2015

ΛCDM Cosmology + Chaotic Inflation

Peter A. Farago
Virginia Commonwealth University, faragopa@vcu.edu

ΛCDM Cosmology + Chaotic Inflation

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

By

Peter Alexander Farago

George Washington University, B.A. 1975

Director: Dr. Robert H. Gowdy

Associate Professor, Department of Physics

Virginia Commonwealth University

Richmond, Virginia

November, 2015

Acknowledgement

I would like to thank my wife, Lucretia, for her love, encouragement, and the support that I needed to undertake this study of the early universe. I wish to express my appreciation for my adviser, Dr. Robert H. Gowdy. I consider myself fortunate to have an adviser with his depth of knowledge in general relativity and cosmology. These are precisely the subjects that motivated me to study physics. Finally, I would like to thank my friend, James Slaughter, for his interest and support in my study of mathematics and physics.

Contents

List of Tables

List of Figures

Abstract

ΛCDM COSMOLOGY + CHAOTIC INFLATION

By Peter Alexander Farago

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2015

Major Director: Dr. Robert H. Gowdy, Associate Professor, Department of Physics

ΛCDM cosmology is described in terms of general relativity and the Robertson-Walker metric. The evolution of the observable universe, currently dominated by dark energy (Λ) and cold dark matter (CDM), is presented in terms of its thermal history. ΛCDM is extended to include an inflation epoch that accelerates the early expansion rate to near exponential levels. It is shown that inflation solves several problems in ΛCDM and produces perturbations in the metric that lead to the observed anisotropies in the Cosmic Microwave Background and the formation of large scale cosmological structures. Various theories of inflation are explored. Predictions of inflation theories are compared to observations published by the Planck Collaboration. The paper concludes with an examination of "α-attractor" theories of inflation based on a modified form of gravity.

Chapter 1

Introduction to Inflationary Cosmology

Inflationary Cosmology is an extension of Hot Big Bang cosmology or "ΛCDM" cosmology where Λ (Lambda) refers to dark energy and CDM refers to cold dark matter. Inflationary Cosmology replaces the dynamics of ΛCDM in the extremely early universe with a brief period of highly accelerated spatial expansion known as the inflation epoch. A model of inflation is a proposed effective field theory that applies during this epoch. Recent, detailed analysis of astrophysical measurements is consistent with a broad but narrowing range of models. Research in astrophysics and particle physics is ongoing to rule out models and to find a single model that is strongly favored by experiments.

The original motivation for inflation was to explain three observations that must otherwise be admitted as initial conditions in the ΛCDM: the apparent flatness of three dimensional space, the unexpected near uniform temperature of the cosmic microwave background (CMB), and the apparent absence of certain relic particles that are predicted in large numbers by grand unification theories [106].

A sufficient amount of inflation smoothes out initially curved space so that it appears flat. Running time backwards in ΛCDM finds that parts of the CMB separated by more than about 1.2° in the sky could not have interacted to reach uniform temperature. Replacing the early dynamics of the universe with inflation changes its causal structure. Inflation

places the precursors of the particles in all regions of the CMB inside a causally connected volume at the start of inflation, allowing the particles to achieve a high degree of thermal equilibrium and maintain near thermal equilibrium throughout inflation. Near thermal equilibrium conditions are transferred to ordinary particles formed at the onset of ΛCDM evolution that produces the observed near uniform temperature of the CMB. Finally, the number density of any relic particles created before or early in the inflation epoch would be diluted by inflation, making it practically impossible to observe their presence.

Additionally, inflation introduces quantum uncertainties into the universe that imprint the primordial curvature perturbation onto the spatial metric. The primordial curvature perturbation causes energy density perturbations that evolve into large-scale cosmological structures through gravitational collapse. Thus, inflation not only obviates the three initial conditions that motivated the theory, but also provides a basis for the origin of structure in the universe [77]. After inflation, the evolution of the universe is in good agreement with ΛCDM.

ΛCDM is the theory of the origin and evolution of the universe based on the assumption that the universe is the same everywhere and looks the same in all directions. Large-scale dynamics are based on general relativity. Thermal history describing the evolution of the content of the universe and the growth of cosmological structure is based on general relativity, statistical mechanics, particle physics, nuclear physics, and quantum field theory.

The components of the universe include ordinary matter and radiation that can be seen and measured, as well as two theoretical components that cannot be seen but which produce measurable effects. The first of these theoretical components is cold dark matter (CDM), an unidentified form of matter that is not composed of ordinary matter (neutrons, protons, and electrons collectively known in cosmology as baryonic matter) or neutrinos. CDM does not emit or absorb light and is measured to move at non-relativistic velocity. CDM is found in large halos around and within galaxies and galaxy clusters. It affects the motion of stars, other matter, and radiation within and around these structures in a manner consistent

with general relativity. The second theoretical component is dark energy ("Λ"), an unidentified form of energy with negative pressure that counteracts the effects of gravity and is responsible for the observed accelerating scale of space.

Cosmological measurements using multiple methods all find that baryonic matter accounts for less than 5% of the energy density of the universe and most of that is dust, not structures that can be seen, like stars and galaxies. Measurements also find that cold dark matter accounts for about 25% of the energy density, dark energy accounts for nearly 70%, and that the curvature of space (spacetime at a moment in time) is flat. Each component of energy density has pressure (possibly zero) related by the component's equation of state.

Most references to observations or measurements in this paper are based on "precision cosmology", the most accurate method currently available to estimate values of cosmological parameters such as energy densities, curvature, and the expansion rate. Precision cosmology is based on the analysis of temperature variations or anisotropies in the oldest observable light, the CMB. The CMB is primordial radiation formed in the very early universe that is still present throughout the universe today.

When the temperature of the universe was too hot to form atoms, primordial photons scattered and re-scattered continuously off free electrons and atomic nuclei. As temperatures dropped, atoms began to form. As ionization fractions dropped significantly, the probability that a primordial photon would scatter again was sharply reduced.

The "last scattering" temperature is the most probable temperature at which photons scattered for the last time. The last scattering temperature was approximately $3,000K$ which corresponds to about $370,000$ years [106]. At this temperature, most primordial photons began to travel freely without interacting with matter. Without interaction, the primordial photons no longer exchanged energy with matter and fell out of thermal equilibrium with matter. Equivalently, primordial photons decoupled from matter. The wavelengths of photons that decoupled at about the last scattering temperature have been stretched by the

expansion of space to a region in the microwave spectrum. Microwave radiation observed from the sky in this spectrum is the CMB.

Maps constructed by systematically scanning the sky for light in the CMB spectrum produce images that show the CMB as a spherical shell centered on the observer. This spherical shell, called the last scattering surface (LSS), is the oldest observable cosmological "structure". Although the last scattering events of photons occurred over many years, it is convenient to consider all last scattering events to have occurred at a single moment in time. The LSS may be viewed as all space at that moment in time. Figure 1.1 is an Aitoff projection of the LSS spherical shell based on data collected over a nine-year period by the Wilkinson Microwave Anisotropy Probe (WMAP). The color variations in the image are constructed to show the anisotropies in the CMB.

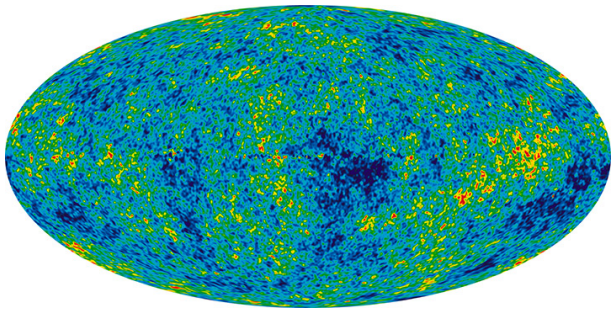


Figure 1.1: The Last Scattering Surface, WMAP9 image from the NASA Website

See §2.2 Λ*CDM Cosmological Parameters*, §2.4 *From Inflation to Photon Decoupling*, §B.15 and *Precision Cosmology* for details.

Flat ΛCDM, sometimes called the Concordance Model, is based on the observation of flat space and as few as six additional cosmological parameters estimated from astronomical observations. Computer codes based on this theoretical model and the observed value of cosmological parameters project the evolution of the universe from initial conditions at a small fraction of a second to about three minutes. The model predicts that atomic nuclei of light elements finished forming at about three minutes. The relative abundances of light element nuclei predicted by the model at three minutes closely match the relative abundances of light elements measured in the universe today [106, 94]. The model further projects the

evolution of the energy content and scale of the universe to the formation of LSS with temperature variations consistent with detailed observations of the CMB made by the Planck Explorer satellite. Additional computer codes developed based on the flat ΛCDM model, project the evolution of the spectrum of anisotropies in the CMB into cosmological structures (galaxies and galaxy clusters) that are statistically consistent with current observations [75, 57, 59, 58].

Flat ΛCDM is successful in matching well-established theories of physics to detailed observations of our universe. The theory traces the evolution of the universe from assumed initial conditions at or somewhat before electroweak symmetry breaking through to the current era. Extending the theory with a model of inflation pushes back the starting point for physical cosmology.

Models of inflation are defined by the potential energy of an unknown real valued scalar field called the *inflaton field* (some models have multiple scalar fields). According to these models, the energy of the universe at the beginning of the inflation epoch, at Planck time or shortly after, is dominated by the energy of the *inflaton field*. Ordinary fields such as matter and radiation fields also exist during this period (excepting short periods of "false vacuum" in some models) but have small energy compared to the *inflaton field*. Time-energy and space-momentum uncertainties or quantum uncertainties are present in all fields including the dominant *inflaton field*. The dominance of the *inflaton field* and its near thermal equilibrium allow the energy of the inflation epoch to be modeled as perfect fluid consisting only of the *inflaton field*. The gradient energy term vanishes and dynamics of the inflation epoch are those of a Friedmann universe. See §2.1 *Dynamics of a Friedmann Universe* for details. The *inflaton field's* equation of motion acquires a friction term based on the expansion rate and the metric used to define curvature in a Friedmann universe. The friction causes the value of the scalar field to change very slowly, like the motion of a "ball in a viscous fluid" [65]. Near exponential inflation commences when the potential energy term is reasonably flat and the potential energy is significantly greater than the kinetic

energy term. Inflation continues while the potential energy is nearly constant. Potential energy functions have at least one local minimum that ends inflation as the amplitude of the scalar field "slowly rolls downhill" [65] along the potential energy curve and settles into rapid oscillation about the bottom of a local potential well. The oscillating scalar field decays, creating a plasma of energetic particles (ordinary particles and dark matter). The universe becomes very hot. The decay processes (referred to in literature as "preheating" and "reheating") that produce these particles are complex and may be the source of the asymmetry between matter and antimatter particles required to produce non-zero baryon and lepton numbers. In order for a model to produce the right amount of inflation to develop the initial conditions for the ΛCDM model, the value of the *inflaton field* must evolve slowly and the potential energy must remain nearly flat over a wide range in the value of the scalar until the end of inflation. These two conditions are formalized as the Slow-Roll Approximation (SRA).

Chaotic inflation theories do not require thermal equilibrium as an assumed initial condition. These theories, defined by a reasonably flat potential energy term, begin at Planck time to describe a universe filled by a chaotically distributed scalar field and composed of causally disconnected Planck scale domains with arbitrary curvature. A domain will inflate if the energy density is uniform and the potential energy in the domain is significantly greater than kinetic and gradient energies. Based on the chaotic distribution of the field, the probability is near unity that the amplitude of the field is large enough in some domains so that the potential energy term approaches the classical density limit (Planck density). Chaotic inflation is sometimes called a large field theory because many forms of the theory require large values of the *inflaton field*. Since the total energy density cannot exceed Planck density, the potential energy density in such domains must be many times greater than the kinetic term, the gradient terms, and square of the scalar curvature. Since spatial inhomogeneity and anisotropy are derived from the gradients and the curvature tensor, classical Planck scale domains that satisfy the homogeneity required for inflation arise

naturally at or immediately after Planck time. Each such domain (a Planck scale inflation patch) inflates, creating "huge homogeneous islands out of initial chaos" (hence the origin of the name) [65]. Depending on the potential energy term and the initial value of the *inflaton field* in an inflation patch, the inflation epoch may be brief, long or endless.

In a simple model of chaotic inflation, the potential energy is proportional to the square of the value of the *inflaton field* (i.e., a "quadratic" model). Inflation commences with the initial value of the field displaced far enough from the local minimum so that potential energy dominates. In this model it is likely that inflation increases the size of the Planck scale domain by a factor of more than$10^{10^{12}}$, creating an enormous volume consisting of a hot, dense plasma of ordinary particles that dwarfs the present scale of the observable universe ($\sim 10^{28}\ cm$). Somewhere within this enormous volume lies a tiny, slightly perturbed region of space containing about 10^{90} particles.

This tiny region corresponds to the initial state of our "Hot Big Bang" which evolved into the observable universe in good agreement with ΛCDM. Based on this theory, the observable portion of the universe grew from a sub-Planck scale region a trillion orders of magnitude smaller than the Planck scale inflation patch. This result is derived in §4.2, *Harmonic Oscillator Theory*.

Classically, during the inflation epoch, the value of the scalar field may be thought of as rolling downhill. In the early part of the inflation, when the value of the field is very large, the quantum uncertainties in the field are also very large and space cannot be described in classical terms [74]. At smaller values, when the potential energy density has dropped below the level that would have been required for inflation to commence, there is a non-zero chance that the fluctuating value of the *inflaton field* attains high enough value to create the conditions for another instance of inflation. This leads to the "endless birth of ever newer regions of the universe" [74], comprising a fractal structure of "pocket" universes (eternal inflation). At lower values, the probability of a new instance of inflation

drops off. Our universe passed through to such a lower level and exited inflation. Smaller quantum uncertainties in the *inflaton field* continue to the end of inflation.

An effective model of inflation (the form of potential energy function and the value of its parameters) ends with oscillations of the scalar field that produce heat energy equivalent to the mass energy of the observable universe excluding dark energy. The heat energy decays into a plasma of ordinary particles that evolve into the elementary and composite particles observed today. Additionally, an effective model generates quantum uncertainties that evolve into energy density perturbations consistent with patterns of gravitational collapse that led to the formation of the observed anisotropies in the CMB and, later, to the formation of stars, galaxies, and galaxy clusters.

As an alternative to an initial, chaotically distributed scalar field, the universe may have emerged from the Planck era as a vacuum in a single Planck scale region of closed space. If the potential energy density of the *inflaton field* within this region acquires near Planck density (again a large field theory) through quantum uncertainties, then the region will inflate. The creation of the universe (or, as above, a fractal structure of "pocket" universes) from quantum uncertainties is called quantum creation. Quantum field theory based analyses have found that the probability of quantum creation is near unity with potential energy density close to Planck density [65].

Quantum creation can also occur in compact topologically flat or open Planck-scale domains (e.g., three-torus in the case of a flat domain). In this process, total energy density is on the order of Planck density, but is dominated by radiation with the potential energy density of the *inflaton field* significantly below Planck density (low scale or small field inflation). As the scale of the domain increases, thermal energy decreases. Inflation commences once the thermal energy density drops below potential energy density of the *inflaton field*. Although the universe created by this process has non-trivial topology, if inflation continues long enough, the scale of the universe at the end of inflation would be large enough so that

measurements of the observable universe would not detect any deviation from Euclidean space [65, 74].

All forms of inflation would have produced gravitational waves of some magnitude [65]. Large field chaotic inflation, such as the quadratic model, would have produced strong primordial gravitational waves with effects that may be observable as circular (B-mode) polarization signatures in the CMB. In 2014, the BICEP2 team reported observations of primordial B-modes to high significance [53]. This finding was withdrawn in 2015 after further review and collaboration with the Planck Explorer Science Team. The *2015 Joint Analysis of BICEP2/Keck Array and Planck Data* [1] concluded that observations of strong B-Mode polarization could be due to foreground dust contamination instead of primordial gravitational waves produced during inflation. The report states that, although high levels of B-mode polarizations have been confirmed, simulations using a dust-only model produce the observed results 8% of the time. Therefore, "this significance is too low to be interpreted as primordial B-modes" and additional data are required to "further constrain or detect" inflationary gravitational waves. The *2015 Joint Analysis* is largely consistent with the *Planck 2013 results. XXII. Constraints on inflation* [2] that favors the Starobinsky $R + R^2$ model described in Chapter 4.

The latest Planck Collaboration report on inflation, *Planck 2015 results. XX. Constraints on inflation* [89], favors the Starobinsky $R + R^2$ model and a new class of models called *alpha*-attractors. These new models are single field slow-roll models of chaotic inflation that can be constructed to predict weak signals of gravitational waves.

Chapter 2 presents the dynamics of a Friedmann universe, describes the ΛCDM cosmological parameters, presents a thermal history from inflation though the formation of the CMB, demonstrates how inflation sets the initial conditions otherwise imposed on ΛCDM by observations, and describes primordial density perturbations. Chapter 3 formalizes the slow-roll approximation (SRA) and describes the quantum origin of perturbations. Chapter 4 begins with a description of a chaotic model of inflation as a harmonic oscillator and

other single field monomial models of chaotic inflation. After demonstrating that these models are inconsistent with current observations, Chapter 4 describes the Starobinsky $R + R^2$ model and *alpha*-attractors models that are "supergravity" models derived from the monomial models of chaotic inflation. Chapter 4 concludes with a demonstration of *alpha*-attractor models of chaotic inflation that have been constructed since 2013 to predict the observations reported in the 2013 Planck reports and in the *2015 Joint Analysis of BICEP2/Keck Array and Planck Data*.

Chapter 2

Inflation in a Friedmann Universe

Inflationary Cosmology is a model of the universe with dynamic scale determined by general relativity and the assumptions that, at each instant of time, space looks the same in all directions (isotropy) and has the same composition everywhere (homogeneous). Although this is clearly not a correct description of our universe on human scale, observations of very large regions of space show that these assumptions are approximately true on cosmological scales. Under these assumptions, space is maximally symmetric and the energy-momentum of the universe can be modeled as a perfect fluid. This model is called a Friedmann universe.

2.1 Dynamics of a Friedmann Universe

Using units commonly applied in cosmology,[1] the Einstein field equation is

$$R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} = 8\,\pi\, G\, T_{\mu\nu} \tag{2.1.1}$$

and the stress-energy tensor for a perfect fluid at rest is

$$T_{\mu\nu} = (\rho + p)\, U_\mu\, U_\nu + p\, g_{\mu\nu} \tag{2.1.2}$$

where ρ represents energy density and p represents pressure. The terms U_μ and U_ν are the four-velocity dual vectors, $(-1, 0, 0, 0)$.[2] See §B.1 *Einstein's Equation for Gravitation* for a derivation of the Einstein field equation in Eq. (B.1.6) and Eq. (B.3.12) for additional information about the stress-energy tensor.

The solution to the Einstein field equation for a Friedmann universe is a Robertson-Walker metric. The line element for a common form[3] of the Robertson-Walker metric is

$$ds^2 = -dt^2 + a^2(t)R_0^2\left(\frac{dr^2}{1 - k\,r^2} + r^2 d\Omega^2\right) \tag{2.1.3}$$

where $a(t)$ is a dynamic dimensionless scale factor normalized so that the factor at the current time is unity ($a(t_0) = a_0 = 1$), R_0 is an expression of the current physical size of the universe, r is a dimensionless radial coordinate, $d\Omega$ is the unit 2-sphere metric, and k has discrete values $1, 0,$ or -1 depending on whether the global structure of three-dimensional space is positively curved, flat, or negatively curved. The coordinates are called co-moving coordinates.

The spatial component of the metric measures the physical distance at time t between objects moving along with the expansion of space and no other motion. These objects are called co-moving objects. An example of a co-moving object in the universe today is a galaxy cluster that is sufficiently distant from any other object so that it is not gravitationally bound to any other object. The co-moving distance between co-moving objects is the physical distance today when, by construction, $a_0 = 1$. The subscript 0 is commonly used in cosmology to represent a value in the current era.

The form of the metric can be deduced from homogeneity and isotropy. Einstein's equation determines the dynamics of expansion expressed by the scale factor. See §A.3.7 *The Robertson-Walker Metric* for details.

The following two equations describe the dynamics of spatial expansion. They are derived from the field equations and are known as the Friedmann equation and the acceler-

ation equation:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2 R_0^2} \tag{2.1.4}$$

and

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p)\ . \tag{2.1.5}$$

See §B.3.2 *Friedmann Equations* for a derivation.

The local conservation of energy-momentum in curved spacetime requires that for each value of ν and stress-energy tensor, $T^{\mu}{}_{\nu}$,

$$\nabla_\mu T^{\mu}{}_{\nu} = 0 \tag{2.1.6}$$

where ∇_μ is the covariant derivative. See §B.2 *Local Energy-Momentum Conservation in Spacetime* for additional information. Applying the local conservation of energy-momentum to the stress-energy tensor for a perfect fluid with $\nu = 0$ yields the "fluid equation",

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

where H is the Hubble spatial expansion parameter defined as $H = \frac{\dot{a}}{a}$ (the fluid equation is derived in Eq. (B.3.16) in an equivalent form). The fluid equation expresses the conservation of energy-momentum in a Friedmann universe in terms of the expansion of space. Spatial expansion may be described as a function of the energy and pressure of the constituents of the universe.

The relationship between pressure, p, and energy density, ρ, for a particular energy constituent is often specified by the equation of state, $p = w\rho$, where w is a conserved quantity.

The fluid equation can be used to derive the following expression for the time evolution of an energy density based on the scale factor and the current density (see derivation of Eq. (B.3.18)):

$$\rho_\alpha(t) = \rho_{\alpha 0} a(t)^{-n_\alpha}$$

where $n_\alpha = 3\,(1 + w_\alpha)$ is the scale exponent. Additionally, during a period when a specific form of energy dominates energy density of the universe, the time evolution of the scale factor is proportional to a power of time (see Eq. (B.3.24)). The time evolution of the Hubble parameter during such a period follows immediately.

Substituting the definition of the Hubble parameter into Eq. (2.1.4) and multiplying through by H^{-2} yields an alternative form of the Friedmann equation:

$$\frac{8\,\pi\,G}{3\,H^2}\rho - \frac{k}{a^2 H^2 R_0^2} = 1\,. \tag{2.1.8}$$

If the total energy density, ρ, is equal to

$$\rho_{critical} = \frac{3\,H^2}{8\pi\,G}\,, \tag{2.1.9}$$

then $k = 0$ and space is flat. The density parameter for energy source α is defined as

$$\Omega_\alpha = \frac{\rho_\alpha}{\rho_{critical}} \tag{2.1.10}$$

and the total density parameter is

$$\Omega = \sum_\alpha \Omega_\alpha = \frac{\rho}{\rho_{critical}}\,. \tag{2.1.11}$$

By substituting Ω for the first term in Eq. (2.1.8), the Friedmann equation may be written

$$1 - \Omega = -\frac{k}{a^2 H^2 R_0^2}\,. \tag{2.1.12}$$

The deviation of the total density parameter, Ω, from unity is called the curvature density parameter, Ω_k. Therefore, from Eq. (2.1.12),

$$\begin{aligned}\Omega_k &= 1 - \Omega \\ &= -\frac{k}{a^2 H^2 R_0^2}\,.\end{aligned} \tag{2.1.13}$$

The corresponding curvature energy density is $\rho_k = -\frac{3k}{8\pi G a^2 R_0^2}$. Including the curvature energy density, ρ_k, into total density, ρ, produces yet another restatement of the Friedmann equation and a concise expression for the Hubble parameter:

$$H = \sqrt{\frac{8\pi G\rho}{3}}\,. \tag{2.1.14}$$

The dimension of the Hubble parameter is the inverse of time since $[H] = [\dot{a}]/[a] = (L/T)/L = 1/T$. Therefore, the inverse of the Hubble parameter, H^{-1}, is called Hubble time, and $H^{-1} = c\,H^{-1}$ is called Hubble length or the Hubble horizon. It is called a horizon because $H_t^{-1} = c\,H_t^{-1}$. Hubble horizon at time t is the distance that light can travel in one unit of Hubble time, H_t^{-1}. The Hubble horizon is on the order of the physical scale of the observable universe and Hubble time is on the order of the age of the universe. Therefore, the Hubble horizon is often referred to as the characteristic scale of the universe.

A spherical region in space with co-moving radius R (i.e., radius measured in co-moving coordinates set in the current era) has physical radius $R\,a(t)$ at time t. The region is inside or outside the Hubble horizon, depending on whether $R\,a(t)/H_t^{-1} = R\,a/H^{-1} = R\,a\,H = R\,a\dot{a}/a = R\,\dot{a}$ is less than or greater than unity. If a is accelerating, then $\dot{a}$, and hence $R\dot{a}$, is increasing. Therefore, during periods of accelerating scale, regions of constant

co-moving size are exiting the Hubble horizon. Conversely, during periods of decelerating scale, regions of constant co-moving size are entering the Hubble horizon.

2.2 ΛCDM Cosmological Parameters

ΛCDM Cosmology is based on a set of primary cosmological parameters determined by observations. The Planck Collaboration reports the following six primary parameters for the base ΛCDM model [87, 88]:

$$\{\,\Omega_{b0}h^2, \Omega_{c0}h^2, 100\theta_{MC}, \tau, \ln\left(10^{10}A_s\right), n_s\,\}\,. \tag{2.2.1}$$

The background temperature of the CMB, $T_0 = 2.7255K$, and the flatness condition, $\Omega_{k0} = 0$, are also used in their analyses of the base ΛCDM model but are not listed as part of the primary parameter set. Other parameters are derived from the primary set and additional parameters are added when considering extensions to ΛCDM.

The physical density parameters for baryonic matter and cold dark matter are given by $\Omega_{b0}h^2$ and $\Omega_{c0}h^2$ where $h = H_0/100$ and H_0 is the estimated value of the Hubble parameter in the current era.

The parameter $100\theta_{MC}$ is one hundred times the co-moving sound horizon usually written as θ_*. The co-moving sound horizon is a measure on the LSS of the maximum co-moving distance that sound waves could have traveled from the formation of the photon-baryon fluid during the first several minutes of cosmic time until photon decoupling at about $370,000$ years. The speed of sound through the photon-baryon fluid [106] is a function of the ratio of baryon energy density to the photon energy density, and is typically $v_s = c/\sqrt{3}$. The subscript MC refers to the Cosmological Monte-Carlo computer codes that are used together with the *Code for Anisotropies in the Microwave Background* (CAMB Codes) to develop the value of the parameter.

The parameter, τ, is the optical depth due to re-ionization caused by initial star formation. It equals $\tau(z_{re})$ where z_{re} is the redshift at which half of the matter in the universe was re-ionized by star formation. The probability that CMB light traveling un-scattered at redshift z is observed in the current era without re-scattering is $e^{-\tau(z)}$. Thus, $\tau = \tau(z_{re})$ is based on the assumption that re-ionization occurred suddenly at z_{re}. See §B.6 *Thomson Scattering and Optical Depth* for additional information.

The Planck Collaboration determines τ in connection with another ΛCDM cosmological parameter, A_s, the primordial curvature perturbation amplitude. The perturbation amplitude is used together with the scalar spectrum power-law index, n_s, to model energy density perturbations at various cosmological scales. The parameters, θ_*, A_s, and n_s, are described further in §2.7 *Primordial Energy Density Perturbations*. There are no tensor mode perturbations in the base ΛCDM model.

The radiation density parameter, Ω_{r0}, can be derived from T_0 using statistical mechanics and particle theory (Eq. (B.13.9)). The density parameters for ordinary matter and dark matter are often combined into a single density parameter for matter, Ω_{m0}. If all sources of energy are accounted for, then the density parameter for dark energy must be $\Omega_{\Lambda 0} = 1 - \Omega_{r0} - \Omega_{m0} - \Omega_{k0}$.

In *Planck 2013 results. XVI. Cosmological parameters* [87], the Planck Collaboration reported the value $T_0 = 2.7255K$, and the following best-fit parameters: $\Omega_{k0} = 0$, $\Omega_{m0} = 0.3183$, and $H_0 = 67.04\frac{km}{Mpc\ sec}$. Based on T_0, H_0 and Eq. (B.13.9), $\Omega_{r0} \simeq 10^{-4}$. Therefore, $\Omega_{\Lambda 0} = 0.6817$. Table 2.1 lists these energy sources together with the symbol, current best fit density parameter, equation of state parameter, scale exponent, energy density function, and, during a period when energy density of the universe is dominated by a single source, the scale factor up to a constant of proportionality, and the Hubble parameter. Details on cosmological parameters are available in *Planck 2013 results. XVI. Cosmological parameters* [87] and *Planck 2015 results. XIII. Cosmological parameters*

Table 2.1: Energy Table

Source	Symbol	Ω_{i0}	w_i	n_i	Energy Density	Scale	Hubble
Radiation	r	10^{-4}	$\frac{1}{3}$	4	$\rho_r = \rho_{r0}\, a^{-4}$	$a \propto t^{\frac{1}{2}}$	$\frac{1}{2t}$
Matter	m	0.3183	0	3	$\rho_m = \rho_{m0}\, a^{-3}$	$a \propto t^{\frac{2}{3}}$	$\frac{2}{3t}$
Curvature	k	0	$-\frac{1}{3}$	2	$\rho_k = \rho_{k0}\, a^{-2}$	$a \propto t$	$\frac{1}{t}$
Dark Energy	Λ	0.6817	-1	0	$\rho_\Lambda = \rho_{\Lambda 0}\, a^{0}$	$a \propto e^{H_0 t}$	H_0

[88]. The complete set of reports published by the Planck Collaboration is available at http://www.cosmos.esa.int/web/planck/publications.

2.3 Calculating Scale, Time, and Distance

Wavelengths of light emitted by distant objects are redshifted (i.e., stretched) by the expansion of space during the period from emission to observation. Redshift is represented by the change in the wavelength, divided by the wavelength at emission, $z = (\lambda_o - \lambda_e)/\lambda_e$. The normalized scale factor at emission is related to redshift as $a(t_e) = \frac{1}{z+1}$. Light from nearby objects has redshift $z = 0$, corresponding to $a_0 = \frac{1}{z+1} = 1$. Light emitted at earlier times has greater redshift. The wavelengths of light emitted by Type 1a supernovae are well understood. The redshift from these objects may be determined by comparing the emitted wavelengths to the observed wavelengths. Using Type 1a supernovae as standard candles, the scale of the universe may be determined at the time light was emitted by distant galaxies.

The definition of the Hubble parameter, the current value of Hubble time (H_0^{-1}), and the density parameters ($\Omega_{r0}, \Omega_{m0}, \Omega_{k0}, \Omega_{\Lambda 0}$) yield an integral expression for cosmological time in terms of scale (Eq. (2.3.4) or, with a change in variable, in terms of redshift. From $\frac{\dot{a}}{a} = H$, it follows that $\frac{da}{dt}\frac{1}{a} = H$, leading to

$$dt = \frac{1}{aH} da \tag{2.3.1}$$

and

$$\Delta t = \int_{a_1}^{a_2} \frac{1}{aH} da \,. \tag{2.3.2}$$

Since $H(t) = \sqrt{\frac{8\pi G\rho(t)}{3}}$, $\rho_{critical\,0} = \frac{3H_0^2}{8\pi G}$, and from Table 2.1

$$\rho(a) = \rho_{r0}a^{-4} + \rho_{m0}a^{-3} + \rho_{k0}a^{-2} + \rho_{\Lambda 0}a^{0} \,, \tag{2.3.3}$$

it follows that

$$\begin{aligned}
\Delta t &= \int_{a_1}^{a_2} \frac{1}{a\sqrt{\frac{8\pi G}{3}}\sqrt{\rho_{r0}a^{-4} + \rho_{m0}a^{-3} + \rho_{k0}a^{-2} + \rho_{\Lambda 0}a^{0}}} da \\
&= \frac{H_0^{-1}}{H_0^{-1}} \int_{a_1}^{a_2} \frac{1}{a\sqrt{\frac{8\pi G}{3}}\sqrt{\rho_{r0}a^{-4} + \rho_{m0}a^{-3} + \rho_{k0}a^{-2} + \rho_{\Lambda 0}a^{0}}} da \\
&= H_0^{-1} \int_{a_1}^{a_2} \frac{1}{\sqrt{a^2}\sqrt{\frac{8\pi G}{3H_0^2}}\sqrt{\rho_{r0}a^{-4} + \rho_{m0}a^{-3} + \rho_{k0}a^{-2} + \rho_{\Lambda 0}a^{0}}} da \\
&= H_0^{-1} \int_{a_1}^{a_2} \frac{1}{\sqrt{\frac{a^2}{\rho_{critical\,0}}}\sqrt{\rho_{r0}a^{-4} + \rho_{m0}a^{-3} + \rho_{k0}a^{-2} + \rho_{\Lambda 0}a^{0}}} da \\
&= H_0^{-1} \int_{a_1}^{a_2} \frac{1}{\sqrt{\Omega_{r0}a^{-2} + \Omega_{m0}a^{-1} + \Omega_{k0}a^{0} + \Omega_{\Lambda 0}a^{2}}} da \,.
\end{aligned} \tag{2.3.4}$$

Since redshift is related to scale as $a = \frac{1}{z+1}$ it follows that

$$da = -\frac{1}{(z+1)^2} dz \,. \tag{2.3.5}$$

Therefore, by change of variable

$$\Delta t = H_0^{-1} \int_{z_2}^{z_1} \frac{dz'}{D(z')} \tag{2.3.6}$$

where $D(z') = \sqrt{\Omega_{r0}(1+z')^6 + \Omega_{m0}(1+z')^5 + \Omega_{\Lambda 0}(1+z')^2 + \Omega_{k0}(1+z')^4}$ (Eq. (2.3.6)). Integrating over scale from zero to one or, equivalently, over redshift from zero to infinity yields the age of the universe today.

The metric given in Eq. (A.3.15) along a null geodesic where coordinates have been rotated so that the path is in the radial direction only, reduces to

$$0 = d\tau^2 = dt^2 - a^2 R_0^2 d\chi^2 \ .$$

From this expression, it follows that

$$d\chi = \frac{dt}{aR_0} \ . \tag{2.3.7}$$

Therefore, if a co-moving object along the path emits light at time t_1 and a second co-moving object along the path receives the light at later time t_2, then the co-moving radial coordinate distance between the two objects is

$$\chi = R_0^{-1} \int_{t_1}^{t_2} \frac{dt'}{a(t')} \ . \tag{2.3.8}$$

A hypersurface, designated by Σ_t, is spacetime at a fixed moment in time t. Therefore, space or hypersurface in the current era is Σ_0. Since co-moving radial distance is $D_R = R_0\chi$, it follows that co-moving radial distance between the two objects, instantaneously measured in Σ_0, is

$$D_R = R_0\chi = \int_{t_1}^{t_2} \frac{dt'}{a(t')} \ . \tag{2.3.9}$$

Set t_1 to the time at the end of inflation, $t_{end} \simeq 0$, and t_2 to t. If a co-moving object along the path emits light at time t_{end} and a second co-moving object along the path receives the light at later time t, then the co-moving radial distance between the two objects measured

in Σ_0 is defined as the co-moving horizon, D_H, which is equivalent to "conformal time", η:

$$D_H(t) = \eta(t) = \int_0^t \frac{dt'}{a(t')}\,. \tag{2.3.10}$$

Substituting $\frac{1}{aH}da$ for dt (see Eq.(2.3.1)) into Eq.(2.3.9) and applying the same procedure used to derive Eq. (2.3.4) yields

$$\begin{aligned} D_R(a) &= \int_{a_1}^{a_2} \frac{da'}{a'^2\,H} \\ &= H_0^{-1} \int_{a_1}^{a_2} \frac{da'}{\sqrt{\Omega_{r0} + \Omega_{m0}\,a' + \Omega_{\Lambda 0}\,a'^4 + \Omega_{k0}\,a'^2}} \end{aligned} \tag{2.3.11}$$

and

$$D_H(a) = \eta(a) = H_0^{-1} \int_0^{a} \frac{da'}{\sqrt{\Omega_{r0} + \Omega_{m0}\,a' + \Omega_{\Lambda 0}\,a'^4 + \Omega_{k0}\,a'^2}}\,. \tag{2.3.12}$$

Change of variable from a to z yields

$$D_R = H_0^{-1} \int_{z_2}^{z_1} \frac{dz'}{E(z')} \tag{2.3.13}$$

and

$$D_H(z) = \eta(z) = H_0^{-1} \int_{z}^{\infty} \frac{dz'}{E(z')} \tag{2.3.14}$$

where, following the notation introduced by Peebles 1993 [83],

$$E(z') = \sqrt{\Omega_{r0}\,(1+z')^4 + \Omega_{m0}\,(1+z')^3 + \Omega_{\Lambda 0} + \Omega_{k0}\,(1+z')^2}\,. \tag{2.3.15}$$

The co-moving angular diameter distance D_A is equal to D_R in flat space. Therefore, for purposes of post-inflation calculations,

$$D_A = D_R \,. \tag{2.3.16}$$

The physical distance, d, measured in $\Sigma_t \equiv \Sigma_a \equiv \Sigma_z$ instead of Σ_0, is the co-moving distance multiplied by the scale factor $a \equiv \frac{1}{z+1}$:

$$d = a(t)D \equiv a\,D \equiv \frac{1}{z+1}D \,. \tag{2.3.17}$$

The "radius" of the universe, called the particle horizon, is $D_H(1)$ expressed in terms of scale (Eq. (2.3.12) and $D_H(0)$ expressed in terms of redshift (Eq. (2.3.14)).

