \quad (\text{B.11})$$

Note that this can be interpreted as the rate of change of the Hubble parameter  $H$  with respect to the number of  $e$ -foldings  $dN = H dt = -\frac{1}{2} \frac{H(\phi)}{H_{,\phi}} d\phi$

$$\boxed{\varepsilon = -\frac{d[\ln H]}{dN} = 2 \left( \frac{H_{,\phi}}{H} \right)^2}. \quad (\text{B.12})$$

Analogously we define the second slow-roll parameter as the rate of change of  $H_{,\phi}$

$$\boxed{\eta = -\frac{d[\ln |H_{,\phi}|]}{dN} = 2 \frac{H_{,\phi\phi}}{H}}. \quad (\text{B.13})$$

Using the Hamilton-Jacobi master equation (B.2) this is also

$$\eta = \frac{d[\ln |\dot{\phi}|]}{dN}. \quad (\text{B.14})$$

### C.3 Slow-Roll Inflation

By definition, slow-roll corresponds to a regime where all dynamical characteristics of the universe, measured in physical (proper) units, change little over a single  $e$ -folding of expansion. This ensures that the primordial perturbations are generated with approximately equal power on all scales, leading to a scale-invariant perturbation spectrum.

Since  $\varepsilon$  and  $\eta$  characterize the rate of change of  $H$  and  $H_{,\phi}$  with  $e$ -foldings, slow-roll is naturally defined by

$$\varepsilon \ll 1, \quad |\eta| \ll 1. \quad (\text{B.15})$$

The first slow-roll condition implies

$$\varepsilon \ll 1 \quad \Rightarrow \quad \mathcal{H}^2 = \frac{1}{3} \frac{(\phi')^2}{\varepsilon} \gg (\phi')^2, \quad (\text{B.16})$$

so that the slow-roll limit of the first Friedmann equation is

$$\mathcal{H}^2 \approx \frac{1}{3} a^2 V. \quad (\text{B.17})$$

The second slow-roll condition implies

$$\eta = \frac{d[\ln |\dot{\phi}|]}{dN} = \frac{1}{H |\dot{\phi}|} \frac{d^2 \phi}{dt^2} \ll 1 \quad \Rightarrow \quad |\dot{\phi}| \ll H |\dot{\phi}|, \quad (\text{B.18})$$

so that the Klein-Gordon equation reduces to

$$\dot{\phi} \approx -\frac{a^2 V'}{3 \mathcal{H}}. \quad (\text{B.19})$$

In **Lecture 1** we defined a second set of common slow-roll parameters in terms of the local shape of the potential  $V(\phi)$

$$\epsilon_v \equiv \frac{1}{2} \left( \frac{V_{,\phi}}{V} \right)^2 \quad (\text{B.20})$$

$$\eta_v \equiv \frac{V_{,\phi\phi}}{V}. \quad (\text{B.21})$$

We note that  $\varepsilon(\phi_{\text{end}}) \equiv 1$  is an exact definition of the end of inflation, while  $\epsilon_v(\phi_{\text{end}}) = 1$  is only an approximation. In the slow-roll regime the following relations hold

$$\varepsilon \approx \epsilon_v \quad (\text{B.22})$$

$$\eta \approx \eta_v - \epsilon_v. \quad (\text{B.23})$$

## C.4 Inflationary Attractor Solution

We now show that the slow-roll condition,  $\varepsilon < 1$ , also implies that inflation is an attractor solution.

Let  $\bar{H}(\phi)$  be a solution of the Hamilton-Jacobi equation (B.4) (at this point we don't demand that this is an inflationary solution). Now consider a small perturbation  $\delta H(\phi)$ , *i.e.*

$$H(\phi) = \bar{H}(\phi) + \delta H(\phi). \quad (\text{B.24})$$

We linearize (B.4) to find

$$\bar{H}_{,\phi} \delta H_{,\phi} \approx \frac{3}{2} \bar{H} \delta H, \quad (\text{B.25})$$

or

$$\frac{d}{d\phi} (\ln \delta H) = \frac{3}{2} \frac{\bar{H}}{\bar{H}_{,\phi}}. \quad (\text{B.26})$$

This has the solution

$$\delta H(\phi) = \delta H(\phi_i) \exp \left[ \frac{3}{2} \int_{\phi_i}^{\phi} \frac{\bar{H}}{\bar{H}_{,\phi}} d\phi \right]. \quad (\text{B.27})$$

Recalling that

$$dN = -\frac{1}{2} \frac{H}{H_{,\phi}} d\phi = \frac{|\mathrm{d}\phi|}{\sqrt{2\varepsilon}} > 0, \quad (\text{B.28})$$

this may be written as

$$\delta H(\phi) = \delta H(\phi_i) \exp [-3(N - N_i)]. \quad (\text{B.29})$$

During inflation,  $\varepsilon < 1$ , the number of *e*-folds of expansion  $N$  rapidly becomes large and any perturbation to the inflationary solution  $\delta H$  gets diluted exponentially.  $H(\phi)$  then approaches  $\bar{H}(\phi)$ .