The dimensions of H_0 are $[H_0] = \frac{L}{T \cdot L} \equiv T^{-1}$. H_0 is usually represented as $H_0 \frac{km}{sec \cdot Mpc}$. The values, $1 Mpc \approx 3.086 \times 10^{24}\, cm$ and $1\, year \approx 3.156 \times 10^7\, sec$, may be used to represent H_0^{-1} in years:

$$t_{Hubble} = {H_0}^{-1} = 9.7782\, h^{-1} \times 10^9\, years \,. \tag{2.3.18}$$

Hubble length is the distance traveled by light in one unit of Hubble time. Hence, one unit of Hubble length (or Hubble radius or Hubble horizon) is

$$d_{Hubble} = {H_0}^{-1} = 9.7782\, h^{-1} \times 10^9\, light\, years \,. \tag{2.3.19}$$

With appropriate conversion factors, the Hubble time and Hubble length may be converted to any system of units.

Applying Eq. (2.3.6) and the parameters in Table 2.1 yield the age of the universe today as

$$age = H_0^{-1} \int_0^{\infty} \frac{dz'}{D(z')} = 13.83 \text{ billion years}\,.$$

Applying Eq. (2.3.14) and the parameters in Table 2.1 yield the particle horizon of the universe today as

$$\begin{aligned} x_{ph} &= H_0^{-1} \int_0^{\infty} \frac{dz'}{E(z')} \\ &\approx 3.23 \text{ Hubble lengths} \\ &\approx 47.06 \text{ billion light years} \\ &\approx 14,428 \text{ Mpc} \\ &\approx 4.45 \times 10^{28} \text{ cm}\,. \end{aligned} \tag{2.3.20}$$

The redshift expressions, with one of the co-moving objects being a present day observer on Earth or a satellite in orbit around the Earth or the Sun,[4] are commonly used to determine distances of galaxies where the redshift of light has been calculated based on the energy of observed light from supernovae within the galaxy. The redshift of the most distant galaxy observed to date is about 7.5 [27].

The redshifts corresponding to several pivotal events in cosmological history can be determined using statistical mechanics and particle physics. For example, the Planck Collaboration reports that the redshift of light from the LSS is $z_* = 1090.48$ [87]. The wavelength evolves as a and the energy of a photon is inversely proportional to its wavelength. Hence, the energy of each photon evolves as $1/a$. Since the temperature of a photon gas is proportional to the average energy of the photons (Eq. (B.5.1)), it follows that $T_\gamma \propto 1/a$, where T_γ is the average temperature of photons in the CMB. The average temperature of the CMB today is $T_{\gamma 0} \simeq 2.7255K$. As cited above, the redshift of light streaming from the

Table 2.2: Approximate Thermal History

Event	**Energy**	**Temperature**	**Time**
End of Planck era	10^{19} GeV	10^{32}K	10^{-43} sec
Inflation begins (e.g.)	$\approx 10^{19}$ GeV	$\approx 10^{32}$ K	$\approx 10^{-43}$ sec
Strong-electroweak unification	$\approx 10^{16}$ GeV	$\approx 10^{29}$ K	$\approx 10^{-38}$ sec
Inflation ends (e.g.)	$\approx 10^{14}$ GeV	$\approx 10^{27}$ K	$\approx 10^{-35}$ sec
Electro-weak unification	100 GeV	10^{15}K	10^{-11} sec
Nucleon formation begins	1 GeV	10^{13}K	10^{-6} sec
Neutrino decoupling begins & Nucleon/anti-nucleon annihilation ends	100 KeV	10^{10}K	1 sec
Electron/positron annihilation ends ($t_\dagger$)	250 KeV	3×10^9K	15 sec
Nucleosynthesis begins & neutrino decoupling ends	86 KeV	10^9K	3 min
Nucleosynthesis ends	26 KeV	3×10^8K	30 min
Matter-radiation equality	26 KeV	3×10^8K	9×10^4 yrs
Hydrogen atom formation begins	0.35 eV	4,090 K	2×10^5 yrs
Light atom formation ends, photon decoupling "complete", & LSS formed	0.25 eV	3.000 K	3.7×10^5 yrs

LSS is $z_* = 1090.48$. Therefore, the average temperature of photons at the time the LSS was formed is $T_{\gamma *} \simeq 3,000K$. By Eq. (2.3.6), the time corresponding to z_* is $t_* \simeq 370,000$ years. The LSS may be regarded as spacetime frozen in time at t_*, i.e., a spacelike hypersurface. The formation of the LSS is described below.

2.4 From Inflation to Photon Decoupling

In ΛCDM cosmology, cosmological history is traced back to Planck time at about 10^{-43} seconds using ΛCDM cosmological parameters. Inflationary cosmology replaces a brief period at the beginning of ΛCDM with an inflation epoch. The inflation epoch is a period of highly accelerated expansion in the scale of the universe. Chapter 1 describes how a period of rapid expansion can establish near uniform temperature in the early universe, reduce the curvature of space to near flatness, and reduce the number density of magnetic monopoles to undetectable levels. In §2.5 *The Required Quantity of Inflation* it will be shown that the minimum amount of inflation required to set these conditions is at least 26

orders of magnitude, corresponding to 60 e-foldings. Therefore, if a_i was the value of the scale factor at the beginning of inflation and a_e is the value at the end of inflation, then $a_e/a_i > e^{60}$. The total expansion during inflation could be much more.

In most models of inflation, it is assumed that a small classical Friedmann universe, called an inflation domain or patch, emerges at Planck time and the inflation epoch begins in this patch almost immediately. The energy density is large and is dominated by the potential energy density of a scalar field called an *inflaton field* associated with an unknown spin-0 particle, the *inflaton*. The potential energy density remains nearly constant during a brief period of inflation, producing a greatly expanded volume filled with the potential energy of the scalar field.

The potential energy begins to decline slowly toward a minimum value as the potential energy converts to thermal (kinetic) energy. Inflation ends when the potential energy no longer dominates the energy density. The thermal energy is transferred to the vacuum field through preheating followed by reheating to produce a volume filled with hot, dense plasma of highly relativistic ordinary particles and dark matter. The ordinary particles are principally photons, electrons, positrons, neutrinos, antineutrinos, quarks, and antiquarks, all highly relativistic in a continuous cycle of formation and annihilation. The end of reheating marks the beginning of the radiation era. Depending on the model of inflation, this volume could be enormous or less than one centimeter in scale. A region within this volume (possibly a very small part) evolves into our observable universe.

Assuming that the current size of the observable universe is R_0, then the physical size observable universe at the end of inflation is $a_e R_0$. The scale factor at the beginning of the radiation era is designated as a_1; therefore the physical size of the observable universe at the beginning of the radiation era is $a_1 R_0$. By this point, the observable universe was nearly homogeneous and isotropic but bore the imprints of quantum uncertainties that developed during inflation. A background of constant but very weak dark energy density was either present during inflation or possibly evolved from the *inflaton field*. This tiny precursor to

the observable universe is a sufficient "initial state" for the "Big Bang" and subsequent evolution into the state of our observable universe based on ΛCDM. Hereafter, "plasma" refers to the initial plasma of ordinary particles established at the beginning of the radiation era in the region $a_1\, R_0$ and the evolution of the initial plasma.

Many models of inflation are constructed so that the temperature of the plasma at the time photons were produced during reheating is below the grand unification energy scale of about $10^{16} GeV$, corresponding to about $10^{29} K$. This may or may not be true, but making this assumption avoids the unknown conditions at grand unification (that are far beyond experimental reach of existing or feasible particle colliders) and eliminates the production of monopoles during and after reheating. For purposes of the present discussion, it will be assumed that the temperature of the plasma produced by reheating is approximately $10^{27} K$. By Eq. (B.13.12), this corresponds to 10^{-35} seconds. By this assumption, reheating ends and the radiation era begins at about 10^{-35} seconds. After this point, the theoretical evolution is based on ΛCDM. The quantum uncertainties produced by inflation play a critical role in modeling the formation of the observed anisotropies in the CMB and the subsequent formation of large scale cosmological structures.

After the beginning of the radiation era, space continues to expand at a much lower rate. Expansion cools the plasma. When the thermal energy of plasma drops to the unification energy scale of the weak and electromagnetic forces ($\sim 100\, GeV$ which corresponds to $10^{15} K$ [74]), symmetry between the weak and electromagnetic forces is broken by the Higgs field. The interaction of W^0 and B^0 bosons produces Z^0 bosons and photons. All weak force bosons (the $W^{\pm}$ bosons that are also mediators of the electroweak force of the newly formed Z^0 bosons) subsequently acquire mass by interaction with the Higgs field. Aspects of this process have been experimentally confirmed by the production of the $W^{\pm}$ and Z^0 bosons and, most recently, the probable production of Higgs boson particles at the Large Hadron Collider. Electroweak symmetry-breaking marks the initial presence of all four separate forces that are measured in the current era. All particles that interact with

the Higgs field have acquired mass. All boson and fermions in the standard model except for the top quark and the Higgs boson have threshold temperatures below $10^{15}K$. It is assumed that dark matter fell out of thermal equilibrium very early in cosmological history because of the negligible interaction between dark matter and photons. Ordinary elementary particles remain in thermal equilibrium through particle-antiparticle annihilation, pair production, and Compton scattering.

The plasma continues to cool under expansion. At about $10^{13}K$, up and down quark pairs quickly form stable composite particle/antiparticle pairs. These stable composite particles are the nucleons: protons and neutrons.[5] Nucleons and anti nucleons annihilate on interaction, releasing photons back into the plasma. An asymmetry between quark and antiquark numbers leaves a residual number of nucleons. The source of this asymmetry is called baryonsynthesis (or baryogenesis). The asymmetry may have occurred during the original production of standard particles at the end of inflation, or it may be related to the grand unification symmetry breaking or electroweak symmetry breaking. (References [106, 77, 74] contain descriptions of baryonsynthesis and leptonsynthesis.) The number of surviving nucleons is called the baryon number. Other composite particles formed from quarks at various temperatures; however, these particles are unstable and decay into lighter, stable particles. Since $10^{13}K$ is above the threshold temperatures for up, down, and strange quark creation (See §B.7), each of these particle pairs continues to form until the temperature falls below their threshold temperatures. These newly created quark pairs bind into hadrons and anti-hadrons of various types at various temperatures. These hadrons and anti-hadrons annihilate on interaction, sending high energy photons into the plasma that form new particle pairs, including light quark pairs. Elementary and composite particles remain in thermal equilibrium through particle-antiparticle annihilation, pair production and Compton scattering.

Once the temperature falls to $10^{10}K$, which is about one-third of the temperature threshold for the lightest quark, quark pair production shuts down. There has been enough time

for nucleons and anti-nucleons to finish annihilating. Neutrinos fall out of thermal equilibrium[6] (called "decoupling") with the rest of the plasma at about $10^{10}K$ [106, Table 3.1]. The universe is a hot, dense plasma in near thermal equilibrium filled with photons, highly relativistic particles with threshold temperature below $10^{10}K$, decoupled neutrinos and anti-neutrinos of all three types, electrons, positrons, a residual number of nucleons, virtually no anti-nucleons, and dark matter.

The number density of electrons, positrons, and neutrinos (three species of neutrino/antineutrino pairs) are on the order of photon number density. The number density of protons and neutrons are each about one-billionth of the number density of photons, but the number density of neutrons is about 25% of the number density of protons. Two body and three body neutron-proton conversions occur in both directions. As the temperature drops below the formation temperature for electron/positron pairs ($\sim 6 \times 10^9 K$), the number density of electrons and positrons begins to drop significantly as annihilation continues. By $3 \times 10^9 K \equiv \frac{1}{4}\, MeV \sim 10^{-1}\, MeV$ effectively all positrons were annihilated, fewer than one in every billion electrons survived due to leptonsynthesis [106, 77, 74], and beta-decay of neutrons was the only remaining neutron-proton conversion process. The remaining electrons are, more or less, the electrons that exist today.

At around $10^9 K \equiv 86\, KeV$, cosmological nucleosynthesis, the conversion of protons and neutrons into light element nuclei through nuclear reaction chains commences. At about $3 \times 10^8 K \equiv 26\, KeV$ nucleosynthesis finishes, effectively all neutrons are bound into nuclei, and proportions of light element nuclei match the proportions of these elements observed today. References [106, 94, 7] describe this complex process. Reference [7] also provides access to computer codes that model the process and closely match the observed abundances of light atomic nuclei based on a few parameters. A single parameter, the ratio of photon number density to baryon number density, is all that is required to model the development of the relative abundances of light atomic nuclei. The success of theories of nucleosynthesis to match observations places important constraints on cosmologies.

As temperatures drop due to expansion, typical photon energy becomes much smaller than the mass energy of electrons. Therefore Compton scattering may be described as Thomson scattering. Photons and ordinary matter remain in thermal equilibrium through high interaction rates based on Thomson scattering between photons and electrons and at much lower interaction rates between photons and atomic nuclei. At temperatures above $20,000K$, the helium ionization fraction is nearly 100% (doubly ionized at $20,000K$), dropping to near zero at $4,400K$. At temperatures above $4,500K$, the hydrogen ionization fraction begins to drop below 100%. Photons and ordinary matter remain in thermal equilibrium through high interaction rates between photons and electrons until most free electrons are bound into atoms. By the time that the temperature falls to $3,000K$, nearly 100% of ordinary matter is bound into neutral atoms. The interaction rate between photons and ordinary matter, now based only on weak force coupling, drops to negligible levels and photons are effectively decoupled. Light streams unimpeded through space, producing the background of primordial photons described as the CMB. Instruments tuned to the frequencies of the highly redshifted light in the CMB, measure heat from streams of primordial photons in small regions across the sky that are combined to form an image of the LSS as shown in Figure 1.1. Although the process of photon decoupling occurs over a relatively long period, the formation of the LSS is generally set to a specific time and, therefore, may be regarded as a spacelike hypersurface.

The Planck Collaboration determined the redshift of last scattering, $z_* = 1090.48$, as the redshift for which optical depth, $\tau(z)$, is unity, applying a measure of optical depth that excludes re-scattering caused by re-ionization of matter at the time of star formation [87]. This means that the probability that a primordial photon, undergoing its last scattering event at $z_* = 1090.48$, has still traveled for roughly 13.5 billion years without scattering, is $e^{-\tau(1090.48)} = e^{-1} \simeq 36\%$. See §B.6 *Thomson Scattering and Optical Depth* for details.

Thermal or energy history is a better guide to the early universe than coordinate time. However, coordinate times corresponding to high temperatures can be estimated using Eq.

(B.13.12) for temperatures above $10^{10}K$.[7] At temperatures below $10^{10}K$ and above $10^{6}K$, Eq. (B.13.12) can be used with adjustments for the lower temperature of neutrinos due to decoupling. These adjustments are described below Eq. (B.13.12). By $3 \times 10^{8}K$, both electron annihilation and nucleosynthesis have settled down. These processes release thermal energy back into the plasma, causing deviations in the normal $T \propto 1/a$ relationship. Therefore, below $3\times10^{8}K$, the scale factor may be estimated as $a \simeq T_0/T$ and the result can be used in Eq. (2.3.4) to estimate time. For example, $3 \times 10^{8}K$ corresponds to 38 minutes according to Eq. (B.13.12) and to 33 minutes according to Eq. (2.3.4).

The temperature thresholds quoted in this section may be found in §B.7 *Standard Model Elementary Particles Table*. The number densities, the temperatures corresponding to the nucleosynthesis, and ionization temperatures/fractions are from Ref. [106].

2.5 The Required Quantity of Inflation

The quantity of expansion required during the inflationary epoch (called "the crucial assumption of inflationary cosmology" in Lythe & Liddle's 2009 *The Primordial Density Perturbation* [77]) must be great enough so that the region that became the observable universe must have been well inside the Hubble horizon at the beginning of inflation and well outside the Hubble horizon at the end of inflation. Letting the subscript i represent the beginning of inflation and the subscript e represent the end of inflation, the quantity of expansion, $\frac{a_e}{a_i} = e^N$, for some positive number N, must have been sufficient so that

$$\frac{a_i R_0}{H_i^{-1}} < 1$$

and

$$\frac{a_e R_0}{H_e^{-1}} = \frac{a_i\, e^N\, R_0}{H_e^{-1}} > 1$$

which requires that

$$N > \ln\left[\frac{H_e^{-1}}{H_i^{-1}}\right] = \ln\left[\frac{H_i}{H_e}\right].$$

If $H(t)^{-1}$ is the Hubble horizon at time t, then the co-moving Hubble horizon is $H(t)^{-1}a_0/a(t) = H^{-1}/a$ (suppressing the time parameter in the last expression). The time derivative of the co-moving Hubble horizon is

$$\begin{aligned}\frac{d}{dt}\left(\frac{H^{-1}}{a}\right) &= \frac{d}{dt}\left(\frac{a}{\dot{a}}\frac{1}{a}\right)\\ &= \frac{d}{dt}\left(\frac{1}{\dot{a}}\right)\\ &= -\dot{a}^{-2}\ddot{a}\,.\end{aligned}$$

Therefore, the *co-moving Hubble horizon is decreasing during inflation*, since $-\dot{a}^{-2}\ddot{a}$ is negative if $\ddot{a}$ is positive. The inflation condition may be written in terms for the co-moving Hubble horizon as

$$\frac{d}{dt}\left(\frac{H^{-1}}{a}\right) < 0$$

[77, Eq. (18.2)]. The time derivative of the Hubble parameter is

$$\begin{aligned}\dot{H} &= \frac{d}{dt}\left(\frac{\dot{a}}{a}\right)\\ &= \frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^a}\\ &= \frac{\ddot{a}}{a} - H^2\,.\end{aligned}$$

Therefore,

$$\ddot{a} = a\left(\dot{H} + H^2\right)$$

and from above

$$\frac{d}{dt}\left(\frac{H^{-1}}{a}\right) = -\dot{a}^{-2}\ddot{a} = -\dot{a}^{-2}a\left(\dot{H} + H^2\right) < 0\,.$$

Since $-\dot{a}^{-2}a < 0$, it follows that

$$\dot{H} + H^2 > 0\,,$$
$$H + \dot{H}H^{-1} > 0\,,$$

and

$$-\frac{\dot{H}}{H^2} < 1 \tag{2.5.1}$$

are all equivalent to the inflation condition $\ddot{a} > 0$.

Under the assumption that H is decreasing during inflation, $\dot{H}$ is negative and the second expression shows that the decrease in the Hubble parameter over a unit of Hubble time, $\dot{H}\,H^{-1}$, is less than H, perhaps much less. If the magnitude of $\dot{H}$ is sufficiently small, the downward slope of H as a function of time is nearly flat and the value of H is decreasing but nearly constant. If H is constant during inflation, then the inflation mechanism may be modeled as a large magnitude cosmological constant with $a \propto e^{H\,t}$, just as during an epoch dominated by dark energy (see Table 2.1). However, this cannot be a physical description of inflation since inflation *in the observable universe* must have ended.

In cosmology based on Einstein's theory of gravity, gravitation is based on energy density and pressure. Under Eq. (2.1.5), it is evident that if $\rho + 3p > 0$ (gravity is attracting

co-moving objects to each other), then scale is decelerating, and if $\rho + 3p < 0$ (gravity is repelling co-moving objects away from each other), then scale is accelerating.

It will be estimated below that a minimum of about 60 e-foldings is required to explain why space appears to be flat, why the LSS is nearly perfectly isotropic at large scales, and why magnetic monopoles and other relics predicted by grand unification theories have not been detected. These puzzles, encountered in ΛCDM cosmology, are generally called the flatness, horizon, and magnetic monopole problems. The "quantity of inflation" required to solve these problems will be estimated based on the dynamics of a Friedmann universe.

2.5.1 The Flatness Problem

From the Friedmann equation in Eq. (2.1.12), the definition of Ω_k in Eq. (2.1.13), and the definition of the Hubble parameter rearranged as $\dot{a} = a\,H$, the dynamic magnitude of the curvature density parameter may be written

$$\left|\Omega_k\right| = \frac{|k|}{a^2 H^2 R_0^2} = \frac{|k|}{\dot{a}^2 R_0^2}\,. \tag{2.5.2}$$

By definition, space expands at an accelerated rate during the inflationary epoch. Hence $\dot{a}$ is positive and increasing rapidly throughout the epoch, showing that inflation drives $\left|\Omega_k\right|$ toward zero and spatial curvature toward flatness during inflation.

After the end of inflation, space continues to expand. The dynamic, positive rate of expansion, $\dot{a}$, can be determined from Eq. (2.1.4). The acceleration equation in Eq. (2.1.5) shows that $\ddot{a}$ switches from deceleration to acceleration at $\rho + 3p = 0$. This may be expressed as a polynomial equation in a using the density parameters and Hubble constant released by the Planck Collaboration and Table 2.1. The real, positive solution to the equation is the value of a at the turning point from deceleration to acceleration. The time corresponding to this scale, about 6 *billion years* ago, may be determined using Eq. (2.3.4).

In ΛCDM, dark energy density is modeled as a "cosmological constant" where ρ_Λ is constant and $p_\Lambda = -\rho_\Lambda$. The scale of the universe begins to accelerate after the turning point at about 8 *billion years* and continues to accelerate through to the present time and beyond with $\dot{a} \to \infty$. Since ρ_r and ρ_m are diluted by expansion but ρ_Λ is constant, eventually H will be nearly constant after $\rho_r + \rho_m \ll \rho_\Lambda$. The universe becomes approximately a de Sitter universe, effectively devoid of all matter with a positive cosmological constant, Λ, where H has constant value $\sqrt{\frac{8\pi G}{3}\rho_\Lambda}$ and $a \propto e^{Ht}$. Therefore, the scale, the velocity of scale, and the acceleration of scale are all proportional to e^{Ht}. Co-moving regions will be exiting the Hubble horizon at increasing rates. Eventually most of the universe will disappear from view, including the LSS leaving only gravitationally bound regions that are merging or have already merged into a single galaxy.

If no other co-moving objects are within the particle horizon of observers living in some supercluster, the observers will conclude that the entire universe is the supercluster. Even though space will be expanding at an enormous rate, since the galaxies in the supercluster will be in the process of merging, future physicists may conclude that the universe is shrinking and that the entire universe may eventually collapse into a black hole.[8]

Decelerating expansion drives Ω_k away from zero. Accelerating expansion drives Ω_k towards zero. A sufficient period of very rapid acceleration during the inflationary epoch would drive Ω_k close enough to zero so that it remained close to zero after the subsequent 8 *billion* year period of mild deceleration and the more recent and shorter period of mild acceleration. The Planck Collaboration reported that the observable universe "is spatially flat to an accuracy of better than a percent" with $\left|\Omega_{k0}\right| < 0.005$ [88, p. 1]. Looking backwards in time, the value of $\left|\Omega_k\right|$ gets ever closer to zero. For example, in order for the present day magnitude of the curvature density parameter to be less than 0.005, its value at 10^{-11} seconds, the time roughly corresponding to the unification of the weak and electromagnetic forces, must have been 10^{-28} or less.[9] Similarly, it can also be shown that a small deviation from flatness at 10^{-11} seconds of $\left|\Omega_k\right| = 10^{-6}$, would have grown to 10^4

by 1 second, producing a universe dramatically inconsistent with observation. Within the first second of its existence, a closed universe would have collapsed to reach near Planck temperature and an open universe would have cooled to a few degrees.

The quantity of inflation is expressed in e-foldings. For example, an increase in the scale of the observable universe by 30 orders of magnitude is equivalent to about 69 e-foldings since $10^{30} = e^{ln(10^{30})} \simeq e^{69}$. This form is convenient since inflation is exponential or quasi-exponential.

The number of e-foldings from the beginning to the end of inflation is N where $a_e/a_i = e^N$. The following expression finds the number of e-foldings required to reduce the magnitude of the curvature density parameter from unity at the beginning of inflation to its maximum observed value, $\left|\Omega_{k0}\right|$:

$$N = \frac{1}{4} \ln\left(\frac{\Omega_{r0}}{\Omega_{k0}^2} \frac{\rho_1}{\rho_{critical,0}}\right). \tag{2.5.3}$$

This dimensionless expression is derived as Eq. (B.8.11) in §B.8 *Expansion: Flatness*. Using the value of Ω_{r0} from Table 2.1, 10^{-4}, the Planck Collaboration finding that $\left|\Omega_{k0}\right| < 0.005$, Eq. (2.5.3) may be written

$$N = \frac{1}{4} \ln\left(\frac{4\rho_1}{\rho_{critical,0}}\right).$$

Therefore, the condition on the minimum number of e-foldings may be written

$$N > \frac{1}{4} \ln\left(\frac{4\rho_1}{\rho_{critical,0}}\right). \tag{2.5.4}$$

The energy density, ρ_1 (a proxy for the value, ρ_e, at the end of inflation), is needed to calculate the number of e-foldings. Energy in any units may be used in this expression since the ratio of the energy densities is dimensionless. Substituting Planck energy density, $\rho_P^E = \frac{c^7}{\hbar G^2} \simeq 4.63 \times 10^{114}\, erg\, cm^{-3}$ (an upper bound on energy density), and the current

critical density calculated from Eq. (B.12.1), $\rho_{critical\,0} = 7.58263 \times 10^{-9}\, erg\, cm^{-3}$, into Eq. (2.5.4) produces a minimum of 71 e-foldings.

The energy density of a fluid in thermal equilibrium at a given temperature may be determined using Eq. (B.13.9) based on the effective degrees of freedom of the elementary particles in free formation determined using Eq. (B.7.1). At temperatures above $10^{16}K$, all elementary particles are in free formation; therefore, the energy density above $10^{16}K$ may be written as

$$\rho(T) = a_B^* T^4 \tag{2.5.5}$$

where $a_B^* := \left(\frac{427}{4}\right)\frac{a_B}{2}$, a_B is the radiation energy constant, and $\frac{427}{4}$ is the effective degrees of freedom above $T = 10^{16}K$ based on §B.7 *Standard Model Elementary Particles Table*. In cgs units, $a_B^* \simeq 4 \times 10^{-13} erg\, cm^{-3}\, K^{-4}$.

Using $10^{27}K$ as the approximate temperature at the end of reheating and Eq. (2.5.5), the energy density at the end of reheating is $\rho_1 = 4 \times 10^{95}\, erg\, cm^{-3}$. Substituting this value into Eq. (2.5.4) yields *at least* 60 e-foldings. Therefore, a minimum of 60 e-foldings is required to reduce the curvature density parameter from $\left|\Omega_k\right| = 1$, its assumed value at beginning of inflation in this estimate, to $\left|\Omega_{k0}\right| < 0.005$, its upper bound in the current era reported by the Planck Collaboration.

2.5.2 The Horizon Problem

In Chapter 1 it was asserted that, based on ΛCDM, parts of the CMB separated by more than about 1.2° in the sky could not have interacted to reach near uniform temperature. Yet the light from the LSS is observed to be nearly perfectly isotropic over solid angles greater than 7°. The demonstration that a sufficient amount of inflation would explain the near uniform temperature of the LSS will begin with the derivation of 1.2° using ΛCDM based distance measurements (see §2.3).

Light traveling from the LSS is visible today just as light emitted billions of years ago from remote galaxies is visible today. To every observer, the LSS appears as a sphere of background microwave radiation, highly redshifted since last scattering at about $3,000K$ by the expansion of space. The light spectrum emitted by the LSS is that of a nearly perfect black body as seen in Figure 2.1. Although the LSS may be observed and mapped as a

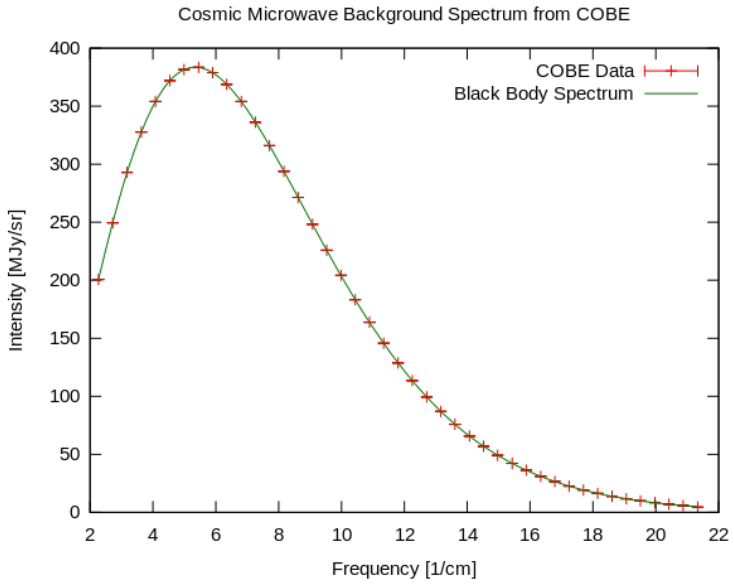


Figure 2.1: The Black Body Spectrum of the CMB based on COBE Dataset, NASA website

"distant" object, it is important to understand that the CMB, which is present throughout space, is the LSS evolved to the present time.[10] Like every spacelike hypersurface, the LSS is "distant" in time only. The physical size of the LSS is the size of today's observable universe scaled down by $a_{t_*} = \frac{1}{z_*+1}$.

The Planck Collaboration reports that redshift of light received from the LSS is $z_* = 1090.48 \rightarrow a_* = \frac{1}{z_*+1}$. This corresponds to $t_* \simeq 370,000$ years, which may be determined using Eq. (2.3.6).

Figure 2.2 shows a present day local observer's location projected backward in time to the LSS hypersurface, Σ_{t_*}, at time t_*. Since we select coordinates where our position is at the center of the observable universe today, the observer's position is also at the center of the LSS hypersurface. Again, LSS is distant in time only. The observer perceives the

LSS as a distant sphere of microwave radiation. When a present day observer picks out a point P on a map of the LSS such as Figure 1.1, then that point is at the outer edge of the LSS, i.e., on the circle in Figure 2.2. A present day extraterrestrial observer, $Obs.X$, at the co-moving location corresponding to point P in our map of the LSS, is at the center of $Obs.X$'s observable universe. If $Obs.X$ scans the sky and produces a map of the LSS, then some point on $Obs.X$'s map corresponds to our co-moving position, so $Obs.X$ would be at the center of his version of Figure 2.2 and we would be on the circle.

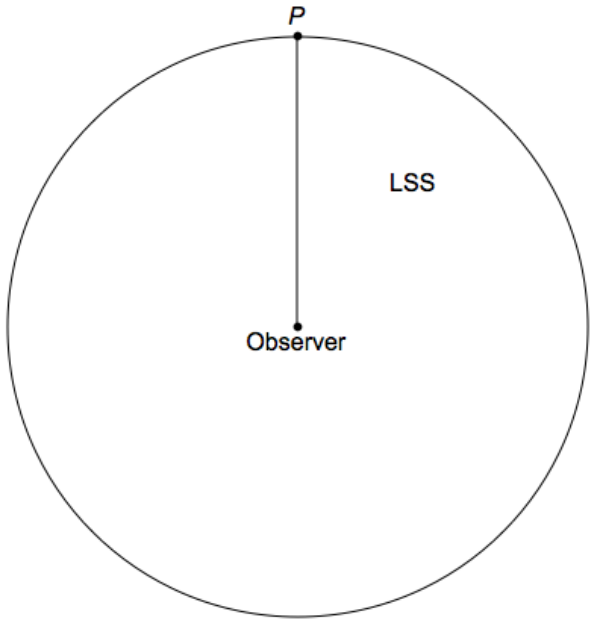


Figure 2.2: An Observer's Position Projected Back to the LSS

The distance light traveled from time $t_i \simeq 0$ to time t through expanding space measured in spacetime fixed at time t (i.e., the hypersurface at time t, Σ_t) is called the physical horizon at time t and is denoted by $d_H(t)$. Therefore, if the distance between two particles in the hypersurface Σ_t is greater than $d_H(t)$, then the two particles could not have exchanged photons or interacted in any way from the beginning of time to time t. The two points are said to be "causally disconnected" at time t. The physical horizon is measured per Eq. (2.3.10) and Eq. (2.3.17) as

$$d_H(t) = a(t) \int_{t_i}^{t} \frac{dt'}{a(t')} . \tag{2.5.6}$$

Therefore, the physical horizon at the time the LSS was formed, t_*, measured in the LSS hypersurface, Σ_{t_*}, is $d_H(t_*)$. Since redshift is observable, the physical horizon is usually measured as $d_H(z_*)$ using Eq. (2.3.14) and Eq. (2.3.17) where $d_H(z_*) = d_H(t_*)$.

The distance light traveled through expanding space from a point P in a hypersurface Σ_t to an observer today measured in the hypersurface Σ_t is called the physical radial distance. The physical radial distance is measured per Eq. (2.3.9) and Eq. (2.3.17) as

$$d_R(t) = a(t) \int_t^{t_0} \frac{dt'}{a(t')} \,. \tag{2.5.7}$$

Therefore, the physical radial distance from a point in the LSS hypersurface, Σ_{t_*}, to an observer today, measured in the LSS hypersurface, Σ_{t_*}, is $d_R(t_*)$. As above, physical radial distances are usually measured as a function of redshift, $d_R(z_*)$, this time using Eq. (2.3.13) and Eq. (2.3.17) where $d_R(z_*) = d_R(t_*)$.

In flat space, the diameter distance, d_A, is the same as the physical radial distance, d_R. In curved space, it is modified to reflect curvature [21, Eq. 8.126]. In positively curved space, $d_A < d_R$ and in negatively curved space, $d_A > d_R$. In slightly curved space $d_A \simeq d_R$. Therefore, after inflation, when space is very flat, $d_A \simeq d_R$. Since the LSS was formed after inflation, $d_A(t_*) \simeq d_R(t_*)$.

Figure 2.3 shows a present day local observer's location projected backward in time to t_* again at the center of the hypersurface Σ_{t_*}. The present day observer is "looking at" a point P on the LSS. Again P is at the edge of the LSS but the circle representing the LSS has been removed from Figure 2.3. The side of the triangle opposite θ is a short segment of an arc in the LSS "circle" in Figure 2.2. The length of the segment is d_H, the physical horizon at the time the LSS was formed. The length of the longer segment is d_A, the diameter distance from the center of the LSS to point P at the "edge" of the LSS. Any particle in the LSS separated from point P by more than d_H was causally disconnected from particles at

point P at time t_* since not even light moving through empty expanding space could have traveled more than distance d_H from the beginning of time to t_*.

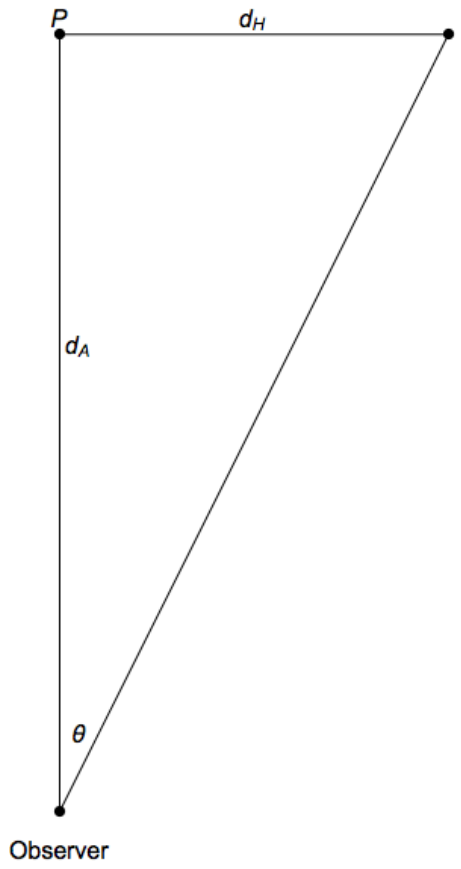


Figure 2.3: ΛCDM Horizon Problem $\theta \simeq 1.2°$

Based on ΛCDM, z_*, and other Planck Collaboration parameters, these measurements are $d_H = \frac{1}{z_*+1} H_0^{-1} \times 6.3 \times 10^{-2}$ and $d_A = \frac{1}{z_*+1} H_0^{-1} \times 3.1$. If θ is small, then, using the small angle approximation, $\theta \simeq \tan\theta = d_H/d_A = 0.02$ radians $\equiv (180/\pi \times 0.02)° \simeq 1.2°$. Thus, based on ΛCDM, a point at P in the LSS could not have interacted with any particle on the LSS separated from P by more than 1.2°. For example, based on ΛCDM, at t_* the observer's position in the LSS was causally disconnected from point P in the LSS. Clearly the observer today is not causally disconnected from point P since the observer is observing P. The light from P has had enough time to reach the observer between $t_* \simeq 370,000$ years and $t_0 \simeq 13.8$ billion years.

Inflationary cosmology changes this situation so that all points in the LSS can reach near thermal equilibrium by requiring that $d_H > d_A$. Figure 2.4 shows the same observer and point but this time inflation has increased d_H so that it is longer than d_A. Again the circle represents the entire LSS and our co-moving position is centered at the origin of the observable universe today and is therefore also located at the center of the LSS. All points are in causal connection with the center since they are within distance d_H of the center. Therefore the LSS moves toward thermal equilibrium. Larger ratios of d_H/d_A lead to more frequent particle interaction and more uniform temperature.

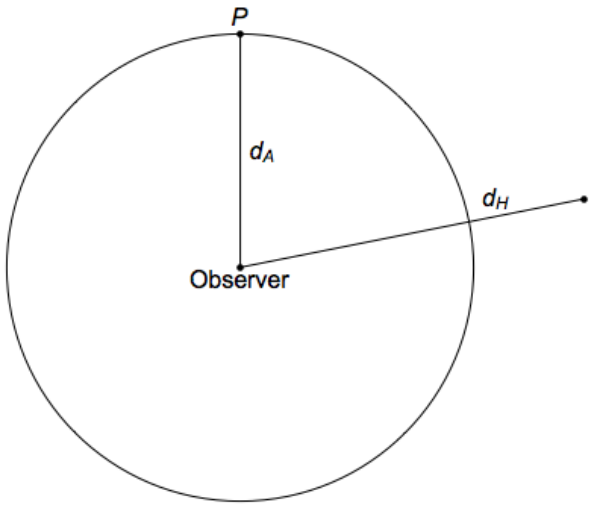


Figure 2.4: Inflation's Horizon Solution: $d_H > d_A$

Selecting the ratio $d_H/d_A = 10$, the following equation expresses the minimum number of e-foldings during the inflation epoch that explains the near perfect isotropy of the LSS. (See derivation of Eq. (B.9.2) in §B.9 *Expansion: Horizon*.)

$$N > \ln\left(10 \cdot \sqrt[4]{\frac{\Omega_{r0}}{\rho_{critical,0}}}\sqrt[4]{\rho_1}\right). \tag{2.5.8}$$

Using $H_0 = 67.040$ from the Planck Collaboration to determine the current value of critical density using Eq. (B.12.1) and the value of Ω_{r0} derived from $T_0 = 2.7255K$ using Eq.

(B.13.9), yields the following condition on the minimum number of e-foldings in cgs units:

$$N > \ln\left(110\sqrt[4]{\rho_1}\right). \tag{2.5.9}$$

After substituting Planck energy density (the upper bound on energy density) into Eq. (2.5.9), the quantity of expansion of the *observable* universe during the inflation epoch required to explain the near uniform temperature of the CMB is at least 71 e-foldings. Substituting $\rho_1 = 4 \times 10^{95}$ corresponding to $10^{27}K$ into Eq. (2.5.9), produces a minimum of about 60 e-foldings. If the ratio $d_H/d_A = 1$ was selected, then the 60 e-foldings are reduced to about 57 e-foldings.

2.5.3 The Magnetic Monopole Problem

Grand unification theories (GUT) describe the unification of the strong force with the electroweak force at about $10^{16}\,GeV$ which corresponds to about $T^{29}K$. Just as symmetry between the weak and electromagnetic forces is broken by the Higgs field (a scalar field), symmetry is broken in GUT by scalar fields. The "scalar fields that break the symmetry can be left in twisted configurations that carry non-zero magnetic charge that cannot be smoothed out by any continuous process" [106, p. 207]. These twisted configurations are magnetic monopoles. If monopoles were created at about $T^{29}K$ and monopole pairs did not subsequently completely annihilate, then it is estimated that the number density of monopoles today would be on the order of the number density of nucleons. If the monopole mass is about $10^{16}\,GeV$, then it is 10^{16} times greater than the nucleon mass (about $1GeV$ [80]). The corresponding monopole density parameter would be about 10^{16} times greater than the density parameter of ordinary matter, Ω_{b0}, which the Planck Collaboration reports as about 0.048. This would result in a highly curved closed universe with $\left|\Omega_{k0}\right| > 10^{14}$ that would have collapsed long ago. Failed attempts to locate monopoles estimate that there must be fewer than 10^{-30} monopoles per nucleon, reducing the monopole density parameter

to an insignificant level of about 10^{-16}. If monopoles were created before inflation or during some intermediate stage of inflation, then the quantity of expansion after monopole creation would need to have been at least 10^{10} to reduce the number density of monopoles by a factor of $10^{-30} = (1/10^{10})^3$. This corresponds to about 23 e-foldings.

Alternately, if reheating produced thermal energy below grand unification energy, monopoles would not have formed. This is the case if $\rho_e \simeq \rho_1 = a_B^* \left(10^{27} K\right)^4$. Experimental production of magnetic monopoles and other evidence of a grand unified gauge group require collision energy of at least $10^{16}\, GeV$ that is about twelve orders of magnitude greater than the energy of collisions produced by the Large Hadron Collider (LHC website data). Grand unification theories could be wrong. There may be no scalar fields left in twisted configurations and, therefore, no magnetic monopoles.

Grand unification energy, the theoretical number density of monopoles, and the upper bound on observed monopole number density used in this section are from Ref. [106].

2.6 Mapping Thermal Anisotropies in the CMB

Maps of the radiation temperature of the CMB may be produced using data collected from satellite, balloon, and land based instruments. Instruments collect data over solid angles in the sky corresponding to the resolution of the instrument. Each solid angle corresponds to the "pixel size" of the instrument. Three generations of satellites with progressively improved resolution have provided data used to construct maps of the CMB and maps of foreground contamination. They are the Cosmic Background Explorer (COBE), the Wilkinson Microwave Anisotropy Probe (WMAP), and the Planck Explorer.

Instruments on these satellites absorb light from the sky. Each instrument consists of multiple receivers (horn assemblies). Radiation is filtered to permit a narrow band of frequencies to strike the surface of a bolometer at the back of the receivers. The frequency at the center of the band is used to label the receiver frequency. The Planck Explorer has

two different instruments: a low frequency instrument with receivers that measure light in four bands each centered on a different low frequency, and a high frequency instrument with receivers that measure light in six bands each centered on a different higher frequency. Bolometer surfaces are cooled to a baseline temperature near absolute zero. The filtered photons received over a time interval increase the temperature of the bolometer surface. The increase in surface temperature is recorded and the surface temperature is reset to the baseline. High frequency light has greater energy so light gathered from smaller solid angles can be used to measure changes in surface temperature. Thus, the high frequency receivers have the best resolution (five arc minutes). Some bolometer surfaces can also measure and record temperature changes due to the electric and magnetic components of polarized light. Bolometers aboard the Planck Explorer can measure polarization at all four low frequency bands and in four out of six high frequency bands.

Maps of CMB light produced at multiple frequencies can be combined and adjusted for foreground contamination. Adjustments are also made for red and blue shifts of light due to peculiar velocities of the instruments relative to the CMB (local orbits, orbit of the solar system within the Milky Way, motion of the Milky Way within the local and super galaxy clusters). Thus in principle, and disregarding the enormous complexity of the actual process, it is possible to obtain a set of temperature fluctuations at nearly all pixels $\mathbf{p} = p(\theta, \phi)$, for each co-moving scale k_p^{-1} corresponding to resolution. Sets may be represented as

$$\{\, \delta T(\mathbf{p}); \mathbf{p} \subset \ LSS \,\}_{g(\nu)} \tag{2.6.1}$$

where $\delta T(\mathbf{p}) = T(\mathbf{p}) - T_0$; $g \in \{T, E, B\}$ referring to the temperature fluctuations of all filtered photons striking the bolometer surface (T), the electric component of filtered polarized photons striking the bolometer surface (E-mode), the magnetic component of filtered polarized photons striking the bolometer surface (B-mode); and ν refers to the frequency of

the receiver. Maps based on scalar temperature variations, E-mode polarized light temperature fluctuations, and B-mode polarized light temperature fluctuations are expanded using spherical harmonics producing sky maps,

$$\delta T(\theta,\phi) = \sum_{\ell>1} \sum_{m=-\ell}^{\ell} a_{\ell m}^{g(\nu)} Y_{\ell m}(\theta,\phi) \ , \tag{2.6.2}$$

where θ, ϕ are the usual polar and azimuthal angles over the LSS sphere. See §B.11 for an example of a technique used to estimate $a_{\ell m}$ from the pixel data sets. The expansions yield observed power spectra, such as the temperature power spectrum, C_ℓ^{TT}, as a function of scale represented by multipole moments, ℓ, where $180/\ell$ corresponds approximately to degrees in great circles of the LSS sphere. Bayesian techniques are used to produce refined values of cosmological parameters that provide the best fit of the observed power spectra to theoretical power spectra simulated by CAMB Codes. See Figure 2.5 comparing the observed temperature power spectrum (red dots) generated from Planck datasets to the theoretical power spectrum (green curve) determined using CAMB Codes. Power spectra are usually graphed as $D_\ell = \frac{\ell(\ell+1)}{2\pi} C_\ell$ where temperature fluctuations are $\delta T/T_0 = D_\ell^{TT} = \frac{\ell(\ell+1)}{2\pi} C_\ell^{TT}$. See §B.11 *Sky Map Expansion in Spherical Harmonics* for an overview of how data sets are expanded using spherical harmonics to produce observed power spectra.

These Bayesian techniques together with CAMB Codes and data sets collected from satellites and other instruments form the basis of "precision cosmology" that is introduced in §B.15 *Precision Cosmology*. For a detailed description of precision cosmology see *Planck 2013 results. XV. CMB power spectra and likelihood* [86]. CAMB Codes used by the Planck Collaboration are described in [57, 59, 58].

The evolution of temperature fluctuations (observed as anisotropies in the CMB) and energy density perturbations both result from the primordial curvature perturbation [77]. As discussed above, the primordial curvature perturbations at all scales were seeded by quantum uncertainties during inflation. The dynamics of the growth of energy density

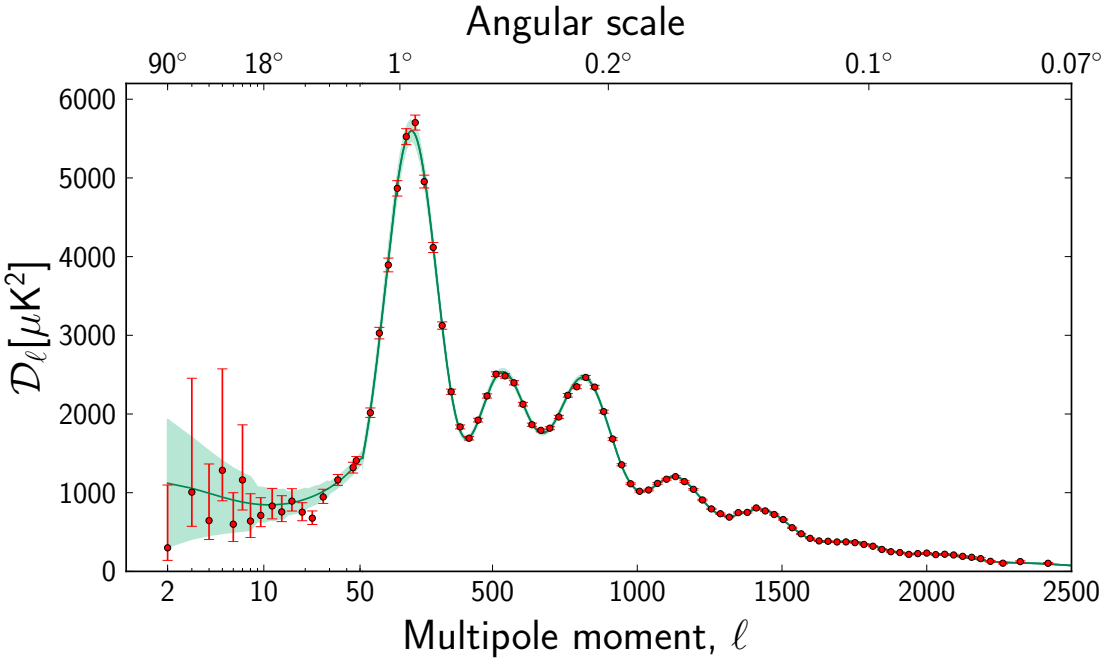


Figure 2.5: Copyright: ESA and the Planck Collaboration, March 2013. "The 2013 *Planck* CMB temperature power spectrum. The error bars include cosmic variance, whose magnitude is indicated by the green shaded area around the best fit model."

variations from the beginning of the radiation era to photon decoupling occurred at all scales. After nucleosynthesis and up to photon decoupling, the contents of the universe may be viewed as photon-baryon fluid (plus decoupled neutrinos). Oscillations of compression and decompression of matter in the photon-baryon fluid occurred. The compression phase was driven by gravity. Radiation pressure increased in response to compression, producing a decompression phase. The oscillations are called sound waves and the cycles are called baryon acoustic oscillations. Oscillations occur at all scales up to the maximum distance sound could travel through the fluid from the formation of the fluid up to photon decoupling. This maximum distance is called the co-moving sound horizon, r_*, and the size is $\theta_* = r_*/D_A(z_*) \simeq r_*/D_R(z_*)$ (Eq. (2.3.13)). The size of the sound horizon, θ_* is one of the base cosmological parameters measured by the Planck Collaboration.

After photon decoupling, the pressure dropped to zero, the oscillations stopped, and the temperature variations in the photon gas, i.e., the CMB, imprinted by the baryon en-

ergy density variations remained fixed other than by the effects of foreground contamination. Thus, the temperature anisotropies are the consequence of events beginning during inflation and ending at last scattering. On the other hand, the polarization of temperature anisotropies is only generated by the last scattering of CMB photons (Thomson scattering) or by foreground contamination such as dust. Identifying foreground contamination is difficult, involving multiple scans at several frequencies.

E-mode and B-mode polarization patterns are produced by energy density perturbations and, if they exist, by gravitational waves. Figure 2.6 shows the azimuthal symmetry in the polarization patterns in the scalar quadrupole moments ($\ell = 2, m = 0$) produced by energy density perturbations.

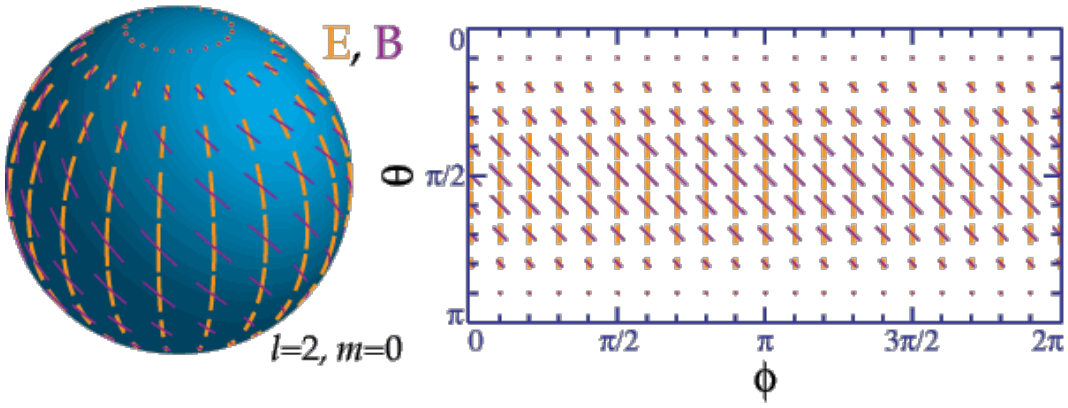


Figure 2.6: E-mode and B-mode Polarization Patterns Produced by Density Perturbations at Last Scattering. Images by Professor Wayne Hu (University of Chicago website) [44]

Figure 2.7 shows the lack of azimuthal symmetry in the polarization patterns of tensor quadrupole moments ($\ell = 2, m = 2$) expected to be produced by gravitational waves.

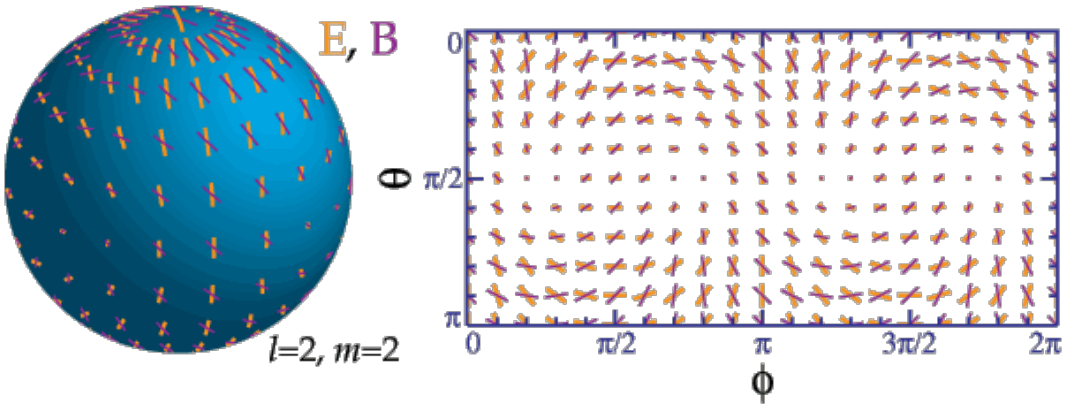


Figure 2.7: E-mode and B-mode Polarization Patterns Produced by Gravitational Waves at Last Scattering. Images by Professor Wayne Hu (University of Chicago website) [44]

If they can be detected, strong B-mode polarization signals would confirm the existence of gravitational waves.

2.7 Primordial Energy Density Perturbations

Primordial energy density perturbations are closely related to the thermal anisotropies in the CMB described in the previous section. These perturbations also caused the development of the large scale cosmological structures seen in the universe today. This section explores the properties of these perturbations and shows how they are caused by "the curvature perturbation". The connection of the curvature perturbation to inflation will be described in §3.2 *Quantum Origin of Perturbations*. The link between the perturbations and the development of cosmological structures is described in Ref. [75].

Photon and neutrino number and energy densities based on temperature may be determined using Eq. (B.13.8). The ratio of baryon number density to photon number density is well known since it drives nucleosynthesis which successfully predicts the relative abundances of light elements [7, 97]. With this information and observed CMB temperature anisotropies, it is possible to probe for signals of non Gaussian distribution primordial energy density perturbations and non adiabatic energy density perturbations (described below).

In *Planck 2015 results. XIII. Cosmological parameters* [88], the Planck Collaboration finds that the "Planck results offer powerful evidence in favor of simple inflationary models, which provide an attractive mechanism for generating the slightly tilted spectrum of (nearly) Gaussian adiabatic perturbations that match our data to such high precision".

For the purposes of this section, primordial density perturbations will be assumed to be adiabatic with Gaussian distribution. The "slightly tilted spectrum" will be used to derive a useful expression for the root mean square ("rms") value of the curvature perturbation.

2.7.1 Cosmological Regions and the Hubble Horizon

The crucial assumption of inflationary cosmology is that the region that became the observable universe was well inside the Hubble horizon at the beginning of inflation and well outside the Hubble horizon at the end of inflation (see §2.5). Figure 2.8 (in logarithmic scales) shows the time that a co-moving region exited the Hubble horizon at the point of the first intersection (left to right). This is called "Hubble exit". The second intersection shows the time that the co-moving region entered the co-moving Hubble horizon after the end of inflation. This is called "Hubble re-entry". A period during which a co-moving region is outside the Hubble horizon is called a super-horizon era and a period within the Hubble horizon is called a sub-horizon era. In Fig. 2.8, the super-horizon era is between the two points of intersection. The upper graph is presented in physical scale and the lower graph is presented in co-moving scale (the term scale should not be confused with scale factor represented by the function $a(t)$).

Significant particle flow (hence heat energy flow) does not occur among non-overlapping regions of some common scale during the super-horizon era.[11] After Hubble re-entry, gravitation and pressure will cause particles to flow among regions within a common Hubble horizon. After photon decoupling, radiation pressure drops to zero and energy density variations among non-overlapping regions of a given scale begin to increase, leading to the formation of stars and later to galaxies and galaxy clusters.

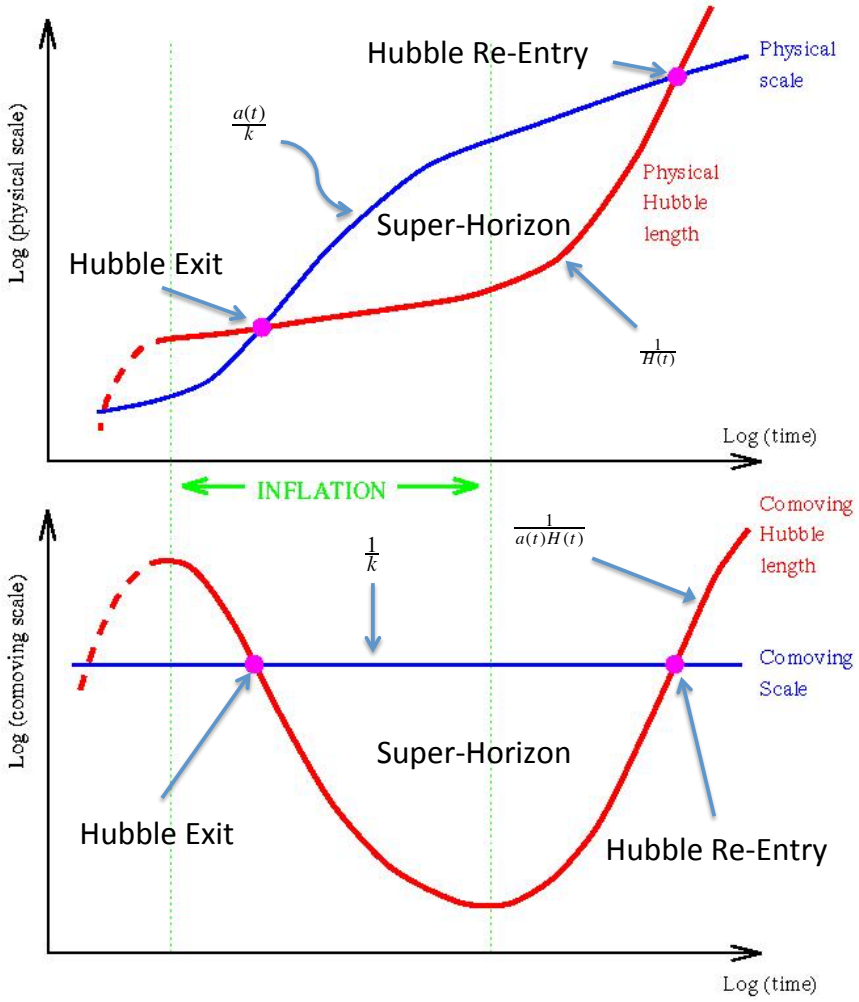


Figure 2.8: Hubble exit & re-entry: physical and co-moving views, from Ref. [60]

Co-moving regions are generally volumes bounded by a sphere of co-moving radius R. In cosmology, it is common to identify the scale of a co-moving region, R, with a co-moving wavelength, λ, and to use the term wavenumber to refer to the inverse of the wavelength, $k = 1/\lambda$, or the inverse of the radius, $k = 1/R$. This definition of wavenumber and the three dimensional counterpart, the wavevector, will be used in this paper, including in the context of Fourier transformations. It is consistent with the "kayser" in cgs units, wavelengths per centimeter. For example, if R is the radius of a spherical co-moving region, then $k = 1/R$ and the time at which the co-moving region exits the co-moving Hubble horizon (i.e., a

co-moving Hubble sphere with radius $1/a\,H$) is given by

$$k = a(t)\,H(t) = \dot{a}(t)\,. \tag{2.7.1}$$

Understanding energy density perturbations at various scales is based on analysis of co-moving "cosmological regions". The smallest cosmological region is the smallest region in which gravitational collapse of baryons (aided by the effects of dark matter) could have occurred. The largest cosmological region is the observable universe. The smallest co-moving cosmological region must contain ordinary and dark matter equivalent to about $10,000$ Solar Masses. Cosmological regions that contain more than 10^{15} Solar Masses have not had enough time to collapse before the current era [77]. Under the usual assumption of near homogeneity on large scales, the smallest cosmological region bounded by a spherical surface must satisfy

$$\frac{4\pi}{3} k^{-3}\,\Omega_{m0}\,\rho_{critical,0} = 10^4 \times \odot$$

where $\odot$ is the symbol for Solar Mass. Using Ω_{m0} from Table 2.1, the value of critical energy density (Eq. (B.12.1)), and Solar Mass from Table A.1, the co-moving radius of the smallest cosmological region is $k^{-1} \simeq 4 \times\ Kpc$. Similarly, the largest cosmological region which has undergone gravitational collapse has co-moving radius $k^{-1} \simeq 18\,Mpc$, putting a limit on the size and the mass of the largest gravitationally bound structures in the current era: galaxy superclusters.

Calculating the scale factor, a, at Hubble re-entry as a function of the co-moving size is straightforward. Based on Table 2.1, the Hubble parameter as a function of the scale factor is

$$H(a) = \sqrt{\frac{8\pi\,G\,(\rho_{r0}a^{-4} + \rho_{m0}a^{-3} + \rho_{\Lambda 0})}{3}}\,. \tag{2.7.2}$$

The scale factor at which Hubble re-entry occurs for a region of co-moving scale, k^{-1}, is given by

$$\frac{a}{k\,H^{-1}} = 1\,. \tag{2.7.3}$$

So the scale factor at which Hubble re-entry occurs is determined by solving

$$\frac{a}{k}\,\frac{1}{c}\,\sqrt{\frac{8\pi\,G\,(\rho_{r0}a^{-4} + \rho_{m0}a^{-3} + \rho_{\Lambda 0})}{3}} = 1 \tag{2.7.4}$$

for a where the constant $1/c$ has been added as a reminder that H^{-1} has dimensions of time. Based on the density parameters in Table 2.1, the previously derived value of critical energy density, and the scale of the smallest cosmological region,

$$R_{10^4\odot} = k^{-1} \simeq 4 \times Kpc\,. \tag{2.7.5}$$

The real solution to Eq. (2.7.4) determined using *Mathematica* is

$$a \simeq 8.4 \times 10^{-9}\,.$$

This scale factor coincides roughly with the scale factor at the end of nucleosynthesis that occurs after electron-positron annihilation. It is, therefore, reasonable to estimate the temperature as $T \simeq T_0/a$. This yields $T \simeq 3 \times 10^8 K$ which is equivalent to plasma thermal energy of $30\,KeV < 10^{-1} MeV$.

Regions of larger scale re-enter later and hence at lower levels of thermal energy. For example, the largest regions that have undergone gravitational collapse re-enter at about $1,300$ years, still well before radiation-matter equality. Regions that had not re-entered the co-moving Hubble horizon when the expansion rate began to accelerate about 6 billion years ago will not re-enter the co-moving Hubble horizon (see p. 33). Therefore, all

relevant regions were super-horizon when the thermal energy of the plasma was about $250KeV$ (corresponding approximately to the end of electron-positron annihilation (p. 28). Based on Eq. (B.13.12), this corresponds to approximately 15 seconds, represented here as $t_\dagger$ (Table 2.2). The precise thermal energy and time is not important. Thermal energy is selected to be high enough so that regions at all scales were super-horizon and below electron-positron annihilation to avoid thermal irregularities which introduce unnecessary complexity to calculations.

2.7.2 Perturbation Analysis by Region

At time t, the total energy density (excluding dark energy) at co-moving location $\mathbf{x}$ is $\rho(\mathbf{x}, t)$, and its components are $\rho_\alpha(\mathbf{x}, t)$ where α represents the symbols ν, γ, b, and c referring to neutrinos, photons, ordinary matter and cold dark matter respectively [77]. The background energy density, $\rho(t)$, is the average total energy density over all space at time t. Similarly, $\rho_\alpha(t)$, is the average energy density of component α over all space at time t. The total energy density perturbation at each position $\mathbf{x}$ at time t is the departure of total energy density at that point from the background energy density:

$$\delta\rho(\mathbf{x}, t) = \rho(\mathbf{x}, t) - \rho(t) \tag{2.7.6}$$

and the total energy density contrast is

$$\frac{\delta\rho(\mathbf{x}, t)}{\rho(t)}\,. \tag{2.7.7}$$

Energy density is "smoothed" to wavenumber k at a location, $\mathbf{x}$, by substituting the average energy density in a co-moving sphere of scale k^{-1} centered on $\mathbf{x}$ for the actual energy density at $\mathbf{x}$. Smoothing to wavenumber k of a density component is given by $\rho_\alpha(\mathbf{x}, t) = \rho_\alpha(k, \mathbf{x}, t)$ and smoothing of total energy density is given by $\rho(\mathbf{x}, t) = \sum_\alpha \rho_\alpha(\mathbf{x}, t)$. Smoothed energy density perturbations of a component [77, Eq. (5.3)] are represented by

the departure of smoothed energy density from $\rho_\alpha(t)$, the average of component α over all space at time t

$$\delta\rho_\alpha(k, \mathbf{x}, t) = \rho_\alpha(k, \mathbf{x}, t) - \rho_\alpha(t).$$

Smoothed total energy density perturbations are represented by the departure of smoothed total energy density from the background energy density:

$$\delta\rho(k, \mathbf{x}, t) = \rho(k, \mathbf{x}, t) - \rho(t) \tag{2.7.8}$$

and the smoothed energy density contrast is

$$\frac{\delta\rho(k, \mathbf{x}, t)}{\rho(t)}. \tag{2.7.9}$$

The sets of energy density perturbations measured at $t_\dagger$ at all positions $\mathbf{x}$ with and without smoothing are sets of *primordial energy density perturbations*:

$$\left\{ \delta\rho(k, \mathbf{x}, t_\dagger) \right\}_{\mathbf{x}\in\mathbb{R}^3} \tag{2.7.10}$$

$$\left\{ \delta\rho(\mathbf{x}, t_\dagger) \right\}_{\mathbf{x}\in\mathbb{R}^3} \tag{2.7.11}$$

where $\mathbb{R}^3$ or Euclidian space is used in place of the three dimensional spatial sub-manifold of spacetime written above as Σ_t. This notation is used for notational simplicity but it is also justified since it is assumed space was driven flat to high precision during the inflation epoch.

Similarly, the sets of energy density contrasts measured at $t_\dagger$ at all positions $\mathbf{x}$ with and without smoothing, are of sets of *primordial energy density contrasts*:

$$\left\{ \frac{\delta\rho(k, \mathbf{x}, t_\dagger)}{\rho(t_\dagger)} \right\}_{\mathbf{x}\in\mathbb{R}^3} \tag{2.7.12}$$

$$\left\{ \frac{\delta\rho(\mathbf{x}, t_\dagger)}{\rho(t_\dagger)} \right\}_{\mathbf{x}\in\mathbb{R}^3}. \tag{2.7.13}$$

Theory, supported by observations, finds that the mean-square values of the elements in the sets defined in Eq. (2.7.10) and Eq. (2.7.12) are nearly scale invariant (i.e., that the mean-square values for different k are nearly the same at $t_\dagger$ and at other times during which all scales were well outside the Hubble horizon).

In §2.7.3 *Statistical Analysis of Primordial Scalar Perturbations* it will be shown that the mean-square values of energy density perturbations (Eq. (2.7.11)) and contrasts (Eq. (2.7.13)) are related to integrals of their power spectra, $\mathcal{P}_{\delta\rho}(k)$ and $\mathcal{P}_{\delta\rho/\rho}(k)$, determined by Fourier expansion of the perturbations and the contrasts. In §3.2 *Quantum Origin of Perturbations* it will be shown that these power spectra are related to the power spectrum of the curvature perturbation, $\mathcal{P}_\zeta(k)$, determined during inflation at about the time that *inflaton* modes with physical wavelength $a(t)/k$ first exceeded Hubble length, $H^{-1}(t)$.

The usual definition of an adiabatic process is one that does not produce exchange of heat energy between the system and its surroundings. Any energy transferred between the system and its surroundings as a result of the process is transferred as work. If the observable universe is regarded a system, then all processes occurring within the observable universe fit the general definition of adiabatic since the observable universe is causally disconnected from surrounding regions. Similarly, a region of some co-moving scale, k^{-1}, is causally disconnected from surrounding regions while it is well outside the Hubble horizon. Therefore, there cannot be any particles, and hence no heat energy, exchanged between the region and surrounding regions while it is well outside the Hubble horizon.

In cosmology, a quantity is described as adiabatic [77, p. 72] during a time interval if it is a unique function of energy density or a component of energy density throughout all space during the time interval (which may be all time). Under the usual assumptions that the universe is homogeneous and isotropic on large scales, the components of energy density are adiabatic based on Table 2.1 since $\rho_\alpha(\mathbf{x}, t) = \rho_{\alpha\,0} a(t)^{-n_\alpha}$ where $\rho_{\alpha\,0}$ and n_α are constants. Additionally, based on Table 2.1, pressure is adiabatic since $p_\alpha(\mathbf{x}, t) = w_\alpha \rho_\alpha$ for each energy density component $\alpha = \gamma, \nu, b, c, \Lambda$ where w_α is constant. It will be shown that

the theory that energy density contrasts are caused by the curvature perturbation, predicts that energy density perturbations are adiabatic. These predictions are supported by the Planck Collaboration conclusion that observations are consistent with "purely adiabatic spectrum of fluctuations" [88].

Observations by the Planck Collaboration are consistent with the hypothesis that elements in the set of primordial energy density contrasts are functions of an initial condition called the primordial curvature perturbation, $\zeta(\mathbf{x})$. §3.2 *Quantum Origin of Perturbations* will show that single field slow-roll models of inflation produce quantum fluctuations of the *inflaton field* in a spectrum of wavelengths during inflation with amplitudes that are fixed at about the time of co-moving Hubble exit, $\delta\phi(\mathbf{k})$, where $\lambda = 1/|\mathbf{k}|$. The amplitude functions of the curvature perturbation, $\zeta(\mathbf{k})$, are linear functions of $\delta\phi(\mathbf{k})$, and $\zeta(\mathbf{x})$ may be produced by inverse Fourier transformation of the amplitudes. Each $\zeta(\mathbf{x})$ represents the curvature perturbation in a co-moving region established within the *inflaton field* but which continues into at least the radiation era. The scale is commonly set to the smallest scale of cosmological interest defined in Eq. (2.7.5) but may be defined on co-moving scales up to the scale of the observable universe. It is shown that the energy density contrasts are directly linked to the curvature perturbation and that the curvature perturbation is conserved from Horizon exit to Horizon re-entry of the scale used to determine the perturbation (between the two intersections in Figure 2.8). It is also shown that curvature perturbations produced in single field slow-roll models are adiabatic with normal Gaussian distribution.

2.7.3 Statistical Analysis of Primordial Scalar Perturbations

Overview. The statistical properties of generic scalar perturbations at a fixed point in time are examined in this section. Scalar perturbations are assumed to have Gaussian distribution. See §B.14 *Gaussian Distribution*. Perturbations are transformed from position space into k-space using Fourier expansion. An expression is derived for the two-point correlator that is used to find the variance and, therefore, the probability distribution of perturbations

in position space. The variance is expressed as the integral sum of the power spectrum per logarithmic unit of the wavenumber.

Random Fields. Up to this point, perturbations, such as the primordial density perturbation, $\delta\rho(\mathbf{x})$, have been discussed as measurable real valued quantities at a fixed time. Although it is not feasible for an observer in the current era to make such measurements, real values of primordial density perturbations existed at each position in space (or at least in a small region about a position) at time $t_\dagger$. The state of primordial density perturbations over all positions in the space was given in Eq. (2.7.11) as $\{\, \delta\rho(\mathbf{x}) \,\}_{\mathbf{x}\in\mathbb{R}^3}$.

According to inflationary cosmology, these perturbations and the curvature perturbation were caused by quantum fluctuations during inflation. Since such quantum fluctuations or uncertainties are random, the state of measurable perturbations in our universe is the outcome or realization of a random process. Over and under densities at each point would have differed based on the realization of the quantum fluctuations. A different spatial outcome of quantum fluctuations with the same statistical distribution of quantum fluctuations would have evolved into a different but statistically equivalent state of primordial density perturbations. These would have evolved into a different but statistically equivalent state of the universe in the current era, i.e., different galaxies but a statistically equivalent distribution of the number and mass densities of galaxies and galaxy clusters. Our universe and, therefore, our galaxy evolved from one realization of quantum fluctuations during inflation.

With this in mind, scalar perturbations such as the primordial density perturbation, the primordial density contrast, and the curvature perturbation are all Gaussian random fields (also known as random processes and stochastic processes) with Eq. (2.7.11) regarded as a realization of the random field of primordial density perturbations.

The symbol $\delta(\mathbf{x})$ will be used to represent the Gaussian random field of a generic perturbation. A Gaussian random field is an uncountably infinite set of uncorrelated (i.e., independent) random variables, X, one for each $\mathbf{x} \in \mathbb{R}^3$. There are an uncountable number

of realizations of the random field $\delta(\mathbf{x})$, each written $\delta_{\mathcal{R}}(\mathbf{x})$ with $\mathcal{R}$ serving as the label for each distinct realization [90]. Each $\delta_{\mathcal{R}}(\mathbf{x})$ is a real valued function, $\delta_{\mathcal{R}} : \mathbb{R}^3 \rightarrow \mathbb{R}$ (generally, but not necessarily, discontinuous given its construction from the outcomes of independent random variables).

The symbol Ω represents the realization space (usually called the outcome space in the theory of random fields [4]) of the uncountable ensemble of realizations, $\delta_{\mathcal{R}}(\mathbf{x})$, of the random field, $\delta(\mathbf{x})$. Each $\delta_{\mathcal{R}}(\mathbf{x})$ represents an ordering (i.e., the realized ordering or the outcome) of real numbers usually, but not necessarily, clustered in a small interval about the origin given the definition of a perturbation. The cardinalities of the domain and codomain of each $\delta_{\mathcal{R}}(\mathbf{x}) \in \Omega$ are $\beth_1$, the cardinality of the set of real numbers. Since Ω is the ensemble of all possible orderings, each one represented by $\delta_{\mathcal{R}}(\mathbf{x})$, the cardinality of Ω is $\beth_2$ which is the cardinality of the uncountably infinite dimensional real number space $\mathbb{R}^{\mathbb{R}}$. The subset $\Omega \subset \Omega_{\mathbf{x}}$, which is the subset in which the domain of $\delta_{\mathcal{R}}(\mathbf{x})$ is limited to a specific position $\mathbf{x}$, also has cardinality $\beth_2$ (analogous to the cardinality of an interval in the real number line being the same as the cardinality of the entire number line).