## References

- [1] A. H. Guth, Phys. Rev. **D23**, 347 (1981).
- [2] A. D. Linde, Phys. Lett. **B108**, 389 (1982).
- [3] A. Albrecht and P. J. Steinhardt, Phys. Rev. Lett. **48**, 1220 (1982).
- [4] CMBPol Study Team, D. Baumann *et al.*, (2008), 0811.3919.
- [5] S. Dodelson, *Modern Cosmology* (Amsterdam, Netherlands: Academic Press, 2003).
- [6] V. Mukhanov, *Physical Foundations of Cosmology* (Cambridge University Press, 2005).
- [7] S. Weinberg, *Cosmology* (Oxford University Press, 2008).
- [8] B. A. Bassett, S. Tsujikawa, and D. Wands, Rev. Mod. Phys. **78**, 537 (2006), astro-ph/0507632.
- [9] C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan, and L. Senatore, JHEP **03**, 014 (2008), 0709.0293.
- [10] S. Weinberg, Phys. Rev. **D77**, 123541 (2008), 0804.4291.
- [11] D. Baumann and L. McAllister, (2009), 0901.0265.
- [12] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, Phys. Rev. **D28**, 679 (1983).
- [13] D. Wands, K. A. Malik, D. H. Lyth, and A. R. Liddle, Phys. Rev. **D62**, 043527 (2000), astro-ph/0003278.
- [14] V. Acquaviva, N. Bartolo, S. Matarrese, and A. Riotto, Nucl. Phys. **B667**, 119 (2003), astro-ph/0209156.
- [15] J. M. Maldacena, JHEP **05**, 013 (2003), astro-ph/0210603.
- [16] J. M. Maldacena, JHEP **05**, 013 (2003), astro-ph/0210603.
- [17] N. D. Birrell and P. C. W. Davies, Cambridge, Uk: Univ. Pr. ( 1982) 340p.
- [18] D. H. Lyth, Phys. Rev. Lett. **78**, 1861 (1997), hep-ph/9606387.
- [19] N. Bartolo, E. Komatsu, S. Matarrese, and A. Riotto, Phys. Rept. **402**, 103 (2004), astro-ph/0406398.
- [20] M. Kamionkowski, A. Kosowsky, and A. Stebbins, Phys. Rev. **D55**, 7368 (1997), astro-ph/9611125.
- [21] M. Zaldarriaga and U. Seljak, Phys. Rev. **D55**, 1830 (1997), astro-ph/9609170.
- [22] E. Newman and R. Penrose, J. Math. Phys. **3**, 566 (1962).
- [23] WMAP, E. Komatsu *et al.*, (2008), 0803.0547.

- [24] D. N. Spergel and M. Zaldarriaga, Phys. Rev. Lett. **79**, 2180 (1997).
- [25] S. Dodelson, AIP Conf. Proc. **689**, 184 (2003), hep-ph/0309057.
- [26] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, Astrophys. J. **304**, 15 (1986).
- [27] D. J. Eisenstein and W. Hu, Astrophys. J. **496**, 605 (1998), astro-ph/9709112.
- [28] K. M. Smith, L. Senatore, and M. Zaldarriaga, (2009), 0901.2572.
- [29] S. Dimopoulos, S. Kachru, J. McGreevy, and J. G. Wacker, (2005), hep-th/0507205.
- [30] L. McAllister, E. Silverstein, and A. Westphal, (2008), 0808.0706.
- [31] E. Silverstein and A. Westphal, (2008), 0803.3085.
- [32] S. Kachru *et al.*, JCAP **0310**, 013 (2003), hep-th/0308055.
- [33] D. Baumann, A. Dymarsky, I. R. Klebanov, and L. McAllister, JCAP **0801**, 024 (2008), 0706.0360.
- [34] D. Baumann, A. Dymarsky, S. Kachru, I. R. Klebanov, and L. McAllister, (2008), 0808.2811.
- [35] J. P. Conlon and F. Quevedo, JHEP **01**, 146 (2006), hep-th/0509012.
- [36] J. J. Blanco-Pillado *et al.*, JHEP **09**, 002 (2006), hep-th/0603129.
- [37] D. Baumann and M. Zaldarriaga, (2009), 0901.0958.
- [38] D. Babich, P. Creminelli, and M. Zaldarriaga, JCAP **0408**, 009 (2004), astro-ph/0405356.
- [39] N. Bartolo, S. Matarrese, and A. Riotto, Phys. Rev. **D65**, 103505 (2002), hep-ph/0112261.
- [40] F. Bernardeau and J.-P. Uzan, Phys. Rev. **D66**, 103506 (2002), hep-ph/0207295.
- [41] F. Bernardeau and J.-P. Uzan, Phys. Rev. **D67**, 121301 (2003), astro-ph/0209330.
- [42] M. Sasaki, Prog. Theor. Phys. **120**, 159 (2008), 0805.0974.
- [43] A. Naruko and M. Sasaki, (2008), 0807.0180.
- [44] C. T. Byrnes, K.-Y. Choi, and L. M. H. Hall, (2008), 0807.1101.
- [45] C. T. Byrnes and D. Wands, Phys. Rev. **D74**, 043529 (2006), astro-ph/0605679.
- [46] D. Langlois, F. Vernizzi, and D. Wands, (2008), 0809.4646.
- [47] J. Valiviita, H. Assadullahi, and D. Wands, (2008), 0806.0623.
- [48] H. Assadullahi, J. Valiviita, and D. Wands, Phys. Rev. **D76**, 103003 (2007), 0708.0223.
- [49] J. Valiviita, M. Sasaki, and D. Wands, (2006), astro-ph/0610001.