Consistent with the homogeneous and isotropic background of unperturbed space in cosmology, random fields of scalar perturbations are assumed to be homogeneous and isotropic. Such random fields are called stationary [4]. The probability densities of stationary random fields are invariant on translation, rotation, and parity transformations [77]. This means that the probability density function for the realization $\delta_{\mathcal{R}}(\mathbf{x})$ is the same for the realization $\delta_{\mathcal{R}}(\mathbf{r} + \mathbf{x})$ for any $\mathbf{r}$. Equivalently, the probability density function for the realization $\delta_{\mathcal{R}}(\mathbf{x})$ is position independent and may be written $\rho(\delta_{\mathcal{R}})$, forming a function space of probability density functions.

The ensemble average of the random field, $\delta(\mathbf{x})$, at fixed position $\mathbf{x}$, is the expected value of the field at position $\mathbf{x}$. The expected value is the probability density weighted

average taken over the subset of the realization space corresponding to position $\mathbf{x}$:

$$\langle \delta(\mathbf{x}) \rangle = \int_{\Omega_{\mathbf{x}}} \delta_{\mathcal{R}}(\mathbf{x})\, \rho(\delta_{\mathcal{R}})\, d\, \delta_{\mathcal{R}}(\mathbf{x}) \tag{2.7.14}$$

where $\Omega_{\mathbf{x}} \subset \Omega$. This expression, which is the mean (first moment) of the field, is shown to vanish on p. 67.

The two point correlator of the field, $\xi(\mathbf{x}, \mathbf{x}')$, is the ensemble average of the product of the field evaluated at two fixed positions $\delta(\mathbf{x})\, \delta(\mathbf{x}')$. The ensemble average is the expected value which is the probability density weighted average of the product:

$$\begin{aligned} \xi(\mathbf{x}, \mathbf{x}') &= \langle \delta(\mathbf{x})\, \delta(\mathbf{x}') \rangle \\ &= \int_{\Omega_{\mathbf{x}}} \int_{\Omega_{\mathbf{x}'}} \delta_{\mathcal{R}}(\mathbf{x})\, \delta_{\mathcal{R}}(\mathbf{x}')\, \rho(\delta_{\mathcal{R}})\, d\, \delta_{\mathcal{R}}(\mathbf{x}')\, d\delta_{\mathcal{R}}(\mathbf{x})\,. \end{aligned} \tag{2.7.15}$$

Let $\mathbf{r} = \mathbf{x}' - \mathbf{x}$, then $\langle \delta(\mathbf{x})\, \delta(\mathbf{x}') \rangle \equiv \langle \delta(\mathbf{x})\, \delta(\mathbf{r} + \mathbf{x}) \rangle$. Since the fixed positions are arbitrary, $\xi(\mathbf{x}, \mathbf{x}')$ depends only on the magnitude of the distance separating the positions, $r = |\mathbf{r}| = |\mathbf{x}' - \mathbf{x}|$. Therefore, $\xi(\mathbf{x}, \mathbf{x}')$ is usually represented as

$$\xi(r) = \langle \delta(\mathbf{x})\, \delta(\mathbf{r} + \mathbf{x}) \rangle \tag{2.7.16}$$

[83, Eq. (21.4)]. Useful versions of this expression are found in Eq. (2.7.27), Eq. (2.7.28), and Eq. (2.7.29). If $\mathbf{r}$ is the null vector, then Eq. (2.7.16) produces the variance (second moment) of the field. See Eq. (2.7.30) and Eq. (2.7.31). When the random fields in the product are the same field, as is the case here, the two-point correlator is called the two-point auto correlator.

The first and second moments of a Gaussian field completely determine the probability distribution of the field. *Hence, the ensemble average and the two-point auto correlator completely determine the distribution of a Gaussian random field.* Higher-order correlators are included in determination of the distribution for non Gaussian random fields.

Perturbations modeled as random fields are used to develop simulations of the evolution of the universe from the primordial epoch through the current era. The perturbations that seeded the evolution of CMB anisotropies and the formation of cosmological structures in our universe correspond to a single realization of a random field. All observation data are collected from a single realization.

Observations can be tested against simulations derived from theory if spatial averages of a random field are equal to ensemble averages. This is known as the ergodic property of a random field. The ergodic property will be demonstrated below for stationary Gaussian random fields.

Application of the Fourier Expansion. The Fourier expansion of the generic random field of a perturbation in three spatial dimensions, $\delta(\mathbf{x})$, is determined by a transformation and an inverse transformation. The transformation and inverse transformation used in this paper are:

$$\hat{\delta}(\mathbf{k}) = \int \delta(\mathbf{x})\, e^{-i\,\mathbf{k}\cdot\mathbf{x}} d\mathbf{x} \tag{2.7.17}$$

and

$$\delta(\mathbf{x}) = \frac{1}{(2\pi)^3} \int \hat{\delta}(\mathbf{k})\, e^{i\,\mathbf{k}\cdot\mathbf{x}} d\mathbf{k} \tag{2.7.18}$$

[77, Eq. (6.2)]. The function, $\delta(\mathbf{x})$ in position space, is represented as a linear combination of wave functions, $e^{i\,\mathbf{k}\cdot\mathbf{x}}$, with wavenumber equal to the magnitude of the wavevector, $k = |\mathbf{k}|$, and coefficient functions, $\hat{\delta}(\mathbf{k})$, in k-space, that serve as wave amplitudes. The magnitude of the wavevector is equal to the inverse of the wavelength (i.e. $k = 1/\lambda$ not $k = 2\pi/\lambda$). The appearance of $(2\pi)^3$ in these expressions results from the dependence of the transformation on the Dirac Delta and is, therefore, independent of the meaning of the wavevector.

In inflationary cosmology, k-space functions are frequently written without differentiation from the position space function other than by the argument; therefore, $\delta(\mathbf{k}) \equiv \hat{\delta}(\mathbf{k})$. This convention will be used in this paper. For example $\delta(\mathbf{k})$, means the amplitude coefficient function in k-space, whereas $\delta(\mathbf{x})$, $\delta(\mathbf{y})$, and $\delta(\mathbf{r})$ all represent the perturbation function in true position space. *Note that δ^3 always represents the Dirac Delta generalized function, not the k-space function.*

Common representations of the three dimensional Dirac Delta generalized function include

$$\delta^3(\mathbf{k}) = \frac{1}{(2\pi)^3} \int e^{\pm i\,\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{x} \tag{2.7.19}$$

and

$$\delta^3(\mathbf{k} \pm \mathbf{k}') = \frac{1}{(2\pi)^3} \int e^{\pm i\,(\mathbf{k}\pm\mathbf{k}')\cdot\mathbf{x}}\, d\mathbf{x} \tag{2.7.20}$$

[83, Eq. (21.45)].

Applying the inverse Fourier transformation and a representation of the Dirac Delta in Eq. (2.7.19), the ensemble average of the random field, $\delta(\mathbf{k})$, may be written

$$\begin{aligned} \langle \delta(\mathbf{k}) \rangle &= \left\langle \int \delta(\mathbf{x})\, e^{-i\,\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{x} \right\rangle \\ &= \int \langle \delta(\mathbf{x}) \rangle\; e^{-i\,\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{x} \\ &= (2\pi)^3 \delta^3(\mathbf{k})\, \langle \delta(\mathbf{x}) \rangle \ . \end{aligned}$$

From this expression, it is evident that $\langle \delta(\mathbf{k}) \rangle = 0$ if $\mathbf{k}$ is not the null vector. Additionally, $\langle \delta(\mathbf{x}) \rangle$ is zero provided that all fluctuations are included in the calculation of the background energy density (see also p. 67). Therefore, for each point in position space and in k-space,

the ensemble averages of the random fields, $\delta(\mathbf{x})$ and $\delta(\mathbf{k})$, vanish:

$$\langle \delta(\mathbf{x}) \rangle = 0 \tag{2.7.21}$$

and

$$\langle \delta(\mathbf{k}) \rangle = 0\,. \tag{2.7.22}$$

The coefficients of the wave amplitudes in the Fourier expansion of any function with Gaussian distribution are uncorrelated except as required by the "reality condition" [77]. Since $\delta(\mathbf{x})$ is a real-valued function, the complex conjugate of the field is equal to itself, i.e., $\delta^*(\mathbf{x}) = \delta(\mathbf{x})$. Applying the transformation in Eq. (2.7.17) yields the "reality condition":

$$\delta^*(\mathbf{k}) = \int \delta^*(\mathbf{x})\, e^{i\,\mathbf{k}\cdot\mathbf{x}} d\mathbf{x} = \int \delta(\mathbf{x})\, e^{i\,\mathbf{k}\cdot\mathbf{x}} d\mathbf{x} = \delta(-\mathbf{k})\,.$$

Since the function $\delta(\mathbf{x})$ has Gaussian distribution, its k-space coefficients are uncorrelated except by the reality condition. The two-point correlator in k-space may be written:

$$\begin{aligned}
\langle \delta(\mathbf{k})\, \delta(\mathbf{k}') \rangle &= \left\langle \int d\mathbf{x}\, \delta(\mathbf{x})\, e^{-i\,\mathbf{k}\cdot\mathbf{x}} \int d\mathbf{y}\, \delta(\mathbf{y})\, e^{-i\,\mathbf{k}'\cdot\mathbf{y}} \right\rangle \\
&= \left\langle \int \int \delta(\mathbf{x})\, \delta(\mathbf{y})\, e^{-i\,\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{x}\, e^{-i\,\mathbf{k}'\cdot\mathbf{y}} d\mathbf{y} \right\rangle \\
&= \int \int \langle \delta(\mathbf{x})\, \delta(\mathbf{y}) \rangle\; e^{-i\,\mathbf{k}\cdot\mathbf{x}}\, e^{-i\,\mathbf{k}'\cdot\mathbf{y}}\, d\mathbf{x}\, d\mathbf{y}
\end{aligned}$$

where the last step is justified since the expectation operator applies only to the random fields, $\delta(\mathbf{x})$ and $\delta(\mathbf{y})$.

Let $\mathbf{r} = \mathbf{y} - \mathbf{x}$ then $\langle \delta(\mathbf{x})\, \delta(\mathbf{y}) \rangle \rightarrow \langle \delta(\mathbf{x})\, \delta(\mathbf{r} + \mathbf{x}) \rangle$. From above, $\xi(r) = \langle \delta(\mathbf{x})\, \delta(\mathbf{r} + \mathbf{x}) \rangle$. Therefore,

$$\begin{aligned}
\langle \delta(\mathbf{k}')\, \delta(\mathbf{k}) \rangle &= \int \int \langle \delta(\mathbf{x})\, \delta(\mathbf{r} + \mathbf{x}) \rangle \; e^{-i\,\mathbf{k}'\cdot\mathbf{x}} \; e^{-i\,\mathbf{k}\cdot(\mathbf{r}+\mathbf{x})} d\mathbf{x}\, d\mathbf{r} \\
&= \int \int \langle \delta(\mathbf{x})\, \delta(\mathbf{r} + \mathbf{x}) \rangle \; e^{-i\,(\mathbf{k}'+\mathbf{k})\cdot\mathbf{x}}\, d\mathbf{x}\, e^{-i\,\mathbf{k}\cdot\mathbf{r}} d\mathbf{r} \\
&= \int \int \xi(r) \; e^{-i\,(\mathbf{k}'+\mathbf{k})\cdot\mathbf{x}}\, d\mathbf{x}\, e^{-i\,\mathbf{k}\cdot\mathbf{r}} d\mathbf{r} \\
&= \int e^{-i\,(\mathbf{k}'+\mathbf{k})\cdot\mathbf{x}}\, d\mathbf{x} \int \xi(r)\, e^{-i\,\mathbf{k}\cdot\mathbf{r}} d\mathbf{r}\,.
\end{aligned}$$

The first term on the right side is a representation of the Dirac Delta in Eq. (2.7.20). Therefore, the expression may be written

$$\begin{aligned}
\langle \delta(\mathbf{k})\, \delta(\mathbf{k}') \rangle &= (2\pi)^3 \delta^3(\mathbf{k}' + \mathbf{k}) \int \xi(r)\, e^{-i\,\mathbf{k}\cdot\mathbf{r}} d\mathbf{r} \\
&= (2\pi)^3 \delta^3(\mathbf{k}' + \mathbf{k}) P_\delta(k) \qquad (2.7.23)
\end{aligned}$$

[77, Eq. (6.18)] where the power spectrum is defined as $P_\delta(k) \equiv \int \xi(r)\, e^{-i\,\mathbf{k}\cdot\mathbf{r}} d\mathbf{r}$. This term is difficult to evaluate directly. However, the expression $\langle \delta(\mathbf{k})\, \delta(\mathbf{k}') \rangle = (2\pi)^3 \delta^3(\mathbf{k}' + \mathbf{k}) P_\delta(k)$ means that the terms in the determination of the expectation are zero unless $k' = -k$. Therefore, applying the reality condition, the power spectrum may be written

$$\begin{aligned}
P_\delta(k) &= \langle \delta(\mathbf{k})\, \delta(\mathbf{k}-) \rangle \\
&= \langle \delta(\mathbf{k})\, \delta^*(\mathbf{k}) \rangle \\
&= \left\langle \left| \delta(\mathbf{k}) \right|^2 \right\rangle \qquad (2.7.24)
\end{aligned}$$

[83, Eq. (21.41)]. The final version of the two-point correlator in k-space may be written

$$\langle \delta(\mathbf{k})\, \delta(\mathbf{k}') \rangle = (2\pi)^3 \delta^3(\mathbf{k}' + \mathbf{k}) P_\delta(k) \qquad (2.7.25)$$

where the power spectrum is $P_\delta(k) \equiv \left\langle \left|\delta(\mathbf{k})\right|^2 \right\rangle$. The determination of $\left\langle \left|\delta(\mathbf{k})\right|^2 \right\rangle$ varies by context. See Ref. [83] p. 509 for an explanation on a galactic scale. See §3.2 *Quantum Origin of Perturbations* for an interpretation at the epoch of Hubble exit during inflation.

The two-point correlator in position space is $\xi(r) = \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}) \rangle$ (Eq. (2.7.16)). Expanding this expression using the inverse Fourier transform yields

$$\begin{aligned}
\xi(r) &= \langle \delta(\mathbf{x})\, \delta(\mathbf{r}+\mathbf{x}) \rangle \\
&= \left\langle \frac{1}{(2\pi)^3} \int \delta(\mathbf{k}')\, e^{i\,\mathbf{k}'\cdot\mathbf{x}} d\mathbf{k}' \frac{1}{(2\pi)^3} \int \delta(\mathbf{k})\, e^{i\,\mathbf{k}\cdot(\mathbf{r}+\mathbf{x})} d\mathbf{k} \right\rangle \\
&= \frac{1}{(2\pi)^6} \left\langle \int \int \delta(\mathbf{k}')\, \delta(\mathbf{k})\, e^{i\,(\mathbf{k}'+\mathbf{k})\cdot\mathbf{x}}\, d\mathbf{k}'\, e^{i\,\mathbf{k}\cdot\mathbf{r}}\, d\mathbf{k} \right\rangle . \qquad (2.7.26)
\end{aligned}$$

Since the expectation operator applies only to random fields, the operator may be moved inside the integrand. This step is followed by the application of Eq. (2.7.25). Therefore, the two-point correlator in position space may be written

$$\begin{aligned}
\xi(r) &= \frac{1}{(2\pi)^6} \int \int \langle \delta(\mathbf{k}')\, \delta(\mathbf{k}) \rangle\; e^{i\,(\mathbf{k}'+\mathbf{k})\cdot\mathbf{x}}\, d\mathbf{k}'\, e^{i\,\mathbf{k}\cdot\mathbf{r}}\, d\mathbf{k} \\
&= \frac{1}{(2\pi)^6} \int \int (2\pi)^3 \delta^3(\mathbf{k}'+\mathbf{k})\, P_\delta(k)\, e^{i\,(\mathbf{k}'+\mathbf{k})\cdot\mathbf{x}}\, d\mathbf{k}'\, e^{i\,\mathbf{k}\cdot\mathbf{r}}\, d\mathbf{k} \\
&= \frac{1}{(2\pi)^3} \int \int \delta^3(\mathbf{k}'+\mathbf{k})\, e^{i\,(\mathbf{k}'+\mathbf{k})\cdot\mathbf{x}}\, d\mathbf{k}'\, P_\delta(k)\, e^{i\,\mathbf{k}\cdot\mathbf{r}}\, d\mathbf{k} \\
&= \frac{1}{(2\pi)^3} \int P_\delta(k)\, e^{i\,\mathbf{k}\cdot\mathbf{r}}\, d\mathbf{k} . \qquad (2.7.27)
\end{aligned}$$

Since the correlator is real-valued, the imaginary part of $e^{i\,\mathbf{k}\cdot\mathbf{r}}$ may be eliminated. By applying spherical coordinates with $d\Omega = \sin\theta\, d\theta\, d\phi$, the correlator may be written

$$\begin{aligned}
\xi(r) &= \frac{1}{(2\pi)^3}\int P_\delta(k)\, e^{i\,\mathbf{k}\cdot\mathbf{r}}\, d\mathbf{k} \\
&= \frac{1}{(2\pi)^3}\int_0^\infty P_\delta(k)\left[\int_\Omega \cos\left(kr\,\cos\theta\right) d\Omega\right] k^2 dk \\
&= \frac{1}{(2\pi)^3}\int_0^\infty P_\delta(k)\left[\frac{4\pi\sin(kr)}{kr}\right] k^2 dk \\
&= \frac{1}{2\pi^2}\int_0^\infty P_\delta(k)\,\frac{\sin(kr)}{kr}\, k^2 dk
\end{aligned} \tag{2.7.28}$$

[83, Eq. (21.40)]. To simplify the expression, it is common to re-define the power spectrum as $\mathcal{P}_\delta(k) = \frac{k^3}{2\pi^2}P_\delta(k)$ yielding

$$\begin{aligned}
\xi(r) &= \frac{1}{2\pi^2}\int_0^\infty P_\delta(k)\,\frac{\sin(kr)}{kr}\, k^2 dk \\
&= \frac{1}{2\pi^2}\int_0^\infty \frac{2\pi^2}{k^3}\frac{k^3}{2\pi^2}P_\delta(k)\,\frac{\sin(kr)}{kr}\, k^2 dk \\
&= \int_0^\infty \mathcal{P}_\delta(k)\,\frac{\sin(kr)}{kr}\,\frac{dk}{k}
\end{aligned} \tag{2.7.29}$$

[77, p. 89].

If $\mathbf{r}$ is the null vector, then Eqs. (2.7.28) and (2.7.29) produce the following equivalent expressions for the variance [77, Eq. (6.26)] of $\delta(\mathbf{x})$ (the mean-square value when the mean is zero):

$$\sigma_\delta^2(\mathbf{x}) = \left\langle \delta^2(\mathbf{x})\right\rangle = \frac{1}{2\pi^2}\int_0^\infty P_\delta(k)\, k^2 dk \tag{2.7.30}$$

and

$$\sigma_\delta^2(\mathbf{x}) = \left\langle \delta^2(\mathbf{x})\right\rangle = \int_0^\infty \mathcal{P}_\delta(k)\,\frac{dk}{k}\,. \tag{2.7.31}$$

The right side shows that the variance is homogeneous since it is position independent. Therefore, the Gaussian distribution of the random field $\delta(\mathbf{x})$ is the same at every point in space. It was shown above that the mean value of a perturbation is zero. Therefore, the probability that the realization of the random field $\delta(\mathbf{x})$ will be in the interval $[a, b]$ is given by the normal form of the Gaussian distribution:

$$Pr[a \leq \delta(\mathbf{x}) \leq b] = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-y^2/2\sigma^2} d\,y \tag{2.7.32}$$

where $\sigma = \sigma_\delta^2$ is determined by Eq. (2.7.31).

The variance may be smoothed to wavelength $\lambda = 1/k$ by setting the upper limit of the integral in Eq. (2.7.31) to k:

$$\sigma_\delta^2(k, \mathbf{x}) = \int_0^k \mathcal{P}_\delta(k') \frac{dk'}{k'} . \tag{2.7.33}$$

This expression omits all terms with wavelengths smaller than $\lambda = 1/k$. This is the Fourier counterpart to smoothing a perturbation to a region of scale $R = \lambda = 1/k$.

Eq. (2.7.31) may be expressed on a logarithmic scale as

$$\sigma_\delta^2(\mathbf{x}) = \int_{-\infty}^{\infty} \mathcal{P}_\delta(\ln k)\, d\ln k \tag{2.7.34}$$

since $u = \ln k \rightarrow d\ln k = \frac{dk}{k}$. Therefore, $\mathcal{P}_\delta(\ln k)$ is the contribution to $\sigma_\delta^2(\mathbf{x})$ per interval $d\ln k$. The logarithmic scale plays an important role in deriving expressions for the power spectra at their origin during inflation.

If a perturbation is nearly scale invariant, and the variance of the perturbation is smoothed to large scale (small k), then the integral may be approximated within a box larger than the smoothing scale [77]. Selecting a box that is e times larger than $1/k$ yields

a simple and useful expression:

$$\begin{aligned}\sigma^2_\delta(\ln k, \mathbf{x}) &\simeq \mathcal{P}_\delta(\ln k) \int_{\ln k/e}^{\ln k} d\ln k \\ &= \mathcal{P}_\delta(\ln k) \int_{\ln k-1}^{\ln k} d\ln k \\ &= \mathcal{P}_\delta(\ln k) \end{aligned} \tag{2.7.35}$$

[77, Eq. (6.24)]. This approximation is valid for most co-moving scales accessible to observation that begin at about $1/k \simeq 5Mpc$.[12]

The Ergodic Property. The ergodic property for stationary Gaussian random fields (p. 58) states that correlators (ensemble averages weighted by probability densities) are equal to spatial averages determined in a single realization within the ensemble.

Spatial average over all points in a single realization. The spatial average within the single realization, $\delta_{\mathcal{R}}(\mathbf{x})$, in the ensemble is determined based on the perturbation of a scalar function, $f(\mathbf{x})$, in position space. Since real measurements can only be made in regions of non-zero scale, the smallest such region may be designated as $1/k$. An ordinary perturbation is defined as $\delta f(k, \mathbf{x}) = f(k, \mathbf{x}) - f$ where f is the spatial average $f = \overline{f(k, \mathbf{x})}$. The average perturbation over a finite set of n non overlapping regions of this scale that covers the co-moving scale of the realization of the observable universe is written

$$\overline{\delta_{\mathcal{R}}(\mathbf{x})} = \overline{\delta f(k, \mathbf{x})} = \frac{\sum_n \delta f(k, \mathbf{x_i})}{n} = \frac{\sum_n \left(f(k, \mathbf{x_i}) - f\right)}{n} = \frac{\sum_n f(k, \mathbf{x_i})}{n} - \frac{nf}{n} = f - f = 0\,. \tag{2.7.36}$$

Similarly, if $\delta(\mathbf{x})$ represents the contrast form of a perturbation, then

$$\overline{\delta_{\mathcal{R}}(\mathbf{x})} = \overline{\left(\frac{\delta f(k, \mathbf{x})}{f}\right)} = \overline{\left(\frac{f(k, \mathbf{x}) - f}{f}\right)} = \frac{\overline{f(k, \mathbf{x})}}{f} - 1 = \frac{f}{f} - 1 = 0\,.$$

Ensemble average at single point. $\delta_{\mathcal{R}}(\mathbf{x})$ represents a realization of the random field evaluated at $\mathbf{x}$, i.e., $\delta(\mathbf{x})$. The set

$$\{\, \delta_{\mathcal{R}}(\mathbf{x}) = y \,\}_{\mathbf{x}\in\mathbb{R}^3}$$

where $y \in \mathbb{R}$ is the realization at all points in space and is, therefore, a real-valued function over $\mathbb{R}^3$. Since its values are random at each point, the function is not likely to be a continuous function. However, if all realizations in the ensemble are considered, then all continuous functions are included in the ensemble. Although the range of the values in the realized perturbation is expected to be small, $[-\epsilon \le y \le \epsilon]$, nothing prevents its taking values over the entire real number line. The spatial average of every realization vanishes as shown above. Thus, $\overline{\delta_{\mathcal{R}}(\mathbf{x})} = 0$. The cardinality of the set is $\beth_1$, which is the cardinality of the set of real numbers, complex numbers, a continuous subset of real numbers, finite dimensional real or complex number spaces (e.g., $\mathbb{R}^3$ and $\mathbb{C}^3$), and, more generally, the cardinality of the continuum.

The realization space, Ω, includes all realizations: $\{\, \delta_{\mathcal{R}}(\mathbf{x}) = y \,\}_{\mathbf{x}\in\mathbb{R}^3}$. Therefore, the realization space may be written

$$\Omega = \left\{ \{\, \delta_{\mathcal{R}}(\mathbf{x}) \,\}_{\mathbf{x}\in\mathbb{R}^3} \right\}_{\mathcal{R}} \quad .$$

The cardinality of Ω is $\beth_2$ since it represents all orderings of real numbers, i.e., it is the set of all real valued functions (continuous functions and functions with any number of discontinuities) over $\mathbb{R}^3$. $\beth_2$ is the cardinality of the infinite dimensional real number space, $\mathbb{R}^{\mathbb{R}}$. The cardinality of subsets of Ω, such as $\Omega_x \subset \Omega$, including only values at a specific position $\mathbf{x}$, is also $\beth_2$.

Thus, the ensemble average, $\langle \delta(\mathbf{x}) \rangle$, is taken over $\mathbb{R}^{\mathbb{R}}$ realizations with cardinality $\beth_2$ at a single point in $\mathbb{R}^3$, while $\overline{\delta_{\mathcal{R}}(\mathbf{x})}$ is averaged over all points in $\mathbb{R}^3$ with cardinality $\beth_1$. Since all orderings are included, all values of $\delta(\mathbf{x})$ at all points in space are included in $\langle \delta(\mathbf{x}) \rangle$ at a single point in space, $\mathbb{R}^{\mathbb{R}}$ realizations each realization with the probability density of the realization, $\rho(\delta_{\mathcal{R}})$. Thus, $\overline{\delta_{\mathcal{R}}(\mathbf{x})} = 0$ requires the ensemble average in Eq. (2.7.14) to be

zero:

$$\langle \delta(\mathbf{x}) \rangle = \int_{\Omega_\mathbf{x}} \delta_\mathcal{R}(\mathbf{x})\, \rho(\delta_\mathcal{R})\, d\, \delta_\mathcal{R}(\mathbf{x}) = 0\,. \tag{2.7.37}$$

Spatial average of all pairs of points with fixed separation in a single realization. The spatial average that corresponds to the two-point correlator is derived as follows. The first step defines the spatial average and the next three steps apply inverse Fourier transformation to the terms in the product and re-arrange the result to form the Dirac Delta.

$$\begin{aligned}
\overline{\delta(\mathbf{y})\delta(\mathbf{y}+\mathbf{x})} &= \frac{1}{(2\pi)^3} \int \delta(\mathbf{y})\delta(\mathbf{y}+\mathbf{x}) d\mathbf{y} \\
&= \frac{1}{(2\pi)^9} \int \left[\int d\mathbf{k}'\, \delta(\mathbf{k}') e^{i\,\mathbf{k}'\cdot\mathbf{y}} \right] \left[\int d\mathbf{k}\, \delta(\mathbf{k}) e^{i\,\mathbf{k}\cdot(\mathbf{y}+\mathbf{x})} \right] d\mathbf{y} \\
&= \frac{1}{(2\pi)^6} \int \int \left[\frac{1}{(2\pi)^3} \int d\mathbf{y}\, e^{i\,(\mathbf{k}'+\mathbf{k})\cdot\mathbf{y}} \right] \delta(\mathbf{k}')\, \delta(\mathbf{k})\, e^{i\,\mathbf{k}\cdot\mathbf{x}} d\mathbf{k}' d\mathbf{k} \\
&= \frac{1}{(2\pi)^6} \int \int \delta^3(\mathbf{k}'+\mathbf{k})\, \delta(\mathbf{k}')\, \delta(\mathbf{k})\, e^{i\,\mathbf{k}\cdot\mathbf{x}} d\mathbf{k}' d\mathbf{k}\,.
\end{aligned} \tag{2.7.38}$$

The terms $\delta(\mathbf{k}')$ and $\delta(\mathbf{k})$ in Eq. (2.7.38) are uncorrelated except as required by the reality condition. Since the product, $\delta(\mathbf{k}')\,\delta(\mathbf{k})$, is contained within the double integral over all values of $\mathbf{k}'$ and $\mathbf{k}$, the correlator $\langle \delta(\mathbf{k}')\,\delta(\mathbf{k}) \rangle$ may be substituted for product within the integral [77]. Applying Eq. (2.7.25) and simplifying the result yields:

$$\begin{aligned}
\overline{\delta(\mathbf{y})\delta(\mathbf{y}+\mathbf{x})} &= \frac{1}{(2\pi)^6} \int \int \left[\delta^3(\mathbf{k}'+\mathbf{k}) \right] \langle \delta(\mathbf{k}')\, \delta(\mathbf{k}) \rangle\; e^{i\,\mathbf{k}\cdot\mathbf{x}} d\mathbf{k}' d\mathbf{k} \\
&= \frac{1}{(2\pi)^6} \int \int \delta^3(\mathbf{k}'+\mathbf{k})\, (2\pi)^3 \delta^3(\mathbf{k}'+\mathbf{k}) P_\delta(k)\, e^{i\,\mathbf{k}\cdot\mathbf{x}} d\mathbf{k}' d\mathbf{k} \\
&= \frac{1}{(2\pi)^3} \int \int \left(\delta^3(\mathbf{k}'+\mathbf{k}) \right)^2 d\mathbf{k}'\, P_\delta(k)\, e^{i\,\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{k} \\
&= \frac{1}{(2\pi)^6} \int P_\delta(k)\, e^{i\,\mathbf{k}\cdot\mathbf{x}} \left[\int d\mathbf{k}' \left(\delta^3(\mathbf{k}'+\mathbf{k}) \right)^2 \right] d\mathbf{k} \\
&= \frac{1}{(2\pi)^3} \int P_\delta(k)\, e^{i\,\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{k}
\end{aligned} \tag{2.7.39}$$

[77, Eq. (6.42)]. The equivalence of the position space two-point correlator in Eq. (2.7.27) to the spatial average in Eq. (2.7.39) completes the demonstration of the ergodic property of stationary Gaussian random fields. For additional information see Appendix D, *The Ergodic Theorem*, in Ref. [106].

Physical Interpretation of the Power Spectrum. The power spectrum is the contribution to the variance of the power or intensity of a quantity (for example light) by components of the quantity. For example, light received from a light source contains a range of frequencies. These frequencies (or equivalently, wavelengths) are the components of light.

The power spectrum of the CMB anisotropies in Figure 2.5 represents the contribution to the variance of anisotropies (measured as temperature fluctuations) by regions of co-moving scales, beginning at the scale of the observable universe down to the smallest observable region corresponding to about $5Mpc$. Regions of the CMB correspond to angular scales that are represented by multipole moments where $\theta \sim 1/\ell$.

The power spectrum of primordial perturbations is the contribution to the variance of the intensity of the perturbation by regions of scales represented by the wavelength λ. Scales may be either physical scales or co-moving scales; however, co-moving scales are used in this exposition. In the Fourier expansions described above, scales are co-moving wavelengths expressed in terms of co-moving wavenumbers where $\lambda = 1/k$.

Consider the representation of the variance of a generic perturbation in Eq.(2.7.30):

$$\sigma_\delta^2(\mathbf{x}) = \frac{1}{2\pi^2}\int_0^\infty P_\delta(k)\, k^2 dk\,.$$

The definition of the power spectrum is $P_\delta(k) = \left\langle \left|\delta(k)\right|^2 \right\rangle$ where $\left\langle \left|\delta(k)\right|^2 \right\rangle$ is the ensemble average of the random field $\left|\delta(k)\right|^2$, or by the ergodic property, the spatial average of $\left|\delta(k)\right|^2$ in one realization of the random field over a fair sample of space. This expression is

equivalent to

$$\sigma_\delta^2(\mathbf{x}) = \int_0^\infty \frac{P_\delta(k)}{(2\pi)^3}\, k^2\, dk$$
$$= \int_0^\infty \frac{\left\langle \left|\delta(k)\right|^2 \right\rangle}{(2\pi)^3}\, 4\pi\, k^2\, dk\,. \tag{2.7.40}$$

The integrand is the contribution to the variance of the perturbation produced by wavenumbers in the infinitesimal range $[k, k + dk]$ over a thin spherical shell with radius k and infinitesimal thickness dk in k-space.

The integral

$$\int_{k_1}^{k_2} \frac{\left\langle \left|\delta(k')\right|^2 \right\rangle}{(2\pi)^3}\, 4\pi\, k'^2\, dk'$$

is an expression for a spherical shell with inner radius k_1 and outer radius k_2 in k-space. The shell "contains" the contributions to total variance in Eq. (2.7.40) produced by wavenumbers from k_1 to k_2. If $k_1 = 1/R_0$, the scale of the observable universe, and $k_2 = 1/R_{10^4\odot}$, the scale of the smallest region that contributes to scalar perturbations (p. 52), then the resulting finite spherical shell in k-space is equivalent to the total variance in Eq. (2.7.40).

In epochs during which these concepts are well defined by classical physics such as $t_\dagger \simeq 15$ seconds (the epoch of primordial perturbations), the physical meaning of k-space is the "space" of scales of spherical regions. Note that at $t_\dagger$, the observable universe was a hot soup simmering at about one billion degrees in a heaping bowl with diameter of about 250 light years; certainly large enough and still cool enough for classical physics.[13]

Section Summary. Observations are consistent with Gaussian distribution of $\zeta(\mathbf{x})$ and its closely related value $\delta\rho(\mathbf{x})/\rho$. In single field slow-roll models of inflation, the amplitudes in the Fourier transformation of perturbations of the field, $\delta\phi(\mathbf{x})$, are caused by quantum fluctuations described in momentum space, $\delta\phi(\mathbf{k})$, which are random and are therefore

uncorrelated, resulting in Gaussian distribution of $\delta\phi(\mathbf{x})$. All these perturbations will be treated as stationary Gaussian random fields.

The statistical properties of stationary Gaussian random fields may be modeled using the position ensemble average to find the vanishing mean, the two-point momentum correlator in Eq. (2.7.25) to determine the power spectrum, and the two-point position correlator in Eq. (2.7.31) to determine the variance. The anisotropies in the CMB are modeled using spherical harmonic expansion instead of Fourier expansion. See §B.11 *Sky Map Expansion in Spherical Harmonics* for additional information.

By the ergodic property of stationary Gaussian random fields, the statistical properties of observable perturbations, such as number and mass densities of galaxies and galaxy clusters and anisotropies in the CMB, may be estimated using spatial averages measured over a large number of samples.

Scalar perturbation models that include higher order correlators and their association spectra are used to examine the statistical properties of non Gaussian scalar perturbations. For example, the three-point correlator yields the bispectrum, $B_\delta(\mathbf{k_1}, \mathbf{k_3}, \mathbf{k_3})$ and the four-point correlator yields the trispectrum, $T_\delta(\mathbf{k_1}, \mathbf{k_3}, \mathbf{k_3}, \mathbf{k_4})$. See References [83, 77, 75, 106] for additional information. Non Gaussian scalar perturbations are not favored by current observations but have not been ruled out. See Ref. [3] *Planck 2015 results. XVII. Constraints on primordial non-Gaussianity* for additional information.

Section §3.2 will present the theory that these perturbations are based on quantum fluctuations in the *inflaton field* at the epoch of Hubble exit during inflation. The quantum fluctuations of the *inflaton field* are conserved until the end of inflation to produce adiabatic Gaussian curvature perturbations that induce energy density perturbations in the plasma of ordinary particles that form after the end of inflation. Therefore, observations of Gaussian adiabatic scalar perturbations are consistent with single field models of inflation that are the focus of the remainder of this chapter and Chapter 3.

2.7.4 Representation of Perturbations with Cosmological Parameters

The power spectra of curvature perturbations, $\mathcal{P}_\zeta(k)$, for the relevant range of wavenumbers are incorporated into computer codes used to simulate the evolution of CMB anisotropies for comparison with observed anisotropies. The codes used by the Planck Collaboration are based on CAMB codes. Through this process it has been determined that $\mathcal{P}_\zeta(k)$ is nearly scale invariant. The variations in $\mathcal{P}_\zeta(k)$ are developed based on the curvature perturbation power spectrum approximated by the power law expression

$$\mathcal{P}_\zeta(k) = A_s \left(\frac{k}{k_*}\right)^{n_s - 1} \tag{2.7.41}$$

[87, Eq. (2)] where the scalar mode perturbation amplitude parameter, A_s, is equal to $\mathcal{P}_\zeta(k_*)$ evaluated at the pivot scale, k_*. The primary pivot scale used by the Planck Collaboration is $k_* = 0.05\, Mpc^{-1}$. This wavenumber corresponds to the multipole moment $\ell \simeq 700$ where ℓ is equal to the co-moving radial distance to the LSS (Eq. (2.3.13) converted to units of Mpc) multiplied by the wavenumber [77]. This value is roughly in the middle of the logarithmic range of multipole moments in the representation of the power spectrum of temperature fluctuations reported by the Planck Collaboration as shown in Figure 2.5.

It is evident from the expression above that the power spectrum of the curvature perturbation is "flat" (or invariant) if $n_s = 1$. Slight departure from a flat power spectrum is an important prediction of inflationary cosmology [65]. For this reason, n_s is frequently called the "primordial tilt". Current observations find that $n_s \simeq 0.96$ and that $n_s \neq 1$ to very high confidence (5σ) [88]. The fitting of A_s and n_s with other parameters to observations is described in §B.15 *Precision Cosmology*. The relationship of these cosmological parameters to the parameters of Slow-Roll Approximation is described in §3.2 *Quantum Origin of Perturbations*.

In some extended models, n_s is a function of k. This is called a running spectral index where

$$\mathcal{P}_\zeta(k) = A_s \left(\frac{k}{k_*}\right)^{n_s - 1 + (1/2)(dn_s/d\ln k)\ln(k/k_*)} \tag{2.7.42}$$

[87, Eq. (2)]. In Planck Collaboration reports, the parameter A_s is called the primordial curvature perturbation, n_s is called the scalar spectrum power law index, and $\mathcal{P}_\zeta(k)$ is the curvature perturbation power spectrum (using the notation $\mathcal{P}_\mathcal{R}(k)$).[14] The most recent Planck Collaboration report, *Planck 2015 results. XIII. Cosmological parameters* [88] finds that the data are consistent with no running spectral index but do not rule out a small negative value of $dn_s/d\ln k < -0.0126^{+0.0098}_{-0.0087}$.

Extensions to the base ΛCDM model include primordial gravitational waves that produce primordial tensor perturbations that are expected to produce the polarization patterns illustrated in Figure 2.7 in the CMB. These patterns have not been observed but the search for them is ongoing. In theory, gravitational waves produce tensor variations in the spatial metric as $\gamma_{ij}(\mathbf{x}) \equiv \left(Ie^{2h}\right)_{ij}$ where h_{ij} is traceless [77]. Tensor mode perturbations are expected to be unobservable except in large field models of inflation [65].

In Planck Collaboration simulations, the power spectrum of tensor mode perturbations is approximated using a power law expression parameterized by the tensor mode perturbation amplitude parameter, A_t, and a tensor mode spectrum power-law index, n_t:

$$\mathcal{P}_t(k) = A_t \left(\frac{k}{k_*}\right)^{n_t} \tag{2.7.43}$$

[87, Eq. (3)]. The ratio of the tensor mode perturbation amplitude parameter, A_t, to the scalar mode perturbation amplitude parameter, A_s, is reported as r labeled by the pivot scale. The Planck Collaboration reports two ratios. The first is $r_{0.05}$ based on $k_* = 0.05\,Mpc^{-1}$ and the second is $r_{0.002}$ based on $k_* = 0.002\,Mpc^{-1}$ (the ratio is more accurately measured at $k_* = 0.002\,Mpc^{-1} \equiv 500\,Mpc$, the larger of the two scales). A large value of

r would confirm the existence of gravitational waves. The most recent value, determined by the 2015 *Joint Analysis of BICEP2/ Keck Array and Planck Data* [3] is $r_{0.05} < 0.12$ with 95% confidence. The report describes this result as "no statistically significant evidence for tensor modes". The fitting of r, A_t, and n_t with other parameters is described in §B.15 *Precision Cosmology*. The relationships of these cosmological parameters to the parameters of Slow-Roll Approximation are described in §3.2 *Quantum Origin of Perturbations*.

Notes

[1]In cosmology it is common to work in units where the reduced Planck constant, the speed of light, and the Boltzmann constant are set to unity. In these units, Planck base units are $m_P = T_P = 1/\sqrt{G}$ and $\ell_p = t_p = m_P^{-1}$.

[2]The metric for a perfect fluid is diagonal with $g_{00} = -1$ in the signature $-+++$. The terms U^μ and U^ν are the four-velocity vectors $(1, 0, 0, 0)^T$. Therefore, the term $U_\rho = g_{\rho\sigma}U^\sigma = (-1, 0, 0, 0)$. The term $U_\mu U_\nu$ is a rank $(0, 2)$ tensor, $T_{\mu\nu}\, dx^\mu \otimes dx^\nu = U_\mu U_\nu\, dx^\mu \otimes dx^\nu$, with $T_{00} = 1$ and all other components set to zero.

[3]This line element does not capture all possible global structures. For example, the global structure of flat space may be a three-torus (a non simply connected compact space) instead of Euclidean space.

[4]The movement of the Sun relative to a co-moving observer has "peculiar motion" that is motion other than motion due to the expansion of space. Adjustments to redshift due to peculiar motion are required for accurate measurements.

[5]The rest energy of nucleons is much greater than the rest energy of the quarks that form the nucleons. The rest energy of nucleons includes the rest energy of the quarks, the kinetic energy of the motion of quarks within the nucleon, and the binding energy of the gluon field. The difference between the rest energy of a nucleon and the rest energy of the quarks within the nucleon is about $928MeV$ that corresponds to about $10^{13}K$. At temperatures above this level, the kinetic energy of quarks is too high to form nucleons.

[6]After decoupling, neutrinos began to cool more rapidly than the rest of the plasma. After about 30 minutes the ratio became constant with $T/T_\nu = 1.401$. After photon decoupling much later at about $3,000K$, the ratio of photon temperature to neutrino temperature remained at $T\gamma/T_\nu = 1.401$. This ratio continues through to the current era.

[7]Ignoring the unknown effects of dark matter and supersymmetry

[8]The theory of dark energy is, however, far from settled. Just as the inflation epoch ended, there could be a phase transition in the action of dark energy that could alter this behavior. Quintessence (the "fifth" force) is an alternate theory of dark energy in which ρ_Λ, ρ_Λ, and w_Λ are all dynamic. Under this theory, the universe could enter a phase of contraction and eventually collapse into a black hole or possible bounce into

another phase of expansion as proposed in cyclic models such as the Steinhardt-Turok model [98] which is an alternative to inflation.

[9]This was calculated by expressing a^2H^2 in terms of the Planck Collaboration density parameters and the Hubble constant in powers of scale similar to the procedure used to derive Eq. (2.3.4). Similar results can be obtained by assuming that the universe has always been dominated by radiation [60]. In this case, $\left|\Omega_{k0}\right| \propto t$ and at 10^{-11} seconds, $\left|\Omega_k\right| = 0.005/(10^{17}/10^{-11}) \simeq 10^{-30}$ where 10^{17} seconds is the present age of the universe.

[10]The evolution of the CMB from the LSS to the present time was disturbed by many processes including re-ionization of matter due to star formation that produced additional scattering of photons emitted by the LSS, and gravitation lensing as LSS photons pass near massive structures. These foreground disturbances need to be distinguished from the anisotropies present in the LSS.

[11]It should be noted that although causal connection co-moving distance is on the order of the co-moving Hubble horizon, the causal limit is the co-moving horizon denoted by D_H or equivalently by η.

[12]Multipole moment $\ell = 2,500$ is at the limit of current observation. This is approximately wavenumber $1/5Mpc^{-1} = 0.20Mpc^{-1}$ which corresponds to co-moving regions of scale $5Mpc$. The Planck Collaboration pivot scale $k_{0.05}$ is wavenumber $k = 0.05Mpc^{-1} = 1/20Mpc^{-1}$ which corresponds to co-moving regions of scale $20Mpc$. In multipoles, this is about $\ell = 700$ which is roughly in the center of the logarithmic scale of observed multipole moments.

[13]The temperature was selected to be well below the formation temperature for electrons, which is much lower than the formation temperature for light quarks let alone baryons; so the thermal energy of the plasma is cool enough for classical treatment. The physical scale is approximately proportional to temperature. The temperature at $t_\dagger$, when the thermal energy of the plasma was about 0.1 MeV, was a little more than 10^9K. The temperature is now a little less than 3 K. The current radius of the universe is about 46 billion light years which can be determined using Eq. (2.3.14). Putting it all together, the radius of the plasma at $t_\dagger$ was about $3/10^9 \times 4.6 \times 10^9 \sim 125$ light years.

[14]The notation $\mathcal{P}_\mathcal{R}(k)$ is used in the literature to include scalar perturbations with non-adiabacity and departature from Gaussian distribution. The Planck Collaboration retains this notation when describing adiabatic curvature perturbations with Gaussian distribution.

Chapter 3

Single Field Slow-Roll Theories of Inflation

This chapter describes a broad class of inflation models based on a single real valued scalar field called the *inflaton field* associated with an unknown spin-0 particle, called the *inflaton*. The field and the particle are usually denoted by the letter, ϕ. Each model is defined by the potential energy of the *inflaton field*, $V(\phi)$. An "effective" model of inflation:

1. specifies the initial conditions for inflation to commence,
2. produces sufficient expansion during inflation to drive space to flatness and the contents of the universe to near homogeneity,
3. specifies the conditions for inflation to end, and
4. produces perturbations in the field that are consistent with the subsequent evolution of the observed anisotropies in the CMB and the subsequent evolution of cosmological structures.

The first three requirements are formalized in the first section of this chapter as the Slow-Roll Approximation (the "SRA") by defining two slow-roll parameters, ϵ and η. The cosmological parameters described in §2.7.4 *Representation of Perturbations with Cosmolog-*

ical Parameters, A_s, n_s, A_t, n_t, and r, are expressed in terms of the potential energy $V(\phi)$ of the *inflaton field* and the slow-roll parameters, ϵ and η. The second section of this chapter shows that the required perturbations are caused by quantum fluctuations of the *inflaton field* during inflation.

SRA conditions may be expressed in terms of either potential energy or the Hubble parameter. The discussion in this paper is based on the potential energy formalism of the SRA following [77, §18.5]. For information on the Hubble parameter formalism see Ref. [63].

3.1 The Slow-Roll Approximation

Section Overview. The central feature in slow-roll models of inflation is that the potential energy of the *inflaton field* dominates the energy density of the field and changes slowly throughout a significant part of the inflation epoch. Ordinary fields such as matter and radiation fields also exist during this period but have negligible energy compared to the *inflaton field*. In later stages of inflation, potential energy decreases as the kinetic energy of the *inflaton field* increases. Once the energy of the *inflaton field* is no longer sufficiently dominated by potential energy, inflation ends.

Two SRA parameters, ϵ and η, will be defined and their corresponding conditions will be developed in this section. The conditions are both "flatness conditions" [77] placed on the potential energy of the *inflaton field*. The first condition, $\epsilon \ll 1$, provides sufficiently flat slope in the graph of $V(\phi)$ over a range in ϕ. The second condition, $|\eta| \ll 1$, provides that sufficiently flat slope is maintained long enough to produce adequate inflation. Inflation begins when the slow-roll conditions are met and inflation ends when the slow-roll conditions no longer hold.

The potential energy curve of a toy slow-roll model of inflation is illustrated in Figure 3.1. $V(\phi)$ begins nearly flat at point a. During this period space expands nearly exponen-

tially. As the scalar field evolves in the direction of point *b* and beyond, inflation continues at reducing rates of acceleration. Inflation ends when the potential energy of the field begins oscillating about the local minimum at point *d*. During the oscillation of the potential energy, the *inflaton field* decays. As the field decays, the kinetic energy gained by the *inflaton field* is transferred as thermal energy to ordinary fields in the vacuum, exciting the fields to produce a very hot plasma of ordinary elementary particles and, if they exist, supersymmetric particles. One or more of these ordinary particles or supersymmetric particles may be dark matter particles. The period of oscillation corresponds to the periods of pre-heating and re-heating. Additional information on the physics of pre-heating and re-heating may be found in references [106, 74, 77, 68].

The potential energy curve of the toy model is an example of a concave model (or concave down) reflecting the shape of the graph during the rapid period of inflationary expansion. Concave models are favored by the 2015 report, *Joint Analysis of BICEP2/ Keck Array and Planck Data* [3]. See Figure 4.5.

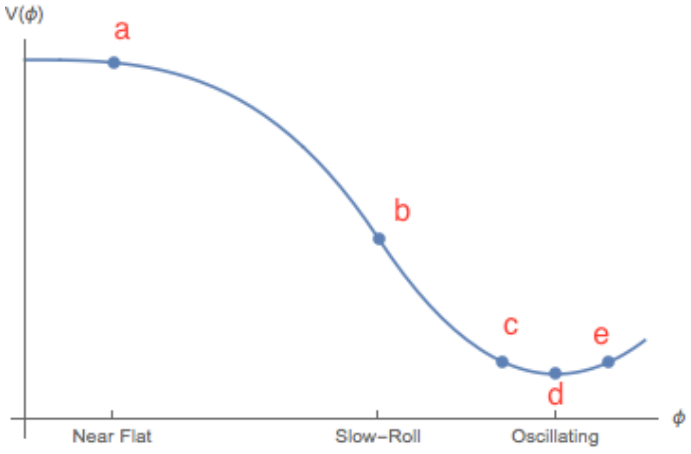


Figure 3.1: The Potential Energy Curve Toy Model of Slow-Roll Inflation

The change in the scale of space during any epoch beginning at time t_1 and ending at time t_2 is a function of the Hubble parameter:

$$a(t_{end})/a(t_{begin}) = \int_{t_1}^{t_2} H(t)dt\,. \tag{3.1.1}$$

The energy density of any scalar field including the *inflaton field* is

$$\rho(\phi) = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}g^{ij}\partial_i\,\phi\,\partial_j\,\phi + V(\phi)\,.$$

See §B.4 *The Stress Energy Tensor of a Scalar Field*.

During the period that includes point *a*, the potential energy of the *inflaton field* dominates the gradient and kinetic energy terms of the *inflaton field* and all other components of energy, including the fictitious curvature energy. At these times, the Hubble parameter, a function of energy density, may be written

$$H(\phi) = \sqrt{\frac{8\pi\,G}{3}\rho(\phi)} \simeq \sqrt{\frac{8\pi\,G}{3}V(\phi)}\,.$$

As rapid inflation drives space to flatness (with density parameter $\Omega \simeq 1$) and the *inflaton field* to homogeneity and isotropy, the field becomes a nearly perfect fluid in flat space. The energy density of a homogeneous scalar field in flat space (Eq. (B.4.8)) is

$$\rho = \frac{1}{2}\dot{\phi}^2 + V(\phi) \tag{3.1.2}$$

with zero gradient energy. The equation of motion of a homogeneous scalar field (Eq. (B.4.4)) is

$$\ddot{\phi} + 3\,H\,\dot{\phi} + V_\phi = 0 \tag{3.1.3}$$

where $V_\phi \equiv \frac{dV}{d\phi}$. This notation and extensions to higher order partial derivatives is consistent with Planck Collaboration reports and other current literature. The Hubble parameter is a friction term in the equation of motion, causing the value of $V(\phi)$ to change slowly.

Based on Eq. (3.1.2), the expansion of space, described by the Friedmann equation in Chapter 2, may be written

$$\frac{3}{8\pi G}H^2 = \frac{1}{2}\dot{\phi}^2 + V(\phi)\,. \tag{3.1.4}$$

In units where $\hbar = c = 1$, Planck Mass is $m_{Pl} = \sqrt{1/G}$ and the Reduced Planck Mass is $M_{Pl} = \sqrt{1/8\pi G}$. Using the latter notation, the expansion of space may be written

$$3\,M_{Pl}^2\,H^2 = \frac{1}{2}\dot{\phi}^2 + V(\phi)\,. \tag{3.1.5}$$

The time derivative of each side yields the following expression for the rate of change of the expansion of space:

$$6\,M_{Pl}^2\,H\,\dot{H} = \dot{\phi}\,\ddot{\phi} + V_\phi\,\dot{\phi}\,. \tag{3.1.6}$$

Substituting $\ddot{\phi} = -3\,H\dot{\phi} - V_\phi$ from Eq. (3.1.3) above into the preceding equation and solving for $\dot{H}$ yields

$$\dot{H} = -\frac{\dot{\phi}^2}{2\,M_{Pl}^2} \tag{3.1.7}$$

which shows that, based on the SRA, $\dot{H}$ is always negative during inflation.

In Eq. (2.5.1) it was determined that the inflation condition may be written

$$-\frac{\dot{H}}{H^2} < 1\,.$$

A slowly changing value of a dominating potential energy term yields slowly changing values of the Hubble parameters which produce inflation exit as illustrated in Figure 3.1. If H is slowly changing, then the magnitude of $\dot{H}$ is small and the inflation condition may be

written

$$\frac{\left|\dot{H}\right|}{H^2} \ll 1\,.$$

In order for the potential energy of the *inflaton field* to dominate energy during the productive part of inflation (the "near flat" part), it must be the case that $V(\phi) >> \frac{1}{2}\dot{\phi}^2$. Therefore, from Eq. (3.1.5), the equation describing the expansion of space is approximately

$$3\,M_{Pl}^2\,H^2 \simeq V(\phi) \tag{3.1.8}$$

and, equivalently,

$$H^2 \simeq \frac{V}{3\,M_{Pl}^2}\,. \tag{3.1.9}$$

3.1.1 Attractor Solutions

In addition to the flatness conditions imposed on the potential energy term defined by the parameters $\epsilon \ll 1$ and $\left|\eta\right| \ll 1$, the equation of motion of the *inflaton field* must be an "attractor solution". This means that the dynamics of ϕ must cause the ϕ to be attracted to the value at the local minimum of the potential energy that ends inflation. An attractor solution satisfies the attractor equation derived in Eq. (3.1.14) below.

The time derivative of the expansion of space in Eq. (3.1.8) is

$$6\,M_{Pl}^2\,H\dot{H} \simeq V_\phi\,\dot{\phi}\,. \tag{3.1.10}$$

By substituting $\dot{H} = -\frac{\dot{\phi}^2}{2\,M_{Pl}^2}$, the equation for the time derivative of the expansion of space may be written

$$6\,M_{Pl}^2\,H\left(-\frac{\dot{\phi}^2}{2\,M_{Pl}^2}\right) \simeq V_\phi\,\dot{\phi} \tag{3.1.11}$$

which reduces to

$$3H\dot{\phi} \simeq -V_{\phi}\,. \tag{3.1.12}$$

This equation may be rearranged into two useful forms:

$$H \simeq -\frac{V_{\phi}}{3\dot{\phi}} \tag{3.1.13}$$

and

$$\dot{\phi} \simeq -\frac{V_{\phi}}{3H} \tag{3.1.14}$$

where Eq. (3.1.14) is called the "attractor equation". Referring to the example in Figure 3.1, it is evident that $V_{\phi} < 0$ at points a, b, c and $V_{\phi} > 0$ at point e. Since $H > 0$, the sign of $\dot{\phi}$ is set by $-V_{\phi}$. The sign of $\dot{\phi}$ causes ϕ to increase at a, b, c and to decrease at e. Hence ϕ is said to be "attracted" to ϕ at d, the value of ϕ at the local minimum of $V(\phi)$. Equivalently, the value of the *inflaton field* changes over time toward values that result in a "graceful exit" from inflation.

3.1.2 The First SRA Condition

By substituting $H^2 \simeq \frac{V}{3\,M_{Pl}^2}$ from above into the square of the attractor equation, $\dot{\phi}^2$ may be written

$$\dot{\phi}^2 \simeq \frac{\left(V_{\phi}\right)^2\, M_{Pl}^2}{3\,V}\,. \tag{3.1.15}$$

By using Eq. (3.1.7), the absolute value of the time derivative of the Hubble parameter is

$$\begin{aligned} \left|\dot{H}\right| &\simeq \frac{\left(V_\phi\right)^2 M_{Pl}^2}{3\,V \cdot 2M_{Pl}^2} \\ &= \frac{\left(V_\phi\right)^2}{6\,V}\,. \end{aligned} \tag{3.1.16}$$

The ratio of Eq. (3.1.16) to Eq. (3.1.9) is

$$\begin{aligned} \frac{\left|\dot{H}\right|}{H^2} &\simeq \frac{\left(V_\phi\right)^2}{6\,V} \Big/ \frac{V}{3\,M_{Pl}^2} \\ &= \frac{M_{Pl}^2}{2}\left(\frac{V_\phi}{V}\right)^2\,. \end{aligned} \tag{3.1.17}$$

The first SRA parameter is defined as

$$\epsilon = \frac{M_{Pl}^2}{2}\left(\frac{V_\phi}{V}\right)^2\,. \tag{3.1.18}$$

Applying the inflation condition, $\frac{|\dot{H}|}{H^2} \ll 1$, yields the first SRA condition

$$\epsilon \ll 1\,. \tag{3.1.19}$$

3.1.3 The Second SRA Condition

The time derivative of Eq. (3.1.12) yields

$$3\left(\dot{H}\dot{\phi} + H\ddot{\phi}\right) = -V_{\phi\phi}\,\dot{\phi}\,.$$

Therefore $\ddot{\phi}$ may be written

$$\ddot{\phi} = -\frac{\dot{H}}{H}\dot{\phi} - \frac{V_{\phi\phi}}{3H}\,\dot{\phi}\,. \tag{3.1.20}$$

From Eq. (3.1.3) above, the exact equation of motion for a homogeneous scalar field is

$$\ddot{\phi} = -3\,H\,\dot{\phi} - V_{\phi}\,. \tag{3.1.21}$$

From Eq. (3.1.12), $V_{\phi} \simeq -3H\dot{\phi}$; therefore, in this approximation, Eq. (3.1.21) yields $\ddot{\phi} \simeq 0$ and Eq. (3.1.20) may be written

$$-\frac{\dot{H}}{H} = \frac{V_{\phi\phi}}{3H}\,. \tag{3.1.22}$$

After multiplying both sides of this equation by $1/H$ and substituting $H^2 = V/3M_{Pl}^2$ from Eq. (3.1.9) on the right side, Eq. (3.1.22) is equivalent to

$$\begin{aligned} -\frac{\dot{H}}{H^2} &= \frac{V_{\phi\phi}}{3\frac{V}{3M_{Pl}^2}} \\ &= M_{Pl}^2\frac{V_{\phi\phi}}{V}\,. \end{aligned} \tag{3.1.23}$$

Defining the second SRA parameter as

$$\eta = M_{Pl}^2\frac{V_{\phi\phi}}{V} \tag{3.1.24}$$

and applying the inflation condition, $\frac{|\dot{H}|}{H^2} \ll 1$, yields the second SRA condition

$$\left|\eta\right| \ll 1\,. \tag{3.1.25}$$

3.1.4 Observable Inflation and the Pivot Scale

The total expansion expressed in e-foldings from scale factor a_1 to scale factor a_2 is $e^N = \frac{a_2}{a_1}$. Therefore, $N = \ln \frac{a_2}{a_1}$. The definition of the Hubble parameter

$$\begin{aligned} H &= \frac{\dot{a}}{a} \\ &= \frac{1}{a}\frac{da}{dt} \end{aligned}$$

yields

$$H\,dt = \frac{1}{a}da$$

and by integration

$$\ln a = \int H\,dt\,.$$

The number of e-foldings from t_1 to later time, t_2, may be written

$$N = \ln \frac{a_2}{a_1} = \int_{t_1}^{t_2} H\,dt\,.$$

The ratio of Eq. (3.1.9) and Eq. (3.1.13) may be written

$$\begin{aligned} H &= H^2/H \\ &\simeq -\frac{V}{3\,M_{Pl}^2}\Big/\frac{V_\phi}{3\dot{\phi}} \\ &= -\frac{\dot{\phi}}{M_{Pl}^2}\frac{V}{V_\phi} \\ &= -\frac{d\phi}{dt}\frac{1}{M_{Pl}^2}\frac{V}{V_\phi}\,. \end{aligned}$$

By rearranging this result, the number of e-foldings from some time during inflation, t, to the time at the end of inflation, t_{end}, may be approximated under the SRA as

$$N \simeq \frac{1}{M_{Pl}^2} \int_{\phi_{end}}^{\phi} \frac{V}{V_\phi} d\phi \tag{3.1.26}$$

where $\phi = \phi(t)$ and $\phi_{end} = \phi(t_{end})$.

The co-moving region corresponding to wavenumber k_* is called the "pivot scale". The precise value is not important for the present discussion, but it will be assumed that k_*^{-1} is small enough to have re-entered the Hubble horizon. Note that $k_* = a_* H_*$ where a_* and H_* represent the values of the scale factor and the Hubble parameter at Hubble exit of k_* (i.e., at the first intersection in Figure 2.8). The pivot scale exits the Hubble horizon N_* e-foldings before the end of inflation:

$$N_* = \ln \frac{a_{end}}{a_*} .$$

From Eq. (3.1.26)

$$N_* \simeq \frac{1}{M_{Pl}^2} \int_{\phi_{end}}^{\phi_*} \frac{V}{V_\phi} d\phi \tag{3.1.27}$$

where ϕ_* is the value of the *inflaton field* at the time k_*^{-1} exits the Hubble horizon.[1] N_* is called *observable inflation*. (X_* is the notation used by the Planck Collaboration to represent X evaluated at Hubble exit during inflation of the region with wavenumber k_*.)

3.1.5 Constraining Models of Inflation

It is evident from the lower graph in Figure 2.8 that the co-moving Hubble horizon at Hubble exit is equal to the co-moving Hubble horizon at Hubble re-entry. Therefore, using the assumption that the Hubble parameter does not change much from Hubble exit to the

end of inflation, i.e., $H_*/H_{end} \simeq 1$, it follows that

$$\begin{aligned} 1 &= \frac{a_* H_*}{a_{re-entry} H_{re-entry}} \\ &= \frac{a_* H_*}{a_{end} H_{end}} \frac{a_{end} H_{end}}{a_{re-heat} H_{re-heat}} \frac{a_{re-heat} H_{re-heat}}{a_{eq} H_{eq}} \frac{a_{eq} H_{eq}}{a_{re-entry} H_{re-entry}} \\ &= e^{-N_*} \frac{a_{end} H_{end}}{a_{re-heat} H_{re-heat}} \frac{a_{re-heat} H_{re-heat}}{a_{eq} H_{eq}} \frac{a_{eq} H_{eq}}{a_{re-entry} H_{re-entry}} \end{aligned} \tag{3.1.28}$$

[77, Eq. (18.7)]. Where *end* represent the end of inflation, *re-heat* represents the epoch of re-heating as if it occurred instantaneously, *eq* represents the epoch of radiation matter equality, and *re-entry* is the epoch of Hubble re-entry. The corresponding co-moving Hubble horizons can be approximated using energy density. The results give e-foldings from Hubble exit to the end of inflation, i.e., *observable inflation*, in the range of $N_* \sim 50 - 60$ e-foldings [63]. This is the range of observable inflation used by the Planck Collaboration [87] to evaluate, constrain, and rule out various models of inflation. The application of observable inflation is described in Chapter 4 and can be seen in Figure 4.5 and several other figures presented in Chapter 4.

3.1.6 Scalar Mode Perturbation Parameters

In §3.2 *Quantum Origin of Perturbations*, the following expression for the power spectrum of the curvature perturbation is derived in Eq. (3.2.20):

$$P_\zeta(k) = \frac{1}{4\pi^2} \left(\frac{H^2(\phi)}{\dot{\phi}} \right)^2 \Bigg|_{k \sim aH}.$$

Applying the SRA equations $\dot{\phi} = -\frac{V_\phi}{3H}$ from the attractor equation in Eq. (3.1.14) and $H^2 = \frac{V}{3\,M_{Pl}^2}$ from Eq. (3.1.9), yields the following expression for the power spectrum evaluated at

wavenumber k at causal exit:

$$\mathcal{P}_\zeta(k) = \frac{1}{12\pi^2\, M_{Pl}^6} \frac{V^3(\phi)}{V_\phi^2}\bigg|_{k\sim aH}. \tag{3.1.29}$$

By reference to Eq. (3.1.18), this result may also be expressed in terms of the SRA parameter ϵ as

$$\mathcal{P}_\zeta(k) = \frac{1}{24\pi^2\, M_{Pl}^4} \frac{V}{\epsilon}\bigg|_{k\sim aH}. \tag{3.1.30}$$

From §2.7, the power law expression used by the Planck Collaboration to approximate the power spectrum of the curvature perturbation in the CAMB Codes is[2]

$$\mathcal{P}_\zeta(k) = A_s \left(\frac{k}{k_*}\right)^{n_s-1}$$

where A_s is the scalar mode amplitude parameter and n_s is the scalar mode spectral index. The amplitude parameter is the power spectrum evaluated at the pivot scale, k_*:

$$A_s = \mathcal{P}_\zeta(k_*) = \frac{1}{24\pi^2\, M_{Pl}^4} \frac{V}{\epsilon}\bigg|_{k_*\sim aH}. \tag{3.1.31}$$

Equating the power law expression for $\mathcal{P}_\zeta(k)$ used by the Planck Collaboration to the SRA derived expression in Eq.(3.1.30) yields

$$A_s \left(\frac{k}{k_*}\right)^{n_s-1} = \frac{1}{24\pi^2\, M_{Pl}^4} \frac{V}{\epsilon}$$

where the right side is evaluated at $k \sim aH$. Applying the natural log to both sides of this equation yields

$$\ln A_s + (n_s - 1)\ln k - (n_s - 1)\ln k_* = \ln\left(\frac{1}{24\pi^2\, M_{Pl}^4}\right) + \ln V - \ln\epsilon\,.$$

Taking the derivative with respect to $\ln k$ yields

$$n_s - 1 = \frac{d \ln V}{d \ln k} - \frac{d \ln \epsilon}{d \ln k}$$

where the constant terms have vanished. By applying the SRA equation $H^2 \simeq \frac{V}{3 M_{Pl}^2}$, the previous expression may be written

$$\begin{aligned} n_s - 1 &= \frac{d \ln \left(3 M_{Pl}^2 H^2\right)}{d \ln k} - \frac{d \ln \epsilon}{d \ln k} \\ &= 2 \frac{d \ln H}{d \ln k} - \frac{d \ln \epsilon}{d \ln k} . \end{aligned} \tag{3.1.32}$$

Using $k = aH$ leads to the following approximation:

$$\begin{aligned} \frac{d \ln k}{dt} &= \frac{d \ln aH}{dt} \\ &= \frac{1}{aH} \left(\dot{a} H + a \dot{H}\right) \\ &= H + \frac{\dot{H}}{H} \\ &\simeq H \end{aligned}$$

since $\dot{H}$ is small compared to H. Therefore, $d \ln k = d \ln aH \simeq H dt$ and the first expression on the right side of Eq. (3.1.32) may be written

$$\begin{aligned} 2 \frac{d \ln H}{d \ln k} &= \frac{2}{H} \frac{d \ln H}{dt} \\ &= 2 \frac{\dot{H}}{H^2} . \end{aligned}$$

In Eq. (3.1.18), ϵ was defined as $\epsilon = \frac{M_{Pl}^2}{2} \left(\frac{V_\phi}{V}\right)^2 \simeq \frac{|\dot{H}|}{H^2}$. $\dot{H}$ is negative during accelerated expansion in the SRA (Eq. (3.1.7)). Therefore,

$$\frac{d \ln H}{d \ln k} \simeq -\epsilon . \tag{3.1.33}$$

The second expression on the right side of Eq. (3.1.32) may be simplified to

$$\begin{aligned}\frac{d\ln\epsilon}{d\ln k} &= \frac{d\left[\ln\left(\frac{M_{Pl}^2}{2}\left(\frac{V_\phi}{V}\right)^2\right)\right]}{d\ln k} \\ &= 4\epsilon - 2\eta\,. \end{aligned} \tag{3.1.34}$$

By substituting Equations (3.1.33) and (3.1.34) into Eq. (3.1.32), the scalar spectral index n_s may be written

$$n_s = 1 - 6\epsilon + 2\eta \tag{3.1.35}$$

[77, Eq. (25.5)] where the parameters are evaluated at causal exit $k \sim aH$. However, under the assumption that the spectral index is not a function of the scale, n_s may be evaluated at the causal exit of the pivot scale, $k_* \sim aH$.

3.1.7 Tensor Mode Perturbation Parameters

The theoretical power spectrum of tensor perturbations is given by

$$\mathcal{P}_\mathcal{T}(k) = \frac{8}{M_{Pl}^2}\left(\frac{H(\phi)}{2\pi}\right)^2\Bigg|_{k\sim aH} \tag{3.1.36}$$

[77, Eq. (24.60)] or, expressed in terms of the potential energy used in the SRA, $V \simeq 3M_{Pl}^2H^2$,

$$\mathcal{P}_\mathcal{T}(k) \simeq \frac{2V}{3\pi^2\,M_{Pl}^4}\Bigg|_{k\sim aH}. \tag{3.1.37}$$

From §2.7, the power spectrum of the tensor perturbation is approximated by the power law expression:

$$\mathcal{P}_t(k) = A_t \left(\frac{k}{k_*}\right)^{n_t} . \tag{3.1.38}$$

Since these expressions represent the same quantity, the tensor mode amplitude parameter used in the power law expression at pivot scale k_* is

$$A_t = \mathcal{P}_{\mathcal{T}}(k_*) = \left.\frac{2V}{3\pi^2 M_{Pl}^4}\right|_{k_* \sim aH} \tag{3.1.39}$$

and

$$A_t \left(\frac{k}{k_*}\right)^{n_t} = \frac{8}{M_{Pl}^2} \left(\frac{H(\phi)}{2\pi}\right)^2 \Bigg|_k . \tag{3.1.40}$$

Taking the natural log of both sides, applying the $d/d\ln k$ to the result (as above), and using Eq. (3.1.33) yields a simple expression for the tensor mode spectral index:

$$n_t = -2\epsilon . \tag{3.1.41}$$

Using the definition of the cosmological parameter, r, as the ratio of the tensor mode amplitude parameter to the scalar mode amplitude parameter yields

$$\begin{aligned} r &= \left.\frac{\frac{2V}{3\pi^2 M_{Pl}^4}}{\frac{1}{24\pi^2 M_{Pl}^4}\frac{V}{\epsilon}}\right|_{k_*} \\ &= 16\epsilon \end{aligned} \tag{3.1.42}$$

[77, Eq. (25.10)] at pivot scale k_*. The value of n_t is frequently reported in the literature as

$$n_t = -\frac{r}{8} \tag{3.1.43}$$

which is known as the consistency relation for single scalar field models of inflation [65].

Section Summary. The slow-roll regime described in this section provides the conditions for inflation to commence and continue in Eqs. (3.1.19) and (3.1.25). The "attractor equation" in Eq. (3.1.14) provides a "graceful exit" from inflation. Eq. (3.1.27) provides a mechanism to determine a range of values of the *inflaton field* at Hubble exit corresponding to a pivot scale, k_*. Eq. (3.1.28) provides a method to test models of inflation by tuning free parameters in the model to produce the expected amount of observable inflation.

The observed values for three cosmological parameters, A_s, n_s, and r, and the estimated range of observable inflation, N_*, can be used to estimate free parameters in potential energy functions for most models of inflation using the expressions for ϵ and η, $\mathcal{P}_\zeta$, $\mathcal{P}_\mathcal{T}$, and N_* that are all functions of potential energy and its derivatives. The parameterized models can be tested in computer simulations of the development of CMB anisotropies and the growth of cosmological structures. Many models are ruled out while others are ranked according to the model's ability to reproduce observations. The search for new models and modifications to existing models continues.

3.2 Quantum Origin of Perturbations

Section Overview. The quantum origin of perturbations is described in the context of slow-roll single field models of inflation. Perturbations of the *inflaton field* are shown to give rise to the curvature perturbation. Energy density contrasts that are responsible for the formation of structure in the universe are given as functions of the curvature perturbation. It is shown that slow-roll single field inflation models produce the adiabatic Gaussian density perturbations and the curvature perturbation with a nearly flat power spectrum, $\mathcal{P}_\zeta(k)$.

Observations by the Planck Collaboration and the BICEP2/Keck array team [87, 89, 3, 1] are consistent with these conditions and, therefore, favor slow-roll single field models of inflation. Additionally, observations favor theories in which tensor mode perturbations are

near zero (parameterized by r), and a scale independent scalar spectral index is parameterized by n_s (so-called "non running tilt").

3.2.1 *Inflaton Field* Perturbations

The models of inflation described in this section are defined by the potential energy of the real valued scalar *inflaton field*. In single field slow-roll models of inflation, the energy of the universe at the beginning of the inflation epoch is dominated by the energy of the *inflaton field*. Ordinary fields, such as matter and radiation fields, also exist during this period but have small energy compared to the *inflaton field*.

As described in Chapter 2, the quantity of expansion required during the inflationary epoch must have been great enough so that the region that became the observable universe must have been well inside the Hubble horizon at the beginning of inflation and well outside the Hubble horizon at the end of inflation. Particles in the field within the region that became the observable universe could interact, moving the field toward homogeneity. In §2.5.2 describing the horizon problem, it was found that about 60-70 e-foldings were needed during inflation to smooth any initial roughness to the slight anisotropies observed in the CMB and that this amount of expansion would have produced a much greater degree of homogeneity during the very early universe (Eq. (2.5.9)). As shown in §3.1.5, the amount of observable inflation (post Hubble exit inflation) is expected to be 50-60 e-foldings.