- [50] F. Vernizzi and D. Wands, JCAP **0605**, 019 (2006), astro-ph/0603799.
- [51] L. E. Allen, S. Gupta, and D. Wands, JCAP **0601**, 006 (2006), astro-ph/0509719.
- [52] A. D. Linde and V. F. Mukhanov, Phys. Rev. **D56**, 535 (1997), astro-ph/9610219.
- [53] D. H. Lyth, C. Ungarelli, and D. Wands, Phys. Rev. **D67**, 023503 (2003), astro-ph/0208055.
- [54] G. Dvali, A. Gruzinov, and M. Zaldarriaga, Phys. Rev. **D69**, 023505 (2004), astro-ph/0303591.
- [55] L. Kofman, (2003), astro-ph/0303614.
- [56] P. Creminelli and L. Senatore, JCAP **0711**, 010 (2007), hep-th/0702165.
- [57] K. Koyama, S. Mizuno, F. Vernizzi, and D. Wands, JCAP **0711**, 024 (2007), 0708.4321.
- [58] E. I. Buchbinder, J. Khoury, and B. A. Ovrut, Phys. Rev. Lett. **100**, 171302 (2008), 0710.5172.
- [59] J.-L. Lehners and P. J. Steinhardt, Phys. Rev. **D77**, 063533 (2008), 0712.3779.
- [60] J.-L. Lehners and P. J. Steinhardt, Phys. Rev. **D78**, 023506 (2008), 0804.1293.
- [61] K. Koyama, S. Mizuno, and D. Wands, Class. Quant. Grav. **24**, 3919 (2007), 0704.1152.
- [62] K. Koyama and D. Wands, JCAP **0704**, 008 (2007), hep-th/0703040.
- [63] M. Alishahiha, E. Silverstein, and D. Tong, Phys. Rev. **D70**, 123505 (2004), hep-th/0404084.
- [64] N. Arkani-Hamed, P. Creminelli, S. Mukohyama, and M. Zaldarriaga, JCAP **0404**, 001 (2004), hep-th/0312100.
- [65] L. Senatore, Phys. Rev. **D71**, 043512 (2005), astro-ph/0406187.
- [66] P. Creminelli, JCAP **0310**, 003 (2003), astro-ph/0306122.
- [67] X. Chen, M.-x. Huang, S. Kachru, and G. Shiu, JCAP **0701**, 002 (2007), hep-th/0605045.
- [68] C. Gordon, D. Wands, B. A. Bassett, and R. Maartens, Phys. Rev. **D63**, 023506 (2001), astro-ph/0009131.
- [69] D. H. Lyth, C. Ungarelli, and D. Wands, Phys. Rev. **D67**, 023503 (2003), astro-ph/0208055.
- [70] S. Weinberg, Phys. Rev. **D70**, 083522 (2004), astro-ph/0405397.
- [71] J. R. Fergusson and E. P. S. Shellard, (2008), 0812.3413.
- [72] P. Creminelli and M. Zaldarriaga, JCAP **0410**, 006 (2004), astro-ph/0407059.
- [73] C. Cheung, A. L. Fitzpatrick, J. Kaplan, and L. Senatore, JCAP **0802**, 021 (2008), 0709.0295.
- [74] R. Holman and A. J. Tolley, JCAP **0805**, 001 (2008), 0710.1302.