Most models of inflation produce significantly more than 70 e-foldings of expansion before Hubble exit, making the *inflaton field* spatially homogeneous and isotropic well before Hubble exit, *excluding the effects of quantum uncertainties*.

Values that exclude quantum uncertainties may be described as the unperturbed background on which quantum uncertainties are expressed. Let $\phi(t)$ represent the unperturbed background value of the *inflaton field* at some time after spatial homogeneity and isotropy have been achieved. The background field is, therefore, a perfect fluid with no gradient

energy. Spatial homogeneity means that the scalar value, $\phi(t)$, is position independent. Similarly, functions of $\phi(t)$ are position independent, including $V(\phi)$, $H(\phi) \simeq \frac{1}{M_{Pl}}\sqrt{V(\phi)}$, and $V_\phi(\phi) \equiv \partial V(\phi)/\partial\phi$ (suppressing the time parameter for clarity). By the attractor equation (Eq. (3.1.14)), the potential energy term produces a slow change in the background value of $\phi(t)$ as a function of time. Since this process occurs throughout the field equally in all directions, isotropy of the background is preserved. Therefore, the background value of $\dot{\phi}(t)$ is small but non zero and position independent, as is $\dot{V}(\phi)$ since $\dot{V}(\phi) = V_\phi\,\dot{\phi}(t)$.

The value of the field at position $\mathbf{x}$ that includes quantum uncertainties is defined as $\phi(\mathbf{x}, t) = \phi(t) + \delta\phi(\mathbf{x}, t)$ where $\delta\phi(\mathbf{x}, t)$ represents the realized quantum uncertainties at position $\mathbf{x}$ at time t. Since this discussion is understood to take place at a specific time, the time parameter will be suppressed except as needed. Rearranging the expression yields

$$\delta\phi(\mathbf{x}) = \phi(\mathbf{x}) - \phi\,. \tag{3.2.1}$$

The term $\delta\phi(\mathbf{x})$ represents a realization of the quantum uncertainties of a random process.

In §2.7.3 *Statistical Analysis of Primordial Scalar Perturbations*, Fourier methods were used to expand the Gaussian random field of a generic perturbation as an integral sum of the products of perturbation amplitude coefficients and classical wave equations, $e^{i\,\mathbf{k}\cdot\mathbf{x}}$, over a continuum of co-moving wavevectors. In the case of the perturbation of the *inflaton field*, the random field of the perturbation is $\delta\phi(\mathbf{x})$ and the random field of the amplitude coefficients is $\delta\phi(\mathbf{k})$. Additionally, in this section it was found that the power spectrum of $\delta\phi(\mathbf{x})$ is $P_{\delta\phi}(k) = \left\langle \left|\delta\phi(\mathbf{k})\right|^2 \right\rangle$, the two-point auto correlator of the amplitude coefficients (Eq. (2.7.24)). The Gaussian distribution of $\delta\phi(\mathbf{x})$ was determined by computing the position independent variance as a function of the power spectrum. The variance on a logarithmic scale of k in Eq. (2.7.35) is an important example:

$$\sigma^2_\delta(\mathbf{x}) \simeq \mathcal{P}_\delta(\ln k)\,.$$

The term $\delta\phi(\mathbf{k})\, e^{i\,\mathbf{k}\cdot\mathbf{x}}$ in the inverse Fourier transformation (Eq. (2.7.18)) is the contribution of the interval $d\mathbf{k}$ in the continuum of wavevectors to the perturbation of the field at position $\mathbf{x}$, namely $\delta\phi(\mathbf{x})$. The perturbation amplitude coefficient, $\delta\phi(\mathbf{k})$, is the fixed realization of the quantum fluctuations with random amplitudes in the *inflaton field* with co-moving wavelength $1/|\mathbf{k}| = 1/k$. According to scalar field theory [74, 66], the amplitude of the quantum fluctuations of co-moving wavelength, k^{-1}, converges to a fixed value at or shortly after causal exit of the co-moving wavelength. Additional quantum fluctuations occur but their effects drop to negligible levels within a few units of Hubble time and the amplitude of fluctuations is essentially fixed.

At the epoch, $k \sim aH$, the fluctuations of the field, $\delta\phi(\mathbf{x})\Big|_k$, are determined by inverse Fourier transformation based on the modes, $\delta\phi(\mathbf{k}')\, e^{i\,\mathbf{k}'\cdot\mathbf{x}}$ for all co-moving wavelengths $k'^{-1} \geq k^{-1}$. Since each of these modes is fixed at or before $k \sim aH$, $\delta\phi(\mathbf{x})\Big|_k$ may be regarded as a classical perturbation.

Physical wavelength $a\,k^{-1}$ corresponds to co-moving wavelength k^{-1}. Physical wavelengths expand enormously during inflation. The epoch when physical wavelength reaches physical Hubble exit may be written

$$a\,k^{-1} = H^{-1}$$

where $a\,k^{-1}$ is the physical wavelength and k^{-1} is the co-moving wavelength. Equivalently, the moment that the co-moving Hubble length shrinks to the size of a co-moving wavelength may be written

$$k^{-1} = (a\,H)^{-1}\,. \tag{3.2.2}$$

Co-moving Hubble exit is usually written

$$k = a\,H\,.$$

The co-moving Hubble length is on the order of causal connection. Therefore, the epochs of "causal exit" and "causal re-entry" may both be written

$$k \sim a\,H\,.$$

Therefore, the epoch that the amplitude $\delta\phi(\mathbf{k})$ is fixed may be written

$$\delta\phi(\mathbf{k})\Big|_{k\sim aH}\,. \tag{3.2.3}$$

From Eq. (3.2.2), it is evident that an amplitude of fluctuations associated with a longer wavelength (smaller wavenumber), $\delta\phi(\mathbf{k})$, is fixed earlier than an amplitude of fluctuations associated with a shorter wavelength (larger wavenumber).

Consequently, the fixed value of the two-point auto correlator $\left\langle \left|\delta\phi(\mathbf{k})\right|^2 \right\rangle$ is also set at or shortly after causal exit. The auto correlator may be determined using the mean square value of the perturbation during a time interval beginning when the co-moving Hubble length was one e-folding greater than the co-moving wavelength of $\phi(\mathbf{k})$ and ending *later*, i.e., when co-moving Hubble length equals the co-moving wavelength of $\phi(\mathbf{k})$ (recalling that the co-moving Hubble horizon is shrinking throughout inflation) [74]. The power spectrum of perturbations of the *inflaton field* may be written in logarithmic units of the wavenumber as

$$\begin{aligned} P_{\delta\phi}(k) &= \left\langle \left|\delta\phi(\mathbf{k})\right|^2 \right\rangle \Big|_{k\sim aH} \\ &= \frac{H^2(\phi)}{4\pi^2} \int_{\ln\frac{k}{aH}}^{\ln\frac{k}{aH}e} d\ln\frac{k}{aH}\Big|_{k\sim aH} \\ &= \frac{H^2(\phi)}{4\pi^2} \int_{\ln\frac{k}{aH}}^{\ln\frac{k}{aH}+1} d\ln\frac{k}{aH}\Big|_{k\sim aH} \\ &= \frac{H^2(\phi)}{4\pi^2}\Big|_{k\sim aH}\,. \end{aligned} \tag{3.2.4}$$

The square root yields the root mean square (the "rms"):

$$P_{\delta\phi}^{\frac{1}{2}}(k) = \left\langle \left|\delta\phi(\mathbf{k})\right| \right\rangle \Big|_{k\sim aH} = \frac{H(\phi)}{2\pi}\Big|_{k\sim aH} \tag{3.2.5}$$

where $\frac{H(\phi)}{2\pi}$ is the Gibbons-Hawking "temperature" that is associated with the event horizon in a de Sitter space, T_H [30, 12]. Since the value of the scalar, ϕ, changes slowly during inflation, $H(\phi)$ also changes slowly. Therefore, the power spectrum of perturbations of the *inflaton field* is nearly scale invariant.

See Chapters 2 and 7 and Eq. (7.5.5) in Ref. [74] and Chapter 24 in Ref. [77] for comprehensive treatments of this result from quantum field theory.

3.2.2 Primordial Perturbations From *Inflaton Field* Perturbations

By Eq. (3.2.3), the amplitude, $\delta\phi(\mathbf{k})$, was fixed at $k \sim a\,H$, for every k that reached causal exit before the end of inflation. The initial perturbation amplitude, $\delta\phi(\mathbf{k})$, that affects the observable universe was fixed at $k_0 \sim a_0\,H_0$, where k_0 represents the largest wavelength and smallest wavenumber. Define k_ω as the co-moving wavenumber of the final $\delta\phi(\mathbf{k})$ fixed before the end of inflation. Thus, k_ω represents the smallest wavelength and the largest wavenumber.

The co-moving region $R_{10^4\odot}$ defined in Eq. (2.7.5) is an estimate of $1/k_\omega$ based on observations. Therefore, it is expected that $k_\omega \sim 1/R_{10^4\odot}$.

If $\delta\phi(\mathbf{x})$ is determined by applying the inverse Fourier transformation at the epoch at $k_\omega \sim aH$, then it will include the contributions from quantum fluctuations of the entire spectrum of fluctuation wavelengths fixed during inflation, $k_\omega^{-1} \rightarrow k_0^{-1}$. Since the terms of $\delta\phi(\mathbf{k})\,e^{i\,\mathbf{k}\cdot\mathbf{x}}$ (sometimes called Fourier modes) are uncorrelated except as required by the reality condition, $\delta\phi(\mathbf{x})$ is a Gaussian random field as described in §2.7.3.

A region of scale $1/k$ reaches causal exit at $k \sim aH$ and is causally disconnected from surrounding regions until causal re-entry at $k \sim aH$. During this period, between first and

second intersections in Figure 2.8, the region is an isolated system that evolves as a separate universe. It is evident that regions of small scale reach causal exit later than regions of larger scale and causal re-entry earlier than regions of larger scale.

At the epoch $k \sim aH$, the amplitudes $\delta\phi(\mathbf{k})$ with $k_0 \leq |\mathbf{k}| \leq k$ have become fixed and

$$\begin{aligned}\delta\phi(\mathbf{x}) &= \frac{1}{(2\pi)^3}\int_{k_0\leq|\mathbf{k}|\leq k}\delta\phi(\mathbf{k})\,e^{i\,\mathbf{k}\cdot\mathbf{x}}\,d\mathbf{k} \\ &\equiv \frac{1}{(2\pi)^3}\int_0^{2\pi}\int_0^{\pi}\int_{k_0}^{k}\delta\phi(k',\theta',\varphi')\,e^{ik'x\left(\sin\theta'\sin\theta\cos(\varphi'-\varphi)-\cos\theta'\cos\theta\right)}\,k^2\sin\theta'\,dk'\,d\theta'\,d\varphi'\end{aligned}$$

is uniform within the region of scale k^{-1}. This represents the smoothing of the region to wavenumber k. There are many such regions within the field. The value of $\delta\phi(\mathbf{x})$ is constant at all positions within each region, but these constant values are different for different regions of the same scale.

Within each isolated region, $\phi(\mathbf{x}) = \phi + \delta\phi(\mathbf{x})$ where ϕ was defined earlier as the constant background value of the unperturbed field. Since $\delta\phi(\mathbf{x})$ is constant within the region, $\phi(\mathbf{x})$ is constant within the region. The region may be regarded as a separate homogeneous universe smoothed to scale k^{-1}.

In section §2.7.2 density perturbations at time t were written as $\delta\rho(k, \mathbf{x}, t)$. To avoid any possible confusion between position space and k space in Fourier expansion, perturbations of the *inflaton field* for regions centered on positions $\mathbf{x}$, each with co-moving scale k^{-1}, are written here as

$$\delta\phi(\mathbf{x})\Big|_k$$

where the relevant scales represented by their wavenumbers are $k_0 \leq k \leq k_\omega$. If $\delta\phi(\mathbf{x})$ is written without a wavenumber, then the smallest scale, k_ω^{-1}, is assumed. Therefore,

$$\begin{aligned}\delta\phi(\mathbf{x}) &= \frac{1}{(2\pi)^3}\int_{k_0\leq|\mathbf{k}|\leq k_\omega}\delta\phi(\mathbf{k})\,e^{i\,\mathbf{k}\cdot\mathbf{x}}\,d\mathbf{k} \\ &\equiv \frac{1}{(2\pi)^3}\int_0^{2\pi}\int_0^{\pi}\int_{k_0}^{k_\omega}\delta\phi(k',\theta',\varphi')\,e^{ik'x\left(\sin\theta'\sin\theta\cos(\varphi'-\varphi)-\cos\theta'\cos\theta\right)}\,k^2\sin\theta'\,dk'\,d\theta'\,d\varphi' \\ &= \delta\phi(\mathbf{x})\Big|_{k_\omega}\,. \end{aligned} \tag{3.2.6}$$

Except at times close to inflation exit, the energy density of the *inflaton field* is dominated by the potential energy; therefore, $\rho(\mathbf{x}) \simeq V(\phi(\mathbf{x}))$. Since $\phi(\mathbf{x})$ is constant within each region of scale k^{-1}, so is $\rho(\mathbf{x}) \simeq V(\phi(\mathbf{x}))$, and the energy density perturbations of smoothed regions may be written

$$\delta\rho(\mathbf{x})\Big|_k\,.$$

Scales that did not reach re-entry before the accelerated rate of expansion restarted about six billion years ago (see p. 33) will not re-enter since the co-moving Hubble horizon shrinks during accelerated expansion. Unless something unexpected happens to dark energy, the co-moving Hubble horizon will shrink indefinitely as the the universe becomes asymptotically de Sitter. This is evident since $\ddot{a} > 0$ means that $\dot{a}$ is increasing. The co-moving Hubble horizon, $1/aH = 1/\dot{a}$, is shrinking when $\dot{a}$ is increasing. This also means that co-moving regions are currently in the process of reaching causal disconnection.

The Curvature Perturbation. In §2.7.2 it was stated that energy density contrasts are directly linked to the "curvature perturbation" and that the curvature perturbation is conserved from Horizon exit to Horizon re-entry of the scale (i.e., the co-moving size of the region) used to determine the perturbation. An expression for the curvature perturbation begins with the following string of equalities that are based on the position independence

of the background values $\dot{\phi}$ and $\dot{V}(\phi)$ and the dominance of potential energy at causal exit during inflation:

$$\frac{\delta\phi(\mathbf{x})}{\dot{\phi}} = \frac{\delta\phi(\mathbf{x})}{\dot{\phi}(\mathbf{x})} = \frac{V_\phi(\phi(\mathbf{x}))}{V_\phi(\phi(\mathbf{x}))}\frac{\delta\phi(\mathbf{x})}{\dot{\phi}(\mathbf{x})} = \frac{\delta V(\phi(\mathbf{x}))}{\dot{V}(\phi(\mathbf{x}))} = \frac{\delta V(\phi(\mathbf{x}))}{\dot{V}(\phi)} \simeq \frac{\delta\rho(\mathbf{x})}{\dot{\rho}}\,. \tag{3.2.7}$$

Applying the continuity equation in Eq. (B.3.22) to Eq. (3.2.7) yields

$$\left.\frac{\delta\phi(\mathbf{x})}{\dot{\phi}}\right|_k \simeq \left.\frac{\delta\rho(\mathbf{x})}{-3H(\rho+p)}\right|_k.$$

The expression $\rho + p = A\,\rho$ for some A determined based on the unknown equation of state of the *inflaton field* that is constant throughout the field (a consequence of the homogeneity and isotropy of the background field). Substituting $B = 3A$ into the previous expression yields

$$\left.\frac{\delta\phi(\mathbf{x})}{\dot{\phi}}\right|_k \simeq \left.\frac{\delta\rho(\mathbf{x})}{-BH\,\rho}\right|_k.$$

Multiplying both side of the equation by $-BH$ yields

$$\left.\frac{\delta\rho(\mathbf{x})}{\rho}\right|_k = B\left[\frac{-H(\phi)\delta\phi(\mathbf{x})}{\dot{\phi}}\right]_k. \tag{3.2.8}$$

The curvature perturbation is defined as the bracketed term in Eq. (3.2.8) where the perturbation is based on regions of co-moving scale k^{-1} and expressed in terms of the *inflaton field* as

$$\left.\zeta(\mathbf{x})\right|_k = \left.\frac{-H(\phi)\delta\phi(\mathbf{x})}{\dot{\phi}}\right|_{k\sim aH} \tag{3.2.9}$$

for every $k \in [k_0, k_\omega]$. On the left side, $\Big|_k$ indicates the co-moving scale for which the curvature perturbation is determined. On the right side, $\Big|_{k\sim aH}$ indicates the epoch of causal exit at which the value of the curvature perturbation is determined. It also indicates the

epoch of re-entry. Since the perturbation of each region includes a spectrum of fluctuation wavelengths fixed during inflation through $\delta\phi(\mathbf{x})$, and the regions are causally disconnected from surrounding regions until re-entry, the perturbation of each region of scale k^{-1} is conserved through re-entry. The perturbation is adiabatic since regions exchange no particles and hence no heat with surrounding regions. See §B.10 *Conservation of the Curvature Pertrubation* for additional information on the conservation of the curvature perturbation and the adiabatic condition. Without an explicit scale indicator, $\Big|_k$, $\zeta(\mathbf{x})$ represents the total curvature perturbation:

$$\zeta(\mathbf{x}) \equiv \zeta(\mathbf{x})\Big|_{k_\omega} = \frac{-H(\phi)\delta\phi(\mathbf{x})}{\dot{\phi}}\Big|_{k_\omega \sim aH}.$$

By applying Eq. (3.2.7), the curvature perturbations may be written in terms of energy density as

$$\zeta(\mathbf{x})\Big|_k = -\frac{H\delta\rho(\mathbf{x})}{\dot{\rho}}\Big|_{k\sim aH}. \tag{3.2.10}$$

This expression is valid for regions of all scales, k^{-1} $k_\omega^{-1} \rightarrow k_0^{-1}$. It is evident that $H(\rho) = H$. If $k \sim aH$ is based as the epoch of re-entry, the curvature perturbation may be determined during the post-inflation era at $t_\dagger$, the time the primordial density perturbations were established in Eq. (2.7.11). Equivalently, this epoch may be referenced by temperature or the thermal energy of the plasma as

$$t_\dagger \sim 15\, seconds \equiv T_\dagger =\sim 3\times 10^8 K \equiv E_\dagger \sim 0.10 MeV\,. \tag{3.2.11}$$

The thermal energy history leading up to nucleosynthesis is well understood since it is highly constrained to produce the relative quantities of the light elements measured in the universe. This period was heavily dominated by radiation (the "radiation era"); therefore,

by the continuity equation:

$$\begin{aligned}\zeta(\mathbf{x})\Big|_{k_\dagger} &= -\frac{H\delta\rho(\mathbf{x})}{\dot{\rho}} \\ &= -\frac{H\delta\rho(\mathbf{x})}{-3H(\rho+p)} \\ &= \frac{\delta\rho(\mathbf{x})}{3(\rho+\frac{1}{3}\rho)} \\ &= \frac{\delta\rho(\mathbf{x})}{4\,\rho}\,. \end{aligned} \tag{3.2.12}$$

Equivalently, primordial density contrast smoothed to scale $k_\dagger$ from Eq. (2.7.12) is four times the curvature perturbation:

$$\frac{\delta\rho(\mathbf{x},t)}{\rho}\Big|_{k_\dagger} = 4\,\zeta(\mathbf{x})\Big|_{k_\dagger}\,. \tag{3.2.13}$$

The expression is valid at all times from causal exit of the region of scale $k_\dagger^{-1}$ during inflation through causal re-entry sometime around the end of nucleosynthesis during the radiation era.

Applying the Fourier transformation to both sides of Eq. (3.2.9) yields an expression for the fixed amplitude coefficients of the curvature perturbation:

$$\zeta(\mathbf{k}) = \frac{-H(\phi)\delta\phi(\mathbf{k})}{\dot{\phi}}\Big|_{k\sim aH}\,. \tag{3.2.14}$$

By Fourier transformation of Eq. (3.2.7), the curvature perturbation amplitudes may be written

$$\zeta(\mathbf{k}) = -\frac{H\delta\rho(\mathbf{k})}{\dot{\rho}}\,. \tag{3.2.15}$$

Substituting Eq. (3.2.9) into Eq. (3.2.8) yields

$$\left.\frac{\delta\rho(\mathbf{x},t)}{\rho}\right|_k = \left. B\,\zeta(\mathbf{x})\right|_k \tag{3.2.16}$$

where the curvature perturbation is determined at $k \sim aH$ (exit or entry). The left side is valid at times from causal exit at $k \sim aH$ to causal re-entry of the scale, and B is calculated based on the constituents of the energy density at time t and the corresponding fixed equations of state. Applying the Fourier transformation:

$$\frac{\delta\rho(\mathbf{k})}{\rho} = B\,\zeta(\mathbf{k})\,. \tag{3.2.17}$$

To develop statistical models of the random process requires the power spectrum of the curvature perturbation to determine the variance of the random field. Applying the two-point auto correlator to both sides of Eq. (3.2.14) yields

$$\langle\zeta(\mathbf{k})\zeta(\mathbf{k}')\rangle = \left(\frac{H}{\dot\phi}\right)^2 \langle\delta\phi(\mathbf{k})\,\delta\phi(\mathbf{k})\rangle\Big|_{k\sim aH}\,.$$

Applying Eq. (2.7.23) and Eq. (2.7.24) to the previous equation yields

$$P_\zeta(k) = \left\langle\left|\zeta(\mathbf{k})\right|^2\right\rangle = \left(\frac{H}{\dot\phi}\right)^2 \left\langle\left|\delta\phi(\mathbf{k})\right|^2\right\rangle\Big|_{k\sim aH}\,. \tag{3.2.18}$$

Similarly, applying the two-point auto correlator to Eq. (3.2.17) followed by application of Eq. (2.7.23) and Eq. (2.7.24) yields

$$P_{\delta\rho/\rho}(k) = \frac{\left\langle\left|\delta\rho(\mathbf{k})\right|^2\right\rangle}{\rho^2} = B^2\left(\frac{H}{\dot\phi}\right)^2 \left\langle\left|\delta\phi(\mathbf{k})\right|^2\right\rangle\Big|_{k\sim aH}\,. \tag{3.2.19}$$

Applying Eq. (3.2.4) to Eq. (3.2.18) and Eq. (3.2.19), the power spectra of the curvature perturbation and the density contrast per unit interval in $\ln \frac{k}{aH}$ are

$$P_\zeta(k) = \frac{1}{4\pi^2}\left(\frac{H^2(\phi)}{\dot{\phi}}\right)^2\Bigg|_{k\sim aH} \tag{3.2.20}$$

and

$$P_{\delta\rho/\rho}(k) = B^2\frac{1}{4\pi^2}\left(\frac{H^2(\phi)}{\dot{\phi}}\right)^2\Bigg|_{k\sim aH}. \tag{3.2.21}$$

Therefore, the rms value of the amplitude coefficients of curvature perturbation $\zeta(\mathbf{k})$ per unit interval in $\ln \frac{k}{aH}$ is

$$P_\zeta^{\frac{1}{2}}(k) = \left\langle \left|\zeta(\mathbf{k})\right| \right\rangle = \frac{\left[H(\phi)\right]^2}{2\pi\dot{\phi}}\Bigg|_{k\sim aH} \tag{3.2.22}$$

and the rms value of each $\rho(\mathbf{k})/\rho$ per unit interval in $\ln \frac{k}{aH}$ is

$$P_{\delta\rho/\rho}^{\frac{1}{2}}(k) = \frac{\left\langle \left|\delta\rho(\mathbf{k})\right| \right\rangle}{\rho} = |B|\,\frac{\left[H(\phi)\right]^2}{2\pi\dot{\phi}}\Bigg|_{k\sim aH}. \tag{3.2.23}$$

Prediction: Near Scale Invariance of the Curvature Perturbation Power Spectrum. According to single field slow-roll theories of inflation and scalar field theory, the power spectrum of the curvature perturbation is

$$P_\zeta(k) = \frac{1}{4\pi^2}\left(\frac{H^2(\phi)}{\dot{\phi}}\right)^2\Bigg|_{k\sim aH}$$

(Eq. (3.2.20)) which is modeled in power-law form as

$$\mathcal{P}_\zeta(k) = A_s\left(\frac{k}{k_*}\right)^{n_s-1}$$

(Eq. (2.7.41)) where k_* is the pivot scale value and $A_s = P_\zeta(k_*)$ is determined using observations of the CMB.

Since the value of the potential energy, $V(\phi)$, must remain sufficiently constant for inflation to commence and continue long enough to produce sufficient quantities of expansion to "solve" the horizon problem, etc., the value of ϕ must change slowly through most of the inflation epoch. Since $H^2(\phi) \propto V(\phi)$, its value must also change slowly. Similarly, $\dot{\phi}$ must be small and evolve slowly through most of the inflation epoch. Since the expression for the power spectrum for each wavenumber is composed of slowly changing terms evaluated at causal exit of k^{-1}, inflation theories predict that the power spectrum of the curvature perturbation is nearly, but not precisely, scale invariant over a broad range of wavenumbers.

Transition to the Post-Inflation Epoch. The energy density of the *inflaton field* was dominated by the potential energy of the field until the later stages of inflation:

$$\delta\rho(\mathbf{x}, t) \simeq \delta V(\mathbf{x}, t) = V_\phi \, \delta\phi(\mathbf{x}, t) \, . \tag{3.2.24}$$

Although the potential energy changes slowly, the attractor equation in the SRA shows that the value of the *inflaton field* is attracted to the value that produces the minimum potential energy. This process occurs throughout all positions in the *inflaton field*. Inflation ends locally when the potential energy of the field begins to oscillate around its effective minimum. At this stage, much of the energy of the field has been transferred from potential energy to the kinetic energy of the oscillations, producing thermal energy. During oscillations of the *inflaton field*, the thermal energy of the *inflaton field* is transferred to ordinary fields in the vacuum. The *inflaton field* loses energy and decays. The excited ordinary fields produce a very hot plasma of ordinary elementary particles and, if they exist, supersymmetric particles. These are the pre-heating and re-heating processes. One or more of these ordinary particles or supersymmetric particles may be dark matter particles. After the end of pre-heating and re-heating, the universe enters the radiation era and expansion is driven

by the pressure of thermal energy (much later to be dominated by dark energy). The thermal energy of the plasma reaches homogeneous background temperature called the re-heat temperature. The re-heat temperature varies by model but is hot enough so that all particles in the plasma are highly relativistic as the universe enters the radiation era. (See §7.9 of [74] and §6 of [65] for information on these processes.)

Prediction: All Scalar Perturbations are Caused by Adiabatic Curvature Perturbations. Energy density perturbations near the end of inflation produce local delays of the end of inflation. Local delays of the end of inflation cause local delays of the decay of the *inflaton field* into the hot-plasma of highly relativistic ordinary particles, leading to local delays in the cooling effect of post-inflation expansion on the temperature of the plasma. This produces local variations in the temperature of the plasma at the beginning of the radiation era marked by the end of pre-heating and re-heating throughout space.

The energy density of each energy component is proportional to T^4, where T is the temperature of a plasma of particles in thermal equilibrium (see Eq. (B.13.9)). Therefore, the temperature perturbations and temperature contrasts in the plasma are related to the primordial energy density perturbations and energy density contrasts at the beginning of the radiation epoch.

In Eq. (3.2.16) the evolution of local energy density contrasts among isolated regions of scale k^{-1} was found to be related to the curvature perturbation throughout the separate universe era of the regions as:

$$\left.\frac{\delta\rho(\mathbf{x},t)}{\rho}\right|_k = \left. B\,\zeta(\mathbf{x})\right|_k .$$

The curvature perturbation on the right side is determined at causal exit during inflation and is conserved through causal re-entry of the region of scale k^{-1}. As described earlier, the value of B changes over time, depending on the contribution of energy components to the total energy density and the related equations of state.

Since temperature contrasts in a plasma of particles in thermal equilibrium are related to energy density contrasts through Eq. (B.13.9), temperature contrasts are also related to the curvature perturbation.

Dark matter fell out of equilibrium with other particles, or decoupled from other particles, at some early, unknown time. It was presumably spread nearly uniformly throughout space since dark matter particles are thought to be minimally coupled to other particles including other dark matter particles. Minimally coupled means only through gravitation.

In the era leading up to photon decoupling, neutrinos had decoupled from the plasma forming two plasmas. The first plasma found throughout all space was composed of photons, electrons, and light atomic nuclei in near thermal equilibrium at some decreasing temperature, T_γ. The second plasma also found throughout all space was composed of neutrinos in near thermal equilibrium at some other decreasing temperature, T_ν (somewhat lower than T_γ). The temperature contrasts among smoothed regions, k^{-1}, in both plasmas during this epoch are related to the curvature perturbation smoothed to wavenumber, k, through distinct coefficients, B_γ and B_ν.

After photon decoupling from ordinary matter and recombination, the temperature contrasts in the CMB continued to be related to the curvature perturbation through B_γ. The energy density of matter, composed of dark matter and "baryonic" matter mostly in the form of atoms, was initially nearly uniformly distributed throughout space. Over time, matter clumped into cosmological structures such as stars, galaxies, and galaxy clusters through the effects of gravitation. Matter-energy density perturbations are related to the curvature perturbation on large enough smoothing scales.

Therefore, based on single field models of inflation, with smoothing at appropriate scales, k^{-1}, all scalar contrasts and hence scalar perturbations are based on curvature perturbations established during inflation smoothed to the wavenumber k. The curvature perturbation is shown to be adiabatic in §B.10.

Local delays of the end of inflation caused local delays of the transition to the radiation era and the beginning of post-inflation expansion. Post inflation expansion is driven by thermal energy pressure. Isolated regions with greater thermal energy have greater pressure and, therefore, more rapid expansion.

The variations in the expansion rates of isolated regions of common co-moving scale, k^{-1}, are formalized as perturbations of the metric smoothed to the same scale. The metric applies to the co-moving regions from causal exit to causal re-entry, i.e., the separate universe era of the regions. The perturbations of the metric are produced by local scale factors, $a(\mathbf{x}, t)$, each applicable to an isolated co-moving region centered on $\mathbf{x}$ during its separate universe era. The local scale factors are functions of the curvature perturbation $a(\mathbf{x}, t) = a(t)\, e^{\zeta(\mathbf{x})}$ where $a(t)$ is the usual scale factor of the homogeneous and isotropic universe and $\zeta(\mathbf{x})$ is written without a time parameter since it is conserved throughout the separate universe era of the scale.

Therefore, in single field theories of inflation, energy density perturbations of the *inflaton field* cause local variations in the expansion rate of space that induce "adiabatic perturbations of the metric" [74, p. 126]. See §7.5 of [74] for additional information.

Single Field Theories of Inflation and Current Observations. Single field theories of inflation predict that the power spectrum of the curvature perturbation is nearly, but not precisely, scale invariant over a broad range of wavenumbers. These theories also predict that temperature perturbations and energy density perturbations are caused by adiabatic curvature perturbations.

Current observations are consistent with these predictions of single field theories of inflation. From *Planck 2015 results. XX. Constraints on inflation* [89]:

- The "robust detection of the departure of the scalar spectral index from exact scale invariance, i.e., $n_s < 1$, at more than 5σ confidence, as well as the lack of the obser-

vation of any statistically significant running of the spectral index, were found to be consistent with simple slow-roll models of inflation". (p. 2)

- "The Planck data are consistent with adiabatic primordial perturbations, and the estimated values for the parameters of the base CDM model are not significantly altered when more general initial conditions are admitted." (p. 1)

- "Adiabaticity is also an important probe of the inflationary paradigm, since any significant detection of isocurvature modes would exclude the possibility that all perturbations in the Universe emerged from quantum fluctuations of a single inflaton field, which can excite only one degree of freedom, the curvature (i.e., adiabatic) perturbation". (p. 41)

Computer models such as CAMB produce maps of the CMB with statistical properties that closely match the statistical properties of the observed scalar mode temperature fluctuations in the CMB. CAMB also include codes that simulate other theoretical processes such as tensor mode perturbations, the re-ionization of matter during star formation, the matter power spectrum, and the variance of matter distribution. The variance of matter distribution describes the statistical properties of the theoretical distribution of cosmological structures. The theoretical distributions are statistically equivalent to the observed distribution of cosmological structure. See Chapter 8 of Ref. [106] for additional information.

Additional Citations. The principal findings in this section were originally presented in four articles published in 1982 and 1983: *The development of irregularities in a single bubble inflationary universe* [39] by S.W. Hawking, *Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations* [96] by A.A. Starobinsky, *Fluctuations in the New Inflationary Universe* [36] by Alan H. Guth, and So-Young Pi, and *Spontaneous creation of almost scale-free density perturbations in an inflationary universe* [12] by James M. Bardeen and Paul J. Steinhardt, and Michael S.Turner. Each of these articles was developed based on what was then called "new" inflation and what is now called

original or old "slow-roll" inflation. Each cites V. F. Mukhanov and G.C. Chibisov's earlier 1981 paper, *Quantum fluctuations and a nonsingular universe* [79], which found equivalent results based on A. A. Starobinsky's model of inflation [95]. An important purpose of all five articles was to determine whether realistic models of inflation can produce perturbations that are consistent with the development of CMB anisotropies and the growth of cosmological structures.

The original slow-roll models have been largely replaced by various versions of chaotic inflation first presented in 1983 by Andrei Linde in *Chaotic Inflation* [73]. A number of Professor Linde's articles on inflation are included in the bibliography as well as his text, *Particle Physics and Inflationary Cosmology* [74], which is cited several times above.

Equations (3.2.4), (3.2.5), (3.2.8), and (3.2.23) correspond to Equations (7.5.5), (7.5.6), (7.5.4), and (7.5.21) in *Particle Physics and Inflationary Cosmology*.

Another text published in 2009, *The Primordial Density Perturbation*, by David H. Lyth and Andrew R. Liddle is also frequently cited in this section and elsewhere in this paper.

The presentation in this section uses the separate universe assumption to establish conservation of the curvature perturbation. This is consistent with the approach used by the authors of *Spontaneous creation of almost scale-free density perturbations in an inflationary universe* and by the authors of *The Primordial Density Perturbation*.

Section Summary. The models considered in this paper are restricted to single field slow-roll models that produce adiabatic Gaussian perturbations and no running spectral index. The Planck Collaboration considered a wider range of models before concluding that observations do not favor a running spectral index, non adiabatic perturbations, or non Gaussian perturbations.

The Planck Collaboration also concluded that the best fit model of inflation is the $R+R^2$ model which is based on a modified theory of gravity. Theories of chaotic inflation, includ-

ing a class of supergravity models called α-attractors that can be constructed to reproduce the $R + R^2$ model, are the subject of Chapter 4.

Notes

[1]The limits of the integrand in Eq. (12) representing N_* are reversed in [87].

[2]The Planck Collaboration uses the notation $\mathcal{P}_{\mathcal{R}}(k)$ to represent the curvature perturbation. This notation is sometimes used to represent a broader class of scalar mode perturbations.

Chapter 4

Chaotic Inflation

4.1 Overview

A general description of chaotic inflation theories is provided in Chapter 1. Chaotic inflation was introduced by Andrei Linde in 1983 [73] to resolve problems with the original models of slow-roll inflation [71, 6]. The principal problem with the original slow-roll models was that the coupling constant of scalar fields consistent with the models was too small to thermalize matter fields sufficiently to produce elementary particles [73]. Chaotic inflation solved this problem by requiring only the conditions described in §3.1 *The Slow-Roll Approximation* for inflation to begin and continue long enough to produce the initial conditions needed as the starting point for ΛCDM cosmology. Initial thermal equilibrium is not required in chaotic inflation theories.

The potential energy terms in the first chaotic inflation models described by Linde were monomials in ϕ. The quadratic monomial, $V(\phi) = \frac{1}{2}m^2\phi^2$, called the harmonic oscillator model, is the most common model used to describe chaotic inflation. This is a large field model in which potential energy is near Planck energy density at the beginning of inflation. §4.2 describes this model and shows how to determine its free parameter, m, based on

cosmological parameters reported by the Planck Collaboration. It also shows why this model is inconsistent with the Planck Collaboration's analysis of current observations.

The chapter concludes in §4.4 with an analysis of supergravity models of chaotic inflation. These models are based on modified versions of the Einstein-Hilbert Lagrangian and are generalizations of the harmonic oscillator chaotic model described in §4.2 and higher order monomial chaotic models described in §4.3. These supergravity models are in excellent agreement with the Planck Collaboration's measurements of the principal inflation observables: A_s, n_s, and r (p. 94).

4.2 Harmonic Oscillator Theory

The application of the SRA will be demonstrated using an inflation model [65, 60] with potential energy in the form of a harmonic oscillator. The form of the potential is

$$V(\phi) = \frac{1}{2} m^2 \phi^2 \tag{4.2.1}$$

where the parameter m has dimensions of mass. This is an example of large field chaotic inflation. It is called a "large field" model because, as will be shown below, the value of the scalar field must be large during inflation.

$$V(\phi) = \frac{m^2}{2}\phi^2$$

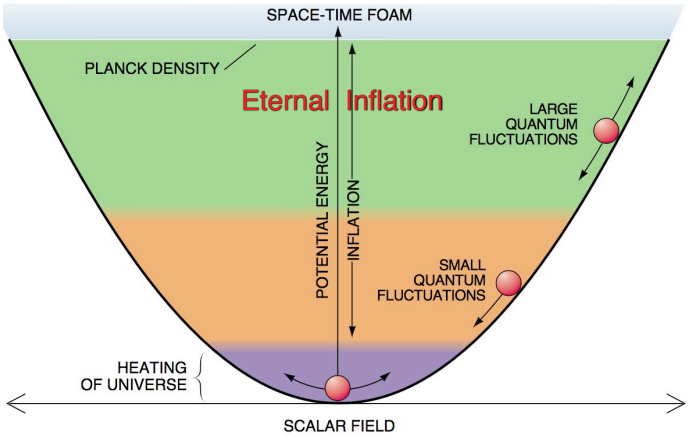


Figure 4.1: Inflation as a Harmonic Oscillator [65]

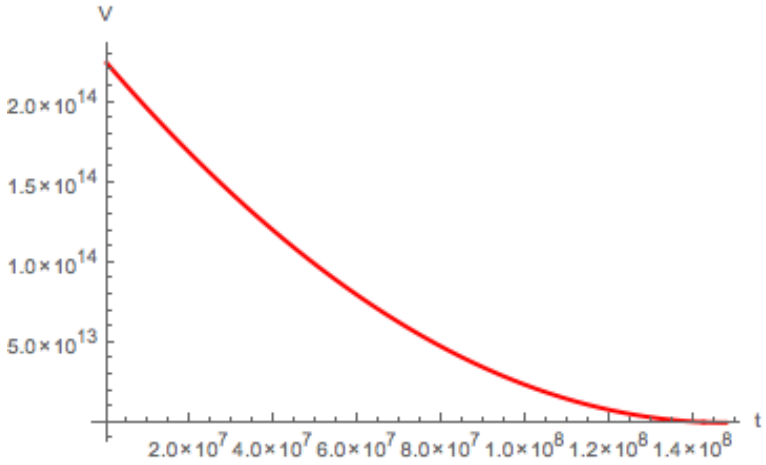


Figure 4.2: This figure shows the decline in potential energy of the harmonic oscillator model as a function of time during inflation. Time ranges from ~ 0 to $5.7 \times 10^{11}\ \sqrt{G}$ in units of time in which $c = \hbar = 1$ and G is in cgs units. Note that the decline in potential energy is less than an order of magnitude over most of the inflation epoch.

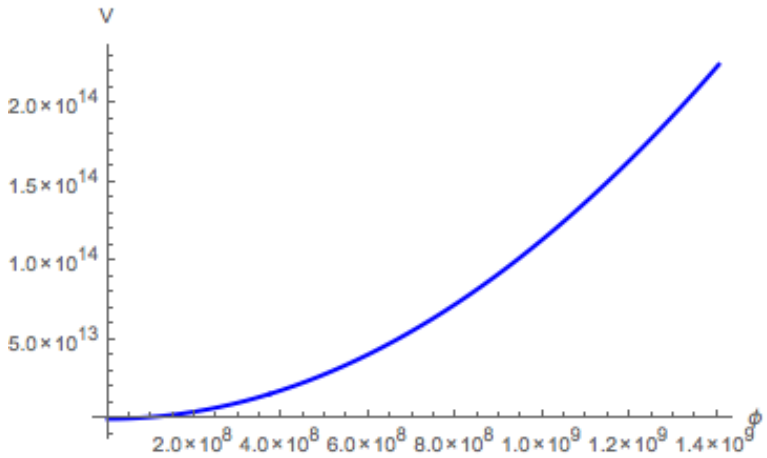


Figure 4.3: This figure shows the decline in potential energy of the harmonic oscillator model as a function of the scalar field during inflation. The initial value is Planck energy density and remains near Planck energy density during most of the inflation epoch. This is because the value of the scalar field decreases slowly during inflation. The scalar field "rolls slowly" towards its minimum value.

Although the graph of a harmonic oscillator shown in Figure 4.1 does not display a flat region, the harmonic oscillator model of inflation satisfies the SRA flatness conditions based on a wide range of the parameter m and the initial value of potential energy. Based on the assumption that the initial potential energy is approximately Planck energy density (m_{Pl}^4) and using a value for the parameter, m, consistent with Planck Collaboration measurements, the value of the potential energy stays within an order of magnitude of Planck energy density during the initial 70% of the inflation epoch. This is constant enough for a long enough period to generate enormous expansion. It will be shown that this large field model produces about a trillion e-foldings of expansion during a period lasting about 10^{-31} seconds. See Figures 4.2 and 4.3.

The SRA parameters may be written as functions of ϕ as:

$$\begin{aligned}\epsilon(\phi) &= \frac{M_{Pl}^2}{2}\left(\frac{V_\phi}{V}\right)^2 \\ &= \frac{M_{Pl}^2}{2}\left(\frac{m^2\phi}{\frac{1}{2}m^2\phi^2}\right)^2 \\ &= 2\,M_{Pl}^2\frac{1}{\phi^2}\end{aligned} \tag{4.2.2}$$

and

$$\begin{aligned}\eta(\phi) &= M_{Pl}^2 \frac{V_{\phi\phi}}{V} \\ &= M_{Pl}^2 \frac{m^2}{\frac{1}{2}m^2\phi^2} \\ &= 2\, M_{Pl}^2 \frac{1}{\phi^2}\,. \end{aligned} \tag{4.2.3}$$

Inflation ends at $\epsilon = 1$; therefore, from the Eq. (4.2.2),

$$2\, M_{Pl}^2 \frac{1}{\phi_{end}^2} = 1 \tag{4.2.4}$$

and

$$\phi_{end} = \sqrt{2}\, M_{Pl}\,. \tag{4.2.5}$$

In this model, the Hubble parameter may be written

$$\begin{aligned} H &\simeq \frac{\sqrt{\frac{1}{2}m^2\phi^2}}{\sqrt{3}M_{Pl}} \\ &= \frac{1}{\sqrt{6}} \frac{m}{M_{Pl}}\, \phi\,. \end{aligned} \tag{4.2.6}$$

After substituting this expression for H and using $\ddot{\phi} = 0$ in the SRA, the equation of motion of the *inflaton field* is

$$3\, \frac{1}{\sqrt{6}} \frac{m}{M_{Pl}}\, \phi\, \dot{\phi} + m^2\, \phi = 0$$

or, more simply,

$$\sqrt{\frac{3}{2}} \frac{1}{M_{Pl}}\, \frac{d\phi}{dt} + m = 0\,.$$

The solution is

$$\phi = \phi_i - \sqrt{\frac{2}{3}} m\, M_{Pl}\, t\,. \tag{4.2.7}$$

From Eq. (4.2.5), the time at which inflation ends is a function of the value of the *inflaton field* at the beginning of inflation, ϕ_i, and the parameter m:

$$t_{end} = \frac{\phi_i - \sqrt{2}\, M_{Pl}}{\sqrt{\frac{2}{3}} m\, M_{Pl}}\,. \tag{4.2.8}$$

By using the Hubble parameter in Eq. (4.2.6) and the solution for $\phi(t)$ in Eq. (4.2.7), the solution to the Einstein equation during inflation is determined as follows. The Hubble parameter may be written

$$H = \frac{\dot{a}}{a} = \frac{1}{a}\frac{d\,a}{d\,t} = \frac{1}{\sqrt{6}}\frac{m}{M_{Pl}}\left(\phi_i - \sqrt{\frac{2}{3}} m\, M_{Pl}\, t\right).$$

Therefore,

$$\frac{1}{a}\, da = \frac{1}{\sqrt{6}}\frac{m}{M_{Pl}}\left(\phi_i - \sqrt{\frac{2}{3}} m\, M_{Pl}\, t\right) dt\,.$$

By integration, the previous expression yields

$$\ln a(t) = \frac{1}{\sqrt{6}}\frac{m}{M_{Pl}}\left(\phi_i\, t - \sqrt{\frac{2}{3}} m\, M_{Pl}\, \frac{t^2}{2}\right).$$

Simplifying the right side and applying e produces the following scale factor during inflation:

$$a(t) = a_i\, e^{\frac{1}{\sqrt{6}}\frac{m}{M_{Pl}}\left(\phi_i\, t - m\, M_{Pl}\, t^2\right)} \tag{4.2.9}$$

where a_i is the value of the scale factor at the beginning of inflation.

The total number of e-foldings from any time during inflation to the end of inflation is given by

$$\begin{aligned} N(t) &= \ln\left(\frac{a(t_{end})}{a(t)}\right) \\ &= \int_{t}^{t_{end}} H\, dt \end{aligned}$$

or, more usefully in terms of the value of the *inflaton field*, by

$$\begin{aligned} N(\phi) &\simeq \frac{1}{M_{Pl}^2}\int_{\phi_{end}}^{\phi}\frac{V}{V_\phi}\, d\phi \\ &= \frac{1}{2\,M_{Pl}^2}\int_{\phi_{end}}^{\phi}\phi\, d\phi \\ &= \frac{1}{4\,M_{Pl}^2}\left(\phi^2 - \phi_{end}^2\right) \\ &= \frac{1}{4\,M_{Pl}^2}\phi^2 - \frac{1}{4\,M_{Pl}^2}\left(\sqrt{2}M_{Pl}\right)^2 \\ &= \frac{1}{4\,M_{Pl}^2}\phi^2 - \frac{1}{2}\,. \end{aligned} \tag{4.2.10}$$

Solving for ϕ yields

$$\phi = M_{Pl}\sqrt{4\left(N + \frac{1}{2}\right)}. \tag{4.2.11}$$

Therefore, Eq. (4.2.11) can be used to determine the value of ϕ at any given e-foldings, N, before the end of inflation.

By Eq. (3.1.29), the curvature perturbation power spectrum is

$$\begin{aligned}\mathcal{P}_{\zeta}(k) &= \frac{1}{12\pi^2\, M_{Pl}^6}\frac{V^3(\phi)}{V_\phi^2}\bigg|_{k\sim aH} \\ &= \frac{1}{12\pi^2\, M_{Pl}^6}\frac{V^3(\phi)}{V_\phi^2(\phi)}\bigg|_{k\sim aH} \\ &= \frac{1}{12\pi^2\, M_{Pl}^6}\frac{\left(\frac{m^2(\phi)^2)}{2}\right)^3}{(m^2\phi)^2}\bigg|_{k\sim aH} \\ &= \frac{1}{96\pi^2\, M_{Pl}^6}m^2\phi^4\bigg|_{k\sim aH}. \end{aligned} \tag{4.2.12}$$

In particular, for pivot scale k_*, and A_s, the power spectrum based on that pivot scale, it follows that:

$$\begin{aligned}A_s &= \frac{1}{96\pi^2\, M_{Pl}^6}m^2\phi^4\bigg|_{k\sim aH} \\ &= \frac{1}{96\pi^2\, M_{Pl}^6}m^2{\phi_*}^4\end{aligned}$$

where ϕ_* is the value of ϕ at $k \sim a\,H$. Solving this equation for m yields the constant value of the free parameter:

$$m = \sqrt{\frac{96\pi^2 M_{Pl}^6\, A_s}{\phi_*^4}}\,. \tag{4.2.13}$$

If inflation begins at Planck time in a Planck-scale region with the potential energy density equal to Planck mass energy density, m_{Pl}^4 (in units with $c = \hbar = 1$, $m_{Pl}/\ell_{Pl}^3 = \sqrt{1/G}/\left(\sqrt{G}\right)^3 = \left(\sqrt{1/G}\right)^4 = m_{Pl}^4$), then

$$V(\phi_i) = \frac{1}{2}m^2\phi_i^2 = m_{Pl}^4 \tag{4.2.14}$$

and

$$\phi_i = \frac{\sqrt{2}m_{Pl}^2}{m}\,. \tag{4.2.15}$$

The value of the free parameter m is determined based on observable inflation which is in the range: $50 \leq N_* \leq 60$ (Eq (3.1.5)). The Planck Collaboration reports $A_s = 2.215{\times}10^{-9}$ [87, 88] as the value of the power spectrum for pivot scale $k_{0.05}$. The value of ϕ at the time the scale $k_{0.05}^{-1}$ reaches causal exit is $\phi = M_{Pl}\sqrt{222}$ based on Eq. (4.2.11) and $N_* = 55$, the mid-point of observable e-foldings. Applying Eq.(4.2.13)

$$\begin{aligned} m &= \sqrt{\frac{96\pi^2 M_{Pl}^6 A_s}{\left(M_{Pl}\sqrt{222}\right)^4}} \\ &= 6.53 \times 10^{-6}\, M_{Pl}\,. \end{aligned} \tag{4.2.16}$$

Therefore, from Eq. (4.2.15) and Eq. (4.2.16), the value of ϕ at the beginning of inflation is

$$\begin{aligned} \phi_i &= \frac{\sqrt{2}m_{Pl}^2}{m} \\ &= 5.45 \times 10^6 M_{Pl}\,. \end{aligned}$$

Substituting this into Eq. (4.2.10) yields

$$\begin{aligned} N(\phi_{end}) &= \frac{1}{4\,M_{Pl}^2}\left(5.45 \times 10^6 M_{Pl}\right)^2 - \frac{1}{2} \\ &\simeq 7.4 \times 10^{12} \\ &> 10^{12}\,. \end{aligned} \tag{4.2.17}$$

The resulting quantity of inflation exceeds $e^{10^{12}} \simeq 10^{10^{12}}$, i.e., one trillion orders of magnitude. Therefore, in this model with initial potential energy density equal to Planck energy

density in a region of Planck scale, the scale of the universe at the end of inflation is $10^{10^{12}} \times 10^{-33} cm \simeq 10^{10^{12}} cm$. In this scenario, the universe (observable and unobservable) begins with energy equivalent to about one milligram of matter ($\rho_{Pl} \times \ell_{Pl}^3$) and expands by a trillion orders of magnitude by the end of inflation.

The time at the end of inflation in this model may be determined using Eq. (4.2.8) and the values of ϕ_i and m from above. The result in the dimension of time in units with $c = \hbar = 1$ is

$$\begin{aligned} t_{end} &= \frac{\phi_i - \sqrt{2}\, M_{Pl}}{\sqrt{\frac{2}{3}} m\, M_{Pl}} \\ &\simeq \frac{5.45 \times 10^6 M_{Pl} - \sqrt{2}\, M_{Pl}}{\sqrt{\frac{2}{3}} (6.53 \times 10^{-6}\, M_{Pl})\, M_{Pl}} \\ &= \frac{5.45 \times 10^6 - \sqrt{2}}{\sqrt{\frac{2}{3}} (6.53 \times 10^{-6}\, M_{Pl})} \\ &\simeq 1.02226 \times 10^{12}\, \sqrt{8\pi G} \\ &\simeq 5.12486 \times 10^{12}\, \sqrt{G} \\ &\simeq 5.12486 \times 10^{12} \sqrt{\frac{\hbar\, G}{c^5}}\,. \end{aligned} \tag{4.2.18}$$

Using the values of the three physical constants in cgs units gives Planck time in units of seconds, $t_{Pl} = \sqrt{\frac{\hbar G}{c^5}} = 5.4 \times 10^{-44}$ seconds. Therefore, the elapsed time in seconds during inflation based on the harmonic oscillator model, the SRA parameters, and the assumption that the potential energy density of the scalar field at the beginning of inflation equals to Planck energy density is

$$\begin{aligned} t_{end} &= 5.12486 \times 10^{12} \sqrt{\frac{\hbar\, G}{c^5}} \\ &\simeq 10^{-31} \text{ seconds}\,. \end{aligned} \tag{4.2.19}$$

In this model, our observable universe was a tiny speck within this enormous volume at about 10^{-31} seconds. The speck grew to its present size of *only* $\sim 10^{28}\,cm$ (Eq. (2.3.20)) containing energy equivalent to $\sim 10^{57}$ grams of matter (Eq. (B.12.2)) after slowly expanding for another 14 billion years. See §B.2 *Local Energy-Momentum Conservation* for a discussion of the conservation of energy in dynamic spacetime. The first SRA parameter is (Eq. (4.2.2))

$$\begin{aligned}\epsilon &= 2\,M_{Pl}^2\frac{1}{\phi^2}\\ &= 2\,M_{Pl}^2\frac{1}{\left(M_{pl}\sqrt{222}\right)^2}\\ &= \frac{1}{111}\end{aligned} \tag{4.2.20}$$

and the second SRA parameter is (Eq. (4.2.3))

$$\eta = \epsilon = \frac{1}{111}\,. \tag{4.2.21}$$

From Eq. (3.1.35), the scalar spectral index is

$$\begin{aligned}n_s &= 1 - 6\epsilon + 2\eta\\ &= 0.9640\end{aligned} \tag{4.2.22}$$

and from Eq. (3.1.42) the ratio of the tensor mode amplitude parameter to the scalar mode amplitude parameter is

$$r = 16\epsilon = 0.144\,. \tag{4.2.23}$$

Note that, in the harmonic oscillator mode, ϵ and η are scale invariant since each is a function solely of N. Therefore, although n_s and r above were determined based on the

pivot scale $k_{0.05}$, the results would be identical based on $k_{0.002}$, the pivot scale illustrated in Figure 4.5.

Recalculating the spectral index and the tensor to scalar index using $N = 50$ and $N = 60$ yields the following table. The most recent best fit values are $n_s = 0.968 \pm 0.0006$, $r_{0.002} < 0.11$, and $r_{0.05} < 0.12$ from *Planck 2015 results. XIII. Cosmological parameters* and Joint Analysis of BICEP2/ *Keck Array* and *Planck* Data [1, 88].

Table 4.1: Harmonic Oscillator Model Predictions

N	n_s	$r_{0.002}$
50	0.9604	0.158
55	0.9640	0.144
60	0.9669	0.132

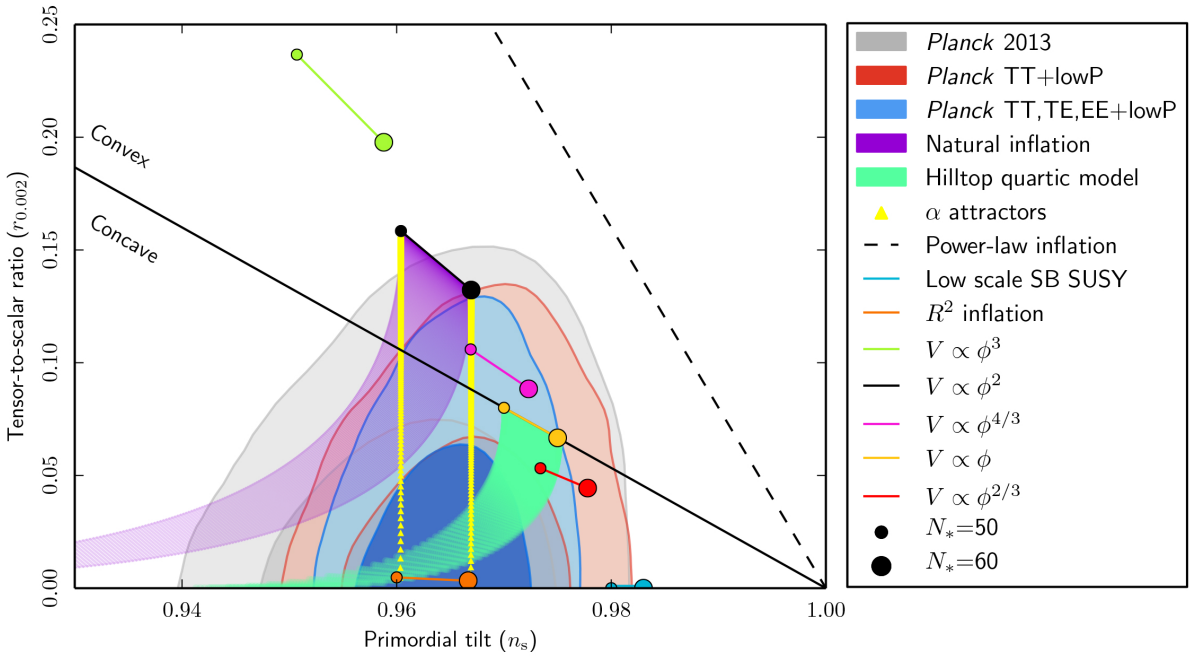


Figure 4.4: Marginalized joint 68 % and 95 % CL regions for n_s and $r_{0.002}$ from *Planck* including TE and EE cross correlations plus low multipole polarization data ($2 \leq \ell \leq 29$) compared to the theoretical predictions of selected inflationary models. Note that concave potentials are favored over convex potentials [89].

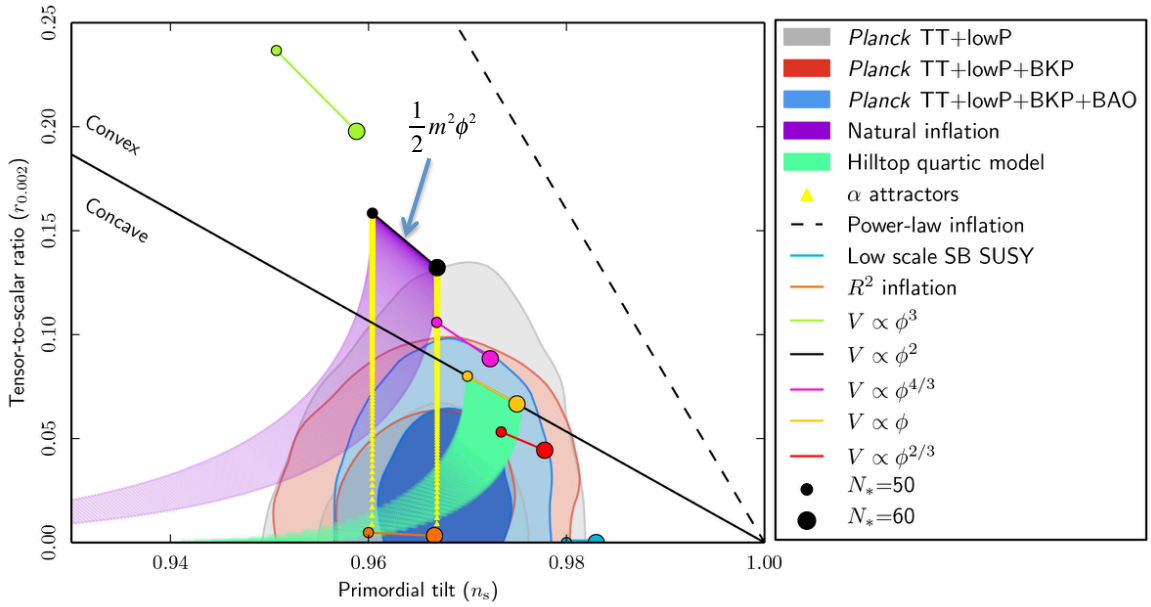


Figure 4.5: Marginalized joint 68 % and 95 % CL regions for n_s and $r_{0.002}$ from *Planck* measurements alone and in combination with its cross-correlation with *BICEP2/Keck Array* and/or *BAO* measurements compared with the theoretical predictions of selected inflationary models. Note that concave potentials are favored over convex potentials [89].

Models of inflation in Figures 4.4 and 4.5 are represented by dumbbells connecting the (n_s, r) value produced by the model based on 50 e-foldings of observable inflation to values based on 60 e-foldings. In Figure 4.5, the light and dark blue likelihood contours are based on analysis of Planck temperature and polarization data, BICEP2/Keck Array data, and baryon acoustic oscillation analysis. A model that falls in the center of the dark blue contour is the most favored model based on the analysis of these data. Such a model has the highest likelihood of being the "correct" model. The likelihood drops to 32% on the perimeter of the dark blue contour, 5% on the perimeter of the light blue contour, and decreases further as its distance outside the light contour increases.

The placement of a 60 e-folding disk closer to the center than the corresponding 50 e-folding disk should not be interpreted to mean that 60 e-foldings of observable inflation is more likely than 50 e-foldings. Consider the R^2 inflation orange dumbbell at the bottom

of each figure. The position of the 50 e-folding disk means that the likelihood of the R^2 model is less than 32% if there were 50 e-foldings of observable inflation. Alternatively, if there were 60 e-foldings of observable inflation, then the likelihood of the R^2 model is quite high.

As shown in Figure 4.5, the harmonic oscillator model $\left(\frac{1}{2}\, m^2\, \phi^2\right)$ produces values that are well outside the light blue contour. Therefore, the likelihood that the harmonic oscillator is the "correct" model is much less than 5%. Two models fall in the sweet spot near the center of the dark blue contour: the $R + R^2$ model (shown as R^2 inflation) and an α-attractor chaotic inflation model that is a variant of the harmonic oscillator model. The α-attractor variant of the harmonic oscillator will be described in §4.4 below.

4.3 Single Field Monomial Theories

The potential energy term in single field monomial theories may be written

$$V(\phi) = \lambda^2\, M_{Pl}^{4-n}\, \phi^n \tag{4.3.1}$$

[50, Eq. (21)]. Applying the procedures developed in §4.2, the slow-roll parameters are

$$\begin{aligned} \epsilon(\phi) &= \frac{M_{Pl}^2}{2}\left(\frac{V_\phi}{V}\right)^2 \\ &= \frac{M_{Pl}^2}{2}\left(\frac{n\,\lambda^2\, M_{Pl}^{4-n}\, \phi^{n-1}}{\lambda^2\, M_{Pl}^{4-n}\, \phi^n}\right)^2 \\ &= \frac{M_{Pl}^2}{2}\frac{n^2}{\phi^2} \end{aligned} \tag{4.3.2}$$

and

$$\begin{aligned}\eta(\phi) &= M_{Pl}^2 \frac{V_{\phi\phi}}{V} \\ &= M_{Pl}^2 \frac{n(n-1)\lambda^2 M_{Pl}^{4-n} \phi^{n-2}}{\lambda^2 M_{Pl}^{4-n} \phi^n} \\ &= M_{Pl}^2 \frac{n(n-1)}{\phi^2} . \end{aligned} \tag{4.3.3}$$

If energy density was dominated by potential energy at the beginning of inflation and energy density was approximately Planck density, then the initial value of ϕ in these theories may be determined by solving the following equation for ϕ where $m_{Pl} = \sqrt{8\pi} M_{Pl}$:

$$\begin{aligned}\lambda^2 M_{Pl}^{4-n} \phi^n &= m_{Pl}^4 \\ &= \left(\sqrt{8\pi} M_{Pl} \right)^4 \\ &= 64\,\pi^2 M_{Pl}^4 . \end{aligned} \tag{4.3.4}$$

Therefore,

$$\begin{aligned}\phi_i &= \left(\frac{64\,\pi^2 M_{Pl}^4}{\lambda^2 M_{Pl}^{4-n}} \right)^{\frac{1}{n}} \\ &= M_{Pl} \left(\frac{64\,\pi^2}{\lambda^2} \right)^{\frac{1}{n}} . \end{aligned} \tag{4.3.5}$$

The value of ϕ at the end of inflation, based on $\epsilon = 1$ and $\left|\eta\right| = 1$, is

$$\phi_{end} = \begin{cases} M_{Pl}\, n/\sqrt{2} & 0 < n \leq 2 \\ M_{Pl}\, \sqrt{n(n-1)} & n > 2 \end{cases} . \tag{4.3.6}$$

The number of e-foldings from some value of ϕ during inflation to the end of inflation is

$$\begin{aligned} N(\phi) &\simeq \frac{1}{M_{Pl}^2}\int_{\phi_{end}}^{\phi}\frac{V}{V_\phi}\,d\phi \\ &= \frac{1}{M_{Pl}^2}\frac{1}{n}\int_{\phi_{end}}^{\phi}\phi\,d\phi \\ &= \frac{1}{M_{Pl}^2}\frac{1}{2\,n}\left(\phi^2-\phi_{end}^2\right). \end{aligned} \tag{4.3.7}$$

This can be used to solve for ϕ at causal exit of a pivot scale based on a given number of observable e-foldings (e.g., $50-60$):

$$\phi_* = \begin{cases} M_{Pl}\,\sqrt{2\,n\,N+n^2/2} & 0<n\leq 2 \\ M_{Pl}\,\sqrt{2\,n\,N+n(n-1)} & n>2 \end{cases}. \tag{4.3.8}$$

Based on these results, the values of ϵ and η at causal exit of the pivot scale are

$$\epsilon = \begin{cases} \frac{n}{4\,N+n} & 0<n\leq 2 \\ \frac{n}{4\,N+2(n-1)} & n>2 \end{cases} \tag{4.3.9}$$

and

$$\eta = \begin{cases} \frac{2(n-1)}{4\,N+n} & 0<n\leq 2 \\ \frac{2(n-1)}{4\,N+2(n-1)} & n>2 \end{cases}. \tag{4.3.10}$$

Therefore, the predicted observables (n_s, r) are

$$n_s = \begin{cases} 1-\frac{2\,n+4}{4\,N+n} \simeq 1-\frac{n+2}{2\,N} & 0<n\leq 2 \\ 1-\frac{2\,n+4}{4\,N+2(n-1)} \simeq 1-\frac{n+2}{2\,N} & n>2 \end{cases} \tag{4.3.11}$$

and

$$r = \begin{cases} \frac{16\,n}{4\,N+n} \simeq \frac{4\,n}{N} & 0<n\leq 2 \\ \frac{16\,n}{4\,N+2(n-1)} \simeq \frac{4\,n}{N} & n>2 \end{cases}. \tag{4.3.12}$$

The value of r above rules out many theories based on the observable constraint $r < 0.11$. For example, the quadratic model, $V(\phi) = \lambda^2 \phi^2$, yields $r \simeq 0.144$ and the quartic model, $V(\phi) = \lambda^2 \phi^4$, yields $r \simeq 0.27$.

By Eq. (3.1.29), the value of the curvature perturbation power spectrum at horizon exit may be determined based on

$$\begin{aligned}
\mathcal{P}_\zeta(k) &= \frac{1}{12\pi^2 M_{Pl}^6} \frac{V^3(\phi)}{V_\phi^2}\bigg|_{k\sim aH} \\
&= \frac{1}{12\pi^2 M_{Pl}^6} \frac{V^3(\phi_*)}{V_\phi^2(\phi_*)} \\
&= \frac{1}{12\pi^2 M_{Pl}^6} \frac{\left(\lambda^2 M_{Pl}^{4-n} \phi_*^n\right)^3}{\left(n \lambda^2 M_{Pl}^{4-n} \phi_*^{n-1}\right)^2} \\
&= \frac{1}{12\pi^2 M_{Pl}^6} \frac{\lambda^6 M_{Pl}^{3(4-n)} \phi_*^{3n}}{n^2 \lambda^4 M_{Pl}^{2(4-n)} \phi_*^{2(n-1)}} \\
&= \frac{1}{12\pi^2 n^2 M_{Pl}^6} \lambda^2 M_{Pl}^{4-n} \phi_*^{n+2} \,. \qquad (4.3.13)
\end{aligned}$$

The SRA Equations (4.3.8) and (4.3.13) may be used to determine the free parameter, λ, based on the observable, A_s:

$$\frac{1}{12\pi^2 n^2 M_{Pl}^6} \lambda^2 M_{Pl}^{4-n} \phi_*^{n+2} = A_s \rightarrow \lambda^2 = \frac{12\pi^2 n^2 M_{Pl}^6 A_s}{M_{Pl}^{4-n} \phi_*^{n+2}} \,. \qquad (4.3.14)$$

Harmonic Oscillator Revisted. The harmonic oscillator is the second degree monomial

$$V(\phi) = \lambda^2 M_{Pl}^2 \phi^2 \,.$$

Applying the constraint in Eq. (4.3.14) yields

$$\lambda^2 = 2.12916 \times 10^{-11}$$

which is equivalent to

$$m = \sqrt{2\,\lambda^2} = 6.53 \times 10^{-6}$$

in §4.2 for $N = 55$. Applying Eq. (4.3.11) and Eq. (4.3.12) yield the same results as before:

$$n_s = 1 - 0.964, \quad r = 0.144\,.$$

The Hubble parameter is

$$H = M_{Pl}^{-1} \sqrt{\frac{\lambda^2 M_{Pl}^{4-n} \phi^n}{3}}\Big|_{n=2} = \frac{\lambda\,\phi}{\sqrt{3}}\,.$$

The equation of motion,

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

together with ϕ_i and ϕ_{end}, yields $t_{end} \simeq 10^{-31}$ seconds, as determined in Eq. (4.2.19).

4.4 Single Field Chaotic Inflation Theory in Supergravity

Overview. The 2013 Planck Collaboration reports [87, 2] effectively rule out many models of chaotic inflation. The reports favor the Starobinsky $R + R^2$ model. As a result, there has been renewed interest in slow-roll inflation theories based on an *inflaton field* that is non-minimally coupled to gravity. These are called supergravity models. One such model is the Starobinsky-Whitt model [107] that was developed in 1984. It is the forerunner of a general class of supergravity models of chaotic inflation called α-attractor E-theories. Another supergravity model, developed in 1984 [31], is the forerunner of another class of supergravity models of chaotic inflation called α-attractor T-theories.

Many articles focusing on α-attractors models have been published since the release of the 2013 Planck Collaboration reports. Versions of these models are in good agreement with 2013 Planck Collaboration reports since they can reproduce the Starobinsky $R + R^2$. The 2015 Planck Collaboration reports [5, 88, 3, 89, 1] include α-attractor chaotic theories as favored models.

The advantage of α-attractor theories over the Starobinsky $R + R^2$ model is that they are large field chaotic models of inflation that have natural initial conditions for inflation. These models do not require assumptions of initial homogeneity or isotropy and they do not require the inflationary domain to have non-trivial topology (e.g., a three-torus in the case of a flat domain).

These supergravity models will be described and corresponding predictions of (n_s, r) will be illustrated within the likelihood contours of current observations from the Planck Collaboration and the BICEP2/Keck Array.

Jordon Frame Lagrangians. A "Jordon frame" Lagrangian is a modified version of the Einstein-Hilbert Lagrangian. The Jordon frame includes a scalar field that is non minimally coupled to gravity though the Ricci scalar. These are a class of so-called supergravity models. This section will explore supergravity models of chaotic inflation based on a single scalar field in which a function of the field is coupled to the Ricci scalar through a coupling constant, ξ.

A generic single field model of chaotic inflation takes the following form:

$$V(\phi) = \lambda^2 f^2(\phi)\,. \tag{4.4.1}$$

The Lagrangian for empty space that is used to develop the Einstein Equation in §B.1 *Einstein's Equation for Gravitation* is given in Eq. (B.1.1) as

$$\frac{\mathcal{L}}{\sqrt{-g}} = \frac{1}{2\kappa}R \tag{4.4.2}$$

where $\kappa = 8\,\pi\,G = M_{Pl}^2$. For consistency with current literature, this section is presented in units where $M_{Pl} = 1$.

The Jordon frame Lagrangian for the supergravity models of chaotic inflation may be written

$$\begin{aligned}\frac{\mathcal{L}}{\sqrt{-g}} &= \frac{1}{2}R + \frac{\xi}{2}\,f(\phi)\,R - \frac{1}{2}\,g^{\mu\nu}\,\partial_\nu\,\phi\,\partial_\mu\,\phi - \lambda^2\,f^2(\phi) \\ &= \frac{1}{2}R + \frac{\xi}{2}\,f(\phi)\,R - \frac{1}{2}\,g^{00}\,\partial_0\,\phi\,\partial_0\,\phi - \frac{1}{2}\,g^{ij}\,\partial_i\,\phi\,\partial_j\,\phi - \lambda^2\,f^2(\phi) \\ &= \frac{1}{2}R + \frac{\xi}{2}\,f(\phi)\,R - \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}\,(\nabla\,\phi)^2 - \lambda^2\,f^2(\phi)\end{aligned} \tag{4.4.3}$$

where ordinary partial derivatives are used in place of covariant derivatives. This is permissible since the covariant derivatives operate on scalar functions and are equivalent to ordinary derivatives in this context. The term ξ is the coupling constant of the scalar field to gravity. Note that space is no longer empty. It includes an unspecified scalar field: the *inflaton field*. It will be assumed that the gradient energy term is negligible since the universe becomes nearly isotropic early during inflation. The kinetic energy terms will be written $\frac{1}{2}(\partial\,\phi)^2$ for consistency with literature. The "Jordon frame" potential energy is defined as

$$V_J(\phi) = \lambda^2\,f^2(\phi) \tag{4.4.4}$$

and

$$\Omega(\phi) \equiv 1 + \xi f(\phi) \tag{4.4.5}$$

[65, Eq. (13.5)]. Putting all this together yields the following form of the Jordon frame Lagrangian which will be labeled $\mathcal{L}_J$:

$$\frac{\mathcal{L}_J}{\sqrt{-g}} = \frac{1}{2}\Omega(\phi)R - \frac{1}{2}\left(\partial\phi\right)^2 - V_J(\phi) \tag{4.4.6}$$

[65, Eq. (13.4)]. An equivalent Jordon frame Lagrangian that highlights the effective potential energy (here in brackets) may be written

$$\frac{\mathcal{L}}{\sqrt{-g}} = \frac{1}{2}R - \frac{1}{2}\left(\partial\phi\right)^2 - \left[\lambda^2 f^2(\phi) - \frac{\xi}{2} f(\phi) R\right]. \tag{4.4.7}$$

If $\xi = 0$, then the Lagrangian reduces to the Lagrangian of a minimally coupled chaotic inflation model. Therefore, $\xi \neq 0$ in the models considered in this section.