- [75] E. J. Copeland, A. R. Liddle, D. H. Lyth, E. D. Stewart, and D. Wands, Phys. Rev. **D49**, 6410 (1994), astro-ph/9401011.
- [76] S. Kachru, R. Kallosh, A. Linde, and S. P. Trivedi, Phys. Rev. **D68**, 046005 (2003), hep-th/0301240.
- [77] D. H. Lyth, Phys. Rev. Lett. **78**, 1861 (1997), hep-ph/9606387.
- [78] A. D. Linde, Phys. Lett. **B129**, 177 (1983).
- [79] C. Vafa, (2005), hep-th/0509212.
- [80] H. Ooguri and C. Vafa, Nucl. Phys. **B766**, 21 (2007), hep-th/0605264.
- [81] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis, and R. Rattazzi, JHEP **10**, 014 (2006), hep-th/0602178.
- [82] M. R. Douglas and S. Kachru, Rev. Mod. Phys. **79**, 733 (2007), hep-th/0610102.
- [83] S. Kachru *et al.*, JCAP **0310**, 013 (2003), hep-th/0308055.
- [84] I. R. Klebanov and M. J. Strassler, JHEP **08**, 052 (2000), hep-th/0007191.
- [85] I. R. Klebanov and A. A. Tseytlin, Nucl. Phys. **B578**, 123 (2000), hep-th/0002159.
- [86] S. B. Giddings, S. Kachru, and J. Polchinski, Phys. Rev. **D66**, 106006 (2002), hep-th/0105097.
- [87] D. Baumann *et al.*, JHEP **11**, 031 (2006), hep-th/0607050.
- [88] D. Baumann, A. Dymarsky, I. R. Klebanov, L. McAllister, and P. J. Steinhardt, Phys. Rev. Lett. **99**, 141601 (2007), 0705.3837.
- [89] D. Baumann, A. Dymarsky, I. R. Klebanov, and L. McAllister, JCAP **0801**, 024 (2008), 0706.0360.
- [90] A. Krause and E. Pajer, JCAP **0807**, 023 (2008), 0705.4682.
- [91] D. Baumann, A. Dymarsky, S. Kachru, I. R. Klebanov, and L. McAllister, JHEP **03**, 093 (2009), 0808.2811.
- [92] K. Freese, J. A. Frieman, and A. V. Olinto, Phys. Rev. Lett. **65**, 3233 (1990).
- [93] T. Banks, M. Dine, P. J. Fox, and E. Gorbatov, JCAP **0306**, 001 (2003), hep-th/0303252.
- [94] P. Svrcek and E. Witten, JHEP **06**, 051 (2006), hep-th/0605206.
- [95] S. Dimopoulos, S. Kachru, J. McGreevy, and J. G. Wacker, JCAP **0808**, 003 (2008), hep-th/0507205.
- [96] R. Easther and L. McAllister, JCAP **0605**, 018 (2006), hep-th/0512102.

- [97] A. R. Liddle, A. Mazumdar, and F. E. Schunck, Phys. Rev. **D58**, 061301 (1998), astro-ph/9804177.
- [98] R. Kallosh, N. Sivanandam, and M. Soroush, Phys. Rev. **D77**, 043501 (2008), 0710.3429.
- [99] T. W. Grimm, Phys. Rev. **D77**, 126007 (2008), 0710.3883.
- [100] L. McAllister, E. Silverstein, and A. Westphal, (2008), 0808.0706.
- [101] M. P. Hertzberg, S. Kachru, W. Taylor, and M. Tegmark, JHEP **12**, 095 (2007), 0711.2512.
- [102] R. Flauger, S. Paban, D. Robbins, and T. Wrane, (2008), 0812.3886.
- [103] C. Caviezel *et al.*, JHEP **04**, 010 (2009), 0812.3551.
- [104] D. Baumann and L. McAllister, Phys. Rev. **D75**, 123508 (2007), hep-th/0610285.
- [105] E. Silverstein and D. Tong, Phys. Rev. **D70**, 103505 (2004), hep-th/0310221.
- [106] M. Alishahiha, E. Silverstein, and D. Tong, Phys. Rev. **D70**, 123505 (2004), hep-th/0404084.
- [107] S. Sarangi and S. H. H. Tye, Phys. Lett. **B536**, 185 (2002), hep-th/0204074.
- [108] E. J. Copeland, R. C. Myers, and J. Polchinski, JHEP **06**, 013 (2004), hep-th/0312067.
- [109] K. A. Malik and D. Wands, (2008), 0809.4944.
- [110] J. M. Bardeen, Phys. Rev. **D22**, 1882 (1980).
- [111] L. Verde, (2007), 0712.3028.
- [112] R. L. Arnowitt, S. Deser, and C. W. Misner, (1962), gr-qc/0405109.