The $\mathcal{L}_J$ supergravity models defined in Eq. (4.4.6) and restated in an Einstein frame under conformal transformation in Eq. (4.4.18) are sometimes called ξ-attractor theories. These are part of a broader category of models called α-attractor E-theories with α=1. α-attractor theories are a part of an even broader class of inflationary theories called "cosmological attractors" [19] that also includes α-attractor T-theories. Both α-attractor E-theories and α-attractor T-theories can be constructed to produce inflation observables that are consistent with current observations.

Under the most recent formulation (July 2015) of these models in Ref. [19], the initial value of the *inflaton field* is extremely large. Therefore, the models retain the "natural" initial conditions for the original large-field models of chaotic inflation: the potential energy is close to Planck energy density, driving the kinetic and gradient terms to near zero values within a Planck-scale inflationary domain. See Figure 4.10 from Ref. [48] for a representation of the potential energy of an α-attractor T-theory.

An *effective supergravity theory* is constructed so that the free parameters (e.g., ξ and λ in $\mathcal{L}_J$ in the supergravity model above) produce potential energy that inflates and predicts inflation observables (A_s, n_s, and r) that are consistent with observations. Additionally, the

inflaton field in an effective theory must have strong enough coupling to matter and radiation fields to thermalize the fields and produce energy density of ordinary particles and dark matter at the end of re-heating that is consistent with the observed components of energy density in our universe. The construction of an effective α-attractor theory requires the introduction of supersymmetric particles and symmetry breaking. A detailed description of this process is given in Ref. [19].

Effective $\mathcal{L}_J$ supergravity models in the Jordon frame as shown in Eq. (4.4.6) above or restated in an Einstein frame under conformal transformation as shown in Eq. (4.4.18) below, produce n_s and r for observable e-foldings, N, in the $50 - 60$ range [49] in the following simple form:

$$n_s \simeq 1 - \frac{2}{N}, \quad r \simeq \frac{12}{N^2}\,. \tag{4.4.8}$$

For example, for $N = 60$,

$$n_s = 1 - 0.967, \quad r = 0.003 \tag{4.4.9}$$

and for $N = 50$,

$$n_s = 1 - 0.960, \quad r = 0.005\,. \tag{4.4.10}$$

These results fit into the "sweet spot" shown in Figures 4.4 and 4.5. It will be shown later in this section that the Starobinsky-Whitt model is a supergravity version of the harmonic oscillator. It will also be seen that the supergravity version of the quartic chaotic model can be constructed to predict the Planck Collaboration's best fit values of inflation observables.

Conformal Transformation to the Einstein Frame. The $\mathcal{L}_J$ Lagrangian may be expressed in the Einstein frame under the conformal transformation, $\tilde{g}_{\mu\nu} = \omega^2(\phi) g_{\mu\nu}$, where

$\omega(\phi) = \sqrt{\Omega(\phi)}$, as

$$\frac{\mathcal{L}_E}{\sqrt{-\tilde{g}}} = \frac{1}{2}\tilde{R} - \frac{1}{2}\left[\Omega(\phi)^{-1} + \frac{3}{2}\left(\frac{\partial}{\partial\phi}\ln\Omega(\phi)\right)^2\right](\partial\phi)^2 - \frac{V_J(\phi)}{\Omega(\phi)^2} \tag{4.4.11}$$

[65, Eq. (13.6)]. See Eq. (A.3.12) for details. The first bracketed term may be neglected if it is much smaller than the second bracketed term. This requirement may be written

$$\Omega(\phi)^{-1} \ll \frac{3}{2}\left(\frac{\partial}{\partial\phi}\ln\Omega(\phi)\right)^2 = \frac{3}{2}\Omega^{-2}\left(\frac{\partial}{\partial\phi}\Omega(\phi)\right)^2 \tag{4.4.12}$$

or, equivalently,

$$\frac{3}{2}\left(\frac{\partial}{\partial\phi}\Omega(\phi)\right)^2 \gg \Omega(\phi)\,. \tag{4.4.13}$$

Since

$$\frac{\partial}{\partial\phi}\left(\ln\Omega(\phi)\right)\partial\phi = \Omega(\phi)^{-1}\partial(\Omega(\phi)\partial\phi = \partial\ln\Omega(\phi)\,, \tag{4.4.14}$$

the Lagrangian may be written

$$\frac{\mathcal{L}_E}{\sqrt{-\tilde{g}}} = \frac{1}{2}\tilde{R} + \frac{3}{4}\left(\partial\ln\Omega(\phi)\right)^2 - \frac{V_J(\phi)}{\Omega(\phi)^2}\,. \tag{4.4.15}$$

By defining

$$\varphi \equiv \pm\sqrt{\frac{3}{2}}\ln\left(\Omega(\phi)\right), \tag{4.4.16}$$

$\Omega(\phi)$ may be written

$$\Omega(\phi) = e^{\sqrt{\frac{2}{3}}\varphi}\,. \tag{4.4.17}$$

Applying the definitions of $\Omega(\phi)$, $f(\phi)$, and φ, the Lagrangian in the Einstein frame may be written

$$\begin{aligned}\frac{\mathcal{L}_E}{\sqrt{-\tilde{g}}} &= \frac{1}{2}\tilde{R} + \frac{1}{2}(\partial\varphi)^2 - \frac{\lambda^2 f^2(\phi)}{\Omega(\phi)^2} \\ &= \frac{1}{2}\tilde{R} + \frac{1}{2}(\partial\varphi)^2 - \frac{\lambda^2 (\Omega(\phi)-1)^2}{\xi^2\Omega(\phi)^2} \\ &= \frac{1}{2}\tilde{R} + \frac{1}{2}(\partial\varphi)^2 - \frac{\lambda^2}{\xi^2}\left(1-\Omega^{-1}(\phi)\right)^2 \\ &= \frac{1}{2}\tilde{R} + \frac{1}{2}(\partial\varphi)^2 - \frac{\lambda^2}{\xi^2}\left(1-e^{-\sqrt{\frac{2}{3}}\varphi}\right)^2 \end{aligned} \tag{4.4.18}$$

[65, Eq. (13.10)] where the second step follows from Eq. (4.4.5).

α-attractor E-theories. The Einstein frame Lagrangian may be extended to include an additional degree of freedom, α, as

$$\frac{\mathcal{L}_E}{\sqrt{-\tilde{g}}} = \frac{1}{2}\tilde{R} + \frac{1}{2}(\partial\varphi)^2 - \alpha\,\frac{\lambda^2}{\xi^2}\left(1-e^{-\sqrt{\frac{2}{3\alpha}}\varphi}\right)^2 \tag{4.4.19}$$

where $\alpha \leq 1$ for *effective* theories. The potential energy of these theories may be written

$$V(\varphi) = \alpha\,\mu^2\left(1-e^{-\sqrt{\frac{2}{3\alpha}}\varphi}\right)^2 \tag{4.4.20}$$

where $\mu = \lambda/\xi$ [19, Eq. (5.5)]. The predicted values of the cosmological parameters, n_s and r, in an *effective* α-attractor E-theory take the simple form for observable e-foldings, N, in the $50-60$ range [19, 92, 70, 49, 26]:

$$n_s = 1 - \frac{2}{N}, \quad r = \alpha\frac{12}{N^2} \tag{4.4.21}$$

where $\alpha \leq 1$. Note that the predictions of the observables do not depend on the form of V_J. However, the value of the free parameters, μ and α, depend on the form of V_J.

Figure 4.6 from Ref. [48] shows the shapes of the potential energy curve based on several values of α. Figure 4.7 from Ref. [48] shows the predicted values of n_s and r in these theories over a large range of α within the likelihood contours of current observations from the Planck Collaboration and the BICEP2/Keck Array. The theory converges to the harmonic oscillator theory, $V(\phi) = \frac{1}{2} m^2 \phi^2$, at large α shown at the upper red star. The theory converges to $R + R^2$ theory at $\alpha = 1$ shown at the lower blue star.

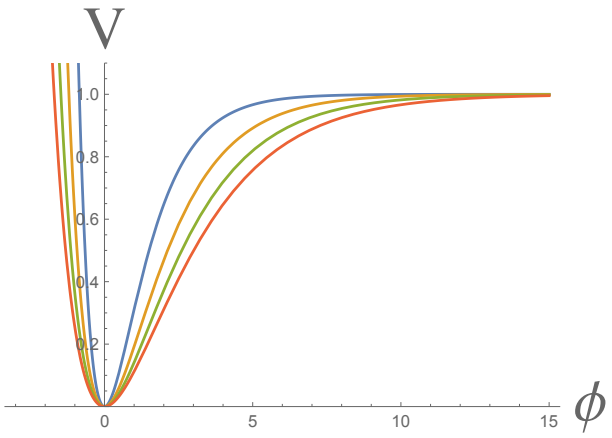


Figure 4.6: α-attractor E-theory potential energy curves, $\alpha\mu^2(1 - e^{-\sqrt{\frac{2}{3\alpha}}\varphi})^2$, in units of $\alpha\mu^2 = 1$ for $\alpha = 1, 2, 3, 4$. Smaller α correspond to more narrow minima of the potentials. The blue line shows the potential of the Starobinsky model, which belongs to the class of E-models with $\alpha = 1$ [48].

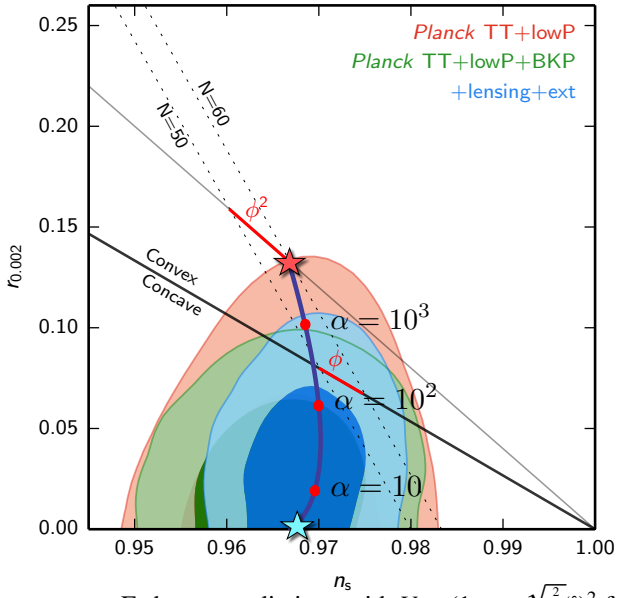


Figure 4.7: α-attractor E-theory predictions with $V \sim (1 - e^{-\sqrt{\frac{2}{3\alpha}}\varphi})^2$ for $N_* = 60$ [48].

α-attractor T-theories. These theories have the Einstein frame Lagrangian:

$$\frac{\mathcal{L}_E}{\sqrt{-\tilde{g}}} = \frac{1}{2}\tilde{R} + \frac{1}{2}(\partial\varphi)^2 - \alpha\,\mu^2\,f^2\left(\tanh^n\frac{\varphi}{\sqrt{6\alpha}}\right) \tag{4.4.22}$$

with potential energy

$$V(\varphi) = \alpha\,\mu^2\,f^2\left(\tanh\frac{\varphi}{\sqrt{6\alpha}}\right) \tag{4.4.23}$$

[19, Eq. (3.2)]. In the simplest form of these models, $f(\phi) = \phi^n$, and the potential energy is

$$V(\varphi) = \alpha\,\mu^2\tanh^{2\,n}\frac{\varphi}{\sqrt{6\alpha}} \tag{4.4.24}$$

[19, Eq. (4.2)], analogous to the monomial potentials $V(\phi) = \lambda^2\,\phi^n$ in units where $M_{Pl} = 1$. Effective α-attractor T-theories with potential in this form predict the same values of the cosmological parameters, n_s and r, as do effective α-attractor E-theories:

$$n_s = 1 - \frac{2}{N}, \quad r = \alpha\frac{12}{N^2} \tag{4.4.25}$$

[19, Eq. (4.4)].

Figure 4.8 from Ref. [48] shows the shapes of the potential energy curve based on several values of α. Figure 4.9 from Ref. [48] shows the predicted values of n_s and r in these theories over a large range of α within the likelihood contours of current observations from the Planck Collaboration and the BICEP2/Keck Array. The theory converges to the harmonic oscillator theory, $V(\phi) = \frac{1}{2}\,m^2\,\phi^2$, at large α shown at the upper red star. The theory converges to $R + R^2$ theory at $\alpha = 1$ shown at the lower blue star.

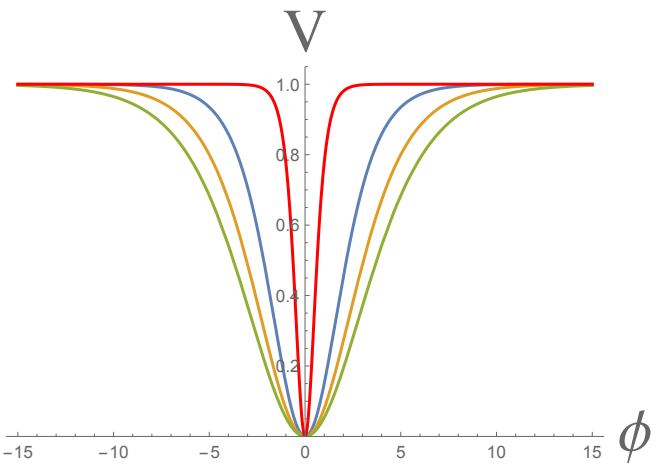


Figure 4.8: α-attractor T-theory potential energy curves, $V(\varphi) = \alpha\mu^2 \tanh^2 \frac{\varphi}{\sqrt{6\alpha}}$ with $\mu = 1$ and $\alpha = 1/9, 1, 2, 3$ corresponding to the red, blue, brown and green lines [48].

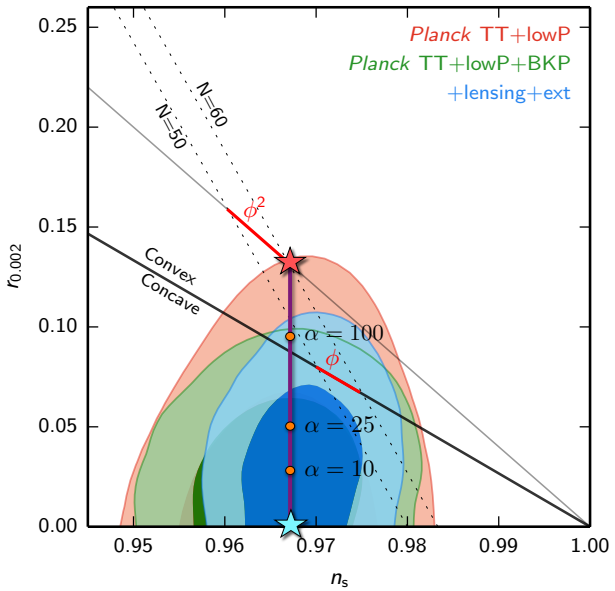


Figure 4.9: Predictions of the simplest α-attractor T-model with the potential $V \sim \tanh^2 \frac{\varphi}{\sqrt{6\alpha}}$ for N = 60 cut through the most interesting part of the Planck 2015 plot for n_s and r [48, 88].

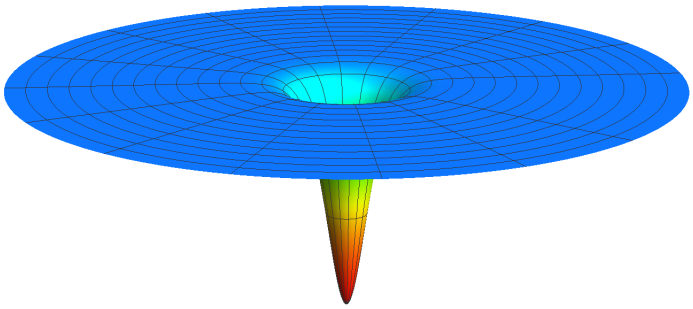


Figure 4.10: The potential energy of an α-attractor T-theory with $\alpha = 1/3$ as a function of the inflaton field, φ. The value of φ vanishes at the center and increases radially from the center over the disk to reach enormous value. The potential energy and hence the *inflaton field* is flat with exponentially good accuracy except near the center, providing natural large-field initial conditions for inflation like those for minimally coupled large field chaotic theories such as $V \propto \phi^2$ and $V \propto \phi^4$ [48].

$\mathbf{R + R^2}$ Starobinsky Model. The Planck Collaboration measurements favor the $R + R^2$ Starobinsky model of inflation developed in the context of quantum gravity by Starobinsky in 1980 [95] and later presented in the classical form below by Kofman, Linde, and Starobinsky in 1985 [52]. The $R + R^2$ model is not a scalar field model of inflation. Switching to units with $M_{Pl} = 1/\sqrt{8\pi G} = 1$, the $R + R^2$ model is based on the Einstein-Hilbert Lagrangian for empty space in Eq.(B.1.1):

$$\frac{\mathcal{L}}{\sqrt{-g}} = \frac{1}{2\kappa} R \equiv M_{Pl}^2 R \equiv R \tag{4.4.26}$$

modified to include a second order term of the Ricci scalar:

$$\frac{\mathcal{L}_{Starobinsky}}{\sqrt{-g}} = \frac{1}{2}\left[R + \frac{R^2}{6\,M^2}\right]. \tag{4.4.27}$$

The $R + R^2$ model includes an early, brief epoch of rapid expansion. The theory requires the universe to be homogeneous and isotropic from its beginning. Starobinsky emphasized that his model was the opposite of Charles Misner's "mixmaster universe" in which Misner attempted to solve the horizon problem based on an initial chaotic epoch.

In 1984, Brian Whitt published a conformally equivalent version of the Starobinsky model based on a single scalar field that is non-minimally coupled to R [107]. The

Starobinsky-Whitt Lagrangian [107, 65] is derived from Eq. (4.4.27) under the conformal transformation $\tilde{g}_{\mu\nu} = \left(1 + \frac{\phi}{3M^2}\right) g_{\mu\nu}$:

$$\frac{\mathcal{L}_{Starobinsky-Whitt}}{\sqrt{-\tilde{g}}} = \frac{1}{2}\tilde{R} - \frac{1}{2}(\partial\varphi)^2 - \frac{3}{4}M^2\left(1 - e^{\sqrt{\frac{2}{3}}\varphi}\right)^2 . \tag{4.4.28}$$

This Lagrangian is an α-attractor E-theory based on the harmonic oscillator model that was rejected as a viable theory of inflation without supergravity in §4.2.

The concave shape of the potential energy term is shown in Figure 4.11. Compare this to the α-attractor E-theory potential energy curves in Figure 4.6.

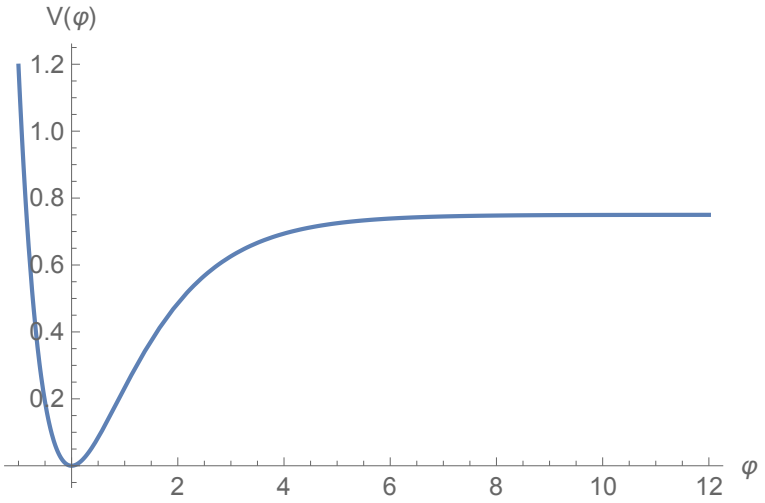


Figure 4.11: Starobinsky-Whitt Potential Energy, $M = 1$.

To derive the α-attractor model of the Starobinsky-Whitt Lagrangian, let

$$\begin{aligned}
\alpha &= 1\,, \\
\xi &= \frac{1}{3M^2}\,, \\
f(\phi) &= \phi\,, \\
\Omega(\phi) &= 1 + \xi\, f(\phi)\,, \\
\lambda &= \frac{1}{\sqrt{12}M}\,, \text{ and} \\
\mu &= \frac{\lambda}{\xi}\,.
\end{aligned}$$

The Jordon frame Lagrangian from Eq. (4.4.6) is

$$\begin{aligned}\frac{\mathcal{L}_J}{\sqrt{-g}} &= \frac{1}{2}\Omega(\phi)R - \frac{1}{2}\left(\partial\phi\right)^2 - V_J(\phi) \\ &= \frac{1}{2}R - \frac{1}{2}\left(\partial\phi\right)^2 - \left(V_J(\phi) - \frac{\xi}{2}\,\phi\,R\right) \\ &= \frac{1}{2}R - \frac{1}{2}\left(\partial\phi\right)^2 - \left(\frac{1}{12\,M^2}\phi^2 - \frac{1}{6\,M^2}\,\phi\,R\right)\end{aligned} \tag{4.4.29}$$

where the minimally coupled potential energy in the Jordan frame is quadratic:

$$\begin{aligned}V_J(\phi) &= \lambda^2\,f^2(\phi) \\ &= \frac{1}{12\,M^2}\phi^2\end{aligned} \tag{4.4.30}$$

and the effective potential energy is

$$\frac{1}{12\,M^2}\phi^2 - \frac{1}{6\,M^2}\,\phi\,R\,.$$

The condition in Eq. (4.4.13) is satisfied after causal exit and well before the end of inflation. Applying the conformation transformation $\tilde{g}_{\mu\nu} = \Omega\,g_{\mu\nu}$ and Eq. (4.4.18), the Lagrangian in the Einstein frame is

$$\begin{aligned}\frac{\mathcal{L}_E}{\sqrt{-\tilde{g}}} &= \frac{1}{2}\tilde{R} + \frac{1}{2}(\partial\varphi)^2 - \alpha\mu^2\left(1 - e^{-\sqrt{\frac{2}{3\alpha}}\varphi}\right)^2 \\ &= \frac{1}{2}\tilde{R} + \frac{1}{2}(\partial\varphi)^2 - \frac{\lambda^2}{\xi^2}\left(1 - e^{-\sqrt{\frac{2}{3}}\varphi}\right)^2 \\ &= \frac{1}{2}\tilde{R} + \frac{1}{2}(\partial\varphi)^2 - \frac{\left(\frac{1}{\sqrt{12}M^2}\right)^2}{\left(\frac{1}{3M^2}\right)^2}\left(1 - e^{-\sqrt{\frac{2}{3}}\varphi}\right)^2 \\ &= \frac{1}{2}\tilde{R} + \frac{1}{2}(\partial\varphi)^2 - \frac{3}{4}M^2\left(1 - e^{-\sqrt{\frac{2}{3}}\varphi}\right)^2\end{aligned} \tag{4.4.31}$$

reproducing the Starobinsky-Whitt Lagrangian in Eq. (4.4.28). The values for n_s and r are as shown in Equations (4.4.9) and (4.4.10).

Therefore, the quadratic potential, $V_J(\phi) = \lambda^2 f^2(\phi)$, and the parameters ξ and λ produce the Starobinsky-Whitt model that is favored by the most recent analysis of the Planck Collaboration. The bottom of the vertical yellow lines in Figures 4.4 and 4.5 demonstrate that α-attractor models based on a harmonic oscillator can be constructed to predict the inflation observables, $(n_s, r)\Big|_{N=50}$ and $(n_s, r)\Big|_{N=60}$, that match the predictions of the Starobinsky $R + R^2$ model.

The $V(\phi) = \lambda^2 \phi^4$ Model in Supergravity. This quartic model was the first chaotic inflation model presented in Linde's 1983 article [73]. The potential was given in the article as $\lambda \phi^4/4$; therefore, $\lambda^2 = \lambda_{1983}/4$. Based on Eq. (4.3.12), $r = 16/60 \simeq 0.27$, placing its predictions well outside the range of models consistent with current observational data. However, with a small coupling to gravity, $-\frac{\xi}{2}\phi^2 R$, where $\xi > 0.1$, the results begin to converge to the Starobinksy-Whitt model as seen in Figure 4.12.

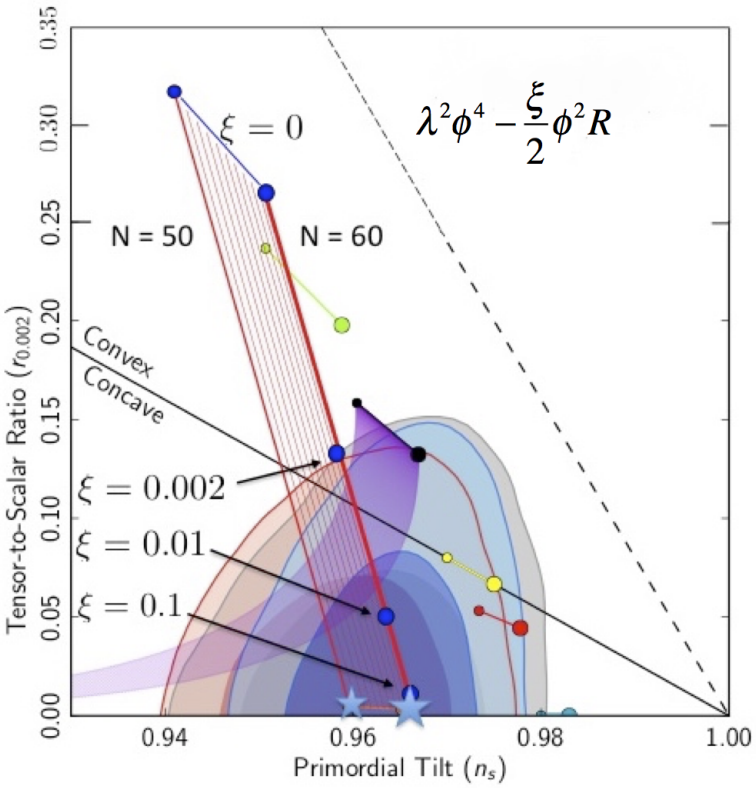


Figure 4.12: Red stripes show inflation parameters n_s and r for the model $\lambda^2 \phi^4$ for different values of ξ in the term $\xi\phi^2 R/2$. The top range corresponds to $\xi = 0$. Different stripes correspond to different number of e-foldings N. The right one corresponds to $N = 60$, the left one corresponds to $N = 50$. The stars show the predictions of the Starobinsky model, that coincide with the predictions of the model $\lambda^2 \phi^4$ with $\xi \gg 0.1$ [65]

The non-minimally coupled quartic model is also an α-attractor E-model with

$$f(\phi) = \phi^2 ,$$

$$\Omega(\phi) = 1 + \xi f(\phi) ,$$

$$\alpha = 1$$

and free parameters λ and ξ, where $\mu = \lambda/\xi$.

The Lagrangian in the Jordon frame is

$$\begin{aligned}\frac{\mathcal{L}_J}{\sqrt{-g}} &= \frac{1}{2}\Omega(\phi)R - \frac{1}{2}\left(\partial\phi\right)^2 - V_J(\phi) \\ &= \frac{1}{2}\left(1+\xi f(\phi)\right)R - \frac{1}{2}\left(\partial\phi\right)^2 - \lambda^2 f^2(\phi) \\ &= \frac{1}{2}\left(1+\xi\,\phi^2\right)R - \frac{1}{2}\left(\partial\phi\right)^2 - \lambda^2\,\phi^4 \\ &= \frac{1}{2}R - \frac{1}{2}\left(\partial\phi\right)^2 - \left(\lambda^2\,\phi^4 - \frac{\xi}{2}\,\phi^2\,R\right). \end{aligned} \tag{4.4.32}$$

Applying Eq. (4.4.18), the equivalent Lagrangian in the Einstein frame under the conformal transformation, $\tilde{g}_{\mu\nu} = \Omega(\phi)\, g_{\mu\nu}$, is

$$\begin{aligned}\frac{\mathcal{L}_E}{\sqrt{-\tilde{g}}} &= \frac{1}{2}\tilde{R} + \frac{1}{2}(\partial\varphi)^2 - \frac{\lambda^2}{\xi^2}\left(1 - e^{-\sqrt{\frac{2}{3}}\varphi}\right)^2 \\ &\equiv \frac{1}{2}\tilde{R} + \frac{1}{2}(\partial\varphi)^2 - \alpha\,\mu^2\left(1 - e^{-\sqrt{\frac{2}{3\alpha}}\varphi}\right)^2 \end{aligned} \tag{4.4.33}$$

where $\alpha = 1$ and $\mu = \lambda/\xi$.

For models that produce large values of ξ, the coefficient $\mu^2 = \frac{\lambda^2}{\xi^2}$, is normalized by the curvature perturbation power spectrum (and if ξ is not large, then the coefficients may be adjusted to satisfy large ξ [50]):

$$\frac{\lambda^2}{\xi^2} \sim A_s\,. \tag{4.4.34}$$

The construction of an effective supergravity theory based on the quartic model, including the determination of ξ, the coupling constant of the *inflaton field* to gravity, was developed in 2013 by Bezrukov, F. and Gorbunov, D. in *Light inflaton after LHC8 and WMAP9 results* [15] based on supersymmetry using methods of quantum field theory. They found that coupling constants $\xi \gtrsim 0.002$ are consistent with observational data.

An effective supergravity theory based on the quartic model was developed in 2014 in a more general context and reported in *Superconformal generalization of the chaotic inflation model* $\frac{\lambda}{4}\phi^4 - \frac{\xi}{2}\phi^2 R$ [46] by Renata Kallosh and Andrei Linde. They found that coupling constants $\xi \gtrsim 0.002$ are consistent with observational data and that $\xi \gg 0.1$ matches the Starobinsky-Whitt model.

Recent analyses of α-attractor and other parameterized supergravity models of chaotic inflation may be found in:

- *Cosmological attractors from α-scale supergravity* [92] published on September 16, 2015 by Diederik Roest and Marco Scalisi;
- *Escher in the Sky* [48] published on May 5, 2015 by Renata Kallosh, and Andrei Linde;
- *Cosmological Attractors and Initial Conditions for Inflation* [19], published on July 17, 2015 by John Joseph M Carrasco, Renata Kallosh, and Andrei Linde; and
- *Unity of Cosmological Inflation Attractors* [28], published on February 2, 2015 by Mario Galante, Renata Kallosh, Andrei Linde, and Diederik Roest;

Based on these new methods, effective theories of chaotic inflation based on cosmological attractors can be constructed to predict virtually any set of inflation observables.

Theoretical research spurred by the reports of the Planck Collaboration is continuing at high speed waiting for the next set of observations to provide additional direction. Observations from the new BICEP3 instrument are expected in 2016 possibly finding stronger signals of tensor mode perturbations and thereby confirming the existence of gravitational waves. The Large Hadron Collider is back online. Observations confirming or ruling out light supersymmetric particles, including dark matter, may be forthcoming. Such results will shape advances in the understanding of the pre-heating and re-heating processes that produced all elementary particles during a small fraction of a second at the end of inflation.

Advances in the understanding of these processes may lead to an improved understanding of the imbalance between matter and anti-matter that produced our universe.

References

[1] P. A. R. Ade, N. Aghanim, and et al. Joint analysis of bicep2/ *Keck Array* and *Planck* data. *Phys. Rev. Lett.*, 114:101301, Mar 2015.

[2] P. A. R. Ade et al. Planck 2013 results. XXII. Constraints on inflation. *Astron. Astrophys.*, 571:A22, 2014.

[3] P. A. R. Ade et al. Planck 2015 results. XVII. Constraints on primordial non-Gaussianity. 2015.

[4] Rober J Adler and Johnathan E Taylor. *Random Fields and Geometry*. Springer, 2007.

[5] N. Aghanim et al. Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters. *Submitted to: Astron. Astrophys.*, 2015.

[6] A. Albrecht and P. J. Steinhardt. Cosmology for grand unified theories with radiatively induced symmetry breaking. *Physical Review Letters*, 48(17):1220–1223, April 1982.

[7] A. Arbey. Alterbbn: A program for calculating the {BBN} abundances of the elements in alternative cosmologies. *Computer Physics Communications*, 183(8):1822 – 1831, 2012.

[8] A. Arbey and F. Mahmoudi. SUSY constraints from relic density: High sensitivity to pre-BBN expansion rate. *Phys.Lett.*, B669:46–51, 2008.

[9] A. Arbey and F. Mahmoudi. SUSY Constraints, Relic Density, and Very Early Universe. *JHEP*, 1005:051, 2010.

[10] Alexandre Arbey, Aldo Deandrea, and Ahmad Tarhini. Anomaly mediated SUSY breaking scenarios in the light of cosmology and in the dark (matter). *JHEP*, 1105:078, 2011.

[11] Arfken, Webber, and Harris. *Mathematical Methods for Physicists*. Academic Press, seventh edition, 2013.

[12] James M. Bardeen, Paul J. Steinhardt, and Michael S. Turner. Spontaneous creation of almost scale-free density perturbations in an inflationary universe. *Phys. Rev. D*, 28:679–693, Aug 1983.

[13] Brian J. Barris and et al. Twenty-three high-redshift supernovae from the institute for astronomy deep survey: Doubling the supernova sample at $z > 0.7$. *The Astrophysical Journal*, 602(2):571, 2004.

[14] N. Bartolo, S. Matarrese, and A. Riotto. Non-gaussianity in the curvaton scenario. *Phys. Rev. D*, 69:043503, Feb 2004.

[15] F. Bezrukov and D. Gorbunov. Light inflaton after LHC8 and WMAP9 results. *JHEP*, 07:140, 2013.

[16] Emory F. Bunn, Andrew R. Liddle, and Martin White. Four-year cobe normalization of inflationary cosmologies. *Phys. Rev. D*, 54:R5917–R5921, Nov 1996.

[17] Emory F. Bunn, Matias Zaldarriaga, Max Tegmark, and Angelica de Oliveira-Costa. *E/B* decomposition of finite pixelized cmb maps. *Phys. Rev. D*, 67:023501, Jan 2003.

[18] A.J. Buras, John R. Ellis, M.K. Gaillard, and Dimitri V. Nanopoulos. Aspects of the Grand Unification of Strong, Weak and Electromagnetic Interactions. *Nucl.Phys.*, B135:66–92, 1978.

[19] John Joseph M. Carrasco, Renata Kallosh, and Andrei Linde. Cosmological Attractors and Initial Conditions for Inflation. 2015.

[20] Sean M. Carroll. The cosmological constant. *Living Rev. Relativity*, 3(1), 2001.

[21] Sean M. Carroll. *Spacetime and Geometry*. Addison Wesley, 2004.

[22] J. Cline. Baryogenesis. *Cornell University Print Archives: High Energy Physics - Phenomlenology*, arXiv:hep-ph/0609145v3, 2006.

[23] Douglas Clowe, Marusa Bradac, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, et al. A direct empirical proof of the existence of dark matter. *Astrophys.J.*, 648:L109–L113, 2006.

[24] A Einstein, H Lorentz, H Minkowski, and H Weyl. *The Principle of Relativity: a collection of original memoirs on the special and general theory of relativity*. Courier Dover Plublications, 1952.

[25] F. Farakos, A. Kehagias, and A. Riotto. On the starobinsky model of inflation from supergravity. *Nuclear Physics B*, 876(1):187 – 200, 2013.

[26] Sergio Ferrara, Renata Kallosh, Andrei Linde, and Massimo Porrati. Minimal Supergravity Models of Inflation. *Phys. Rev.*, D88(8):085038, 2013.

[27] S. L. Finkelstein and et al. A galaxy rapidly forming stars 700 million years after the big bang at redshift 7.51. *Nature*, 502:524–527, 2013.

[28] Mario Galante, Renata Kallosh, Andrei Linde, and Diederik Roest. Unity of Cosmological Inflation Attractors. *Phys. Rev. Lett.*, 114(14):141302, 2015.

[29] Howard Georgi and S. L. Glashow. Unity of all elementary-particle forces. *Phys. Rev. Lett.*, 32:438–441, Feb 1974.

[30] G. W. Gibbons and S. W. Hawking. Cosmological event horizons, thermodynamics, and particle creation. *Phys. Rev. D*, 15:2738–2751, May 1977.

[31] A.S. Goncharov and A.D. Linde. Chaotic inflation in supergravity. *Physics Letters B*, 139(1–2):27 – 30, 1984.

[32] R.H. Gowdy. Gravitational waves in closed universes. *Phys. Rev. Lett.*, 27:826–829, 1971.

[33] David Griffiths. *Introduction to Elementary Particles*, volume 2014. Wiley-VCH, second, revised edition edition.

[34] Alan H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, 23(2):347–356, January 1980.

[35] Alan H. Guth. *The Inflationary Universe*. Helix Books, 1997.

[36] Alan H. Guth and So-Young Pi. Fluctuations in the new inflationary universe. *Phys. Rev. Lett.*, 49:1110–1113, Oct 1982.

[37] J Safco H Goldstein, C Poole. *Classical Mechanics*. Pearsom, third edition, 2002.

[38] J. B. Hartle and S. W. Hawking. Wave function of the universe. *Physical Review D*, 28(12), 1983.

[39] S.W. Hawking. The development of irregularities in a single bubble inflationary universe. *Physics Letters B*, 115(4):295 – 297, 1982.

[40] S.W. Hawking and I.L. Moss. Supercooled phase transitions in the very early universe. *Physics Letters B*, 110(1):35 – 38, 1982.

[41] G. Hinshaw et al. First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: The Angular power spectrum. *Astrophys. J. Suppl.*, 148:135, 2003.

[42] G. Hinshaw and et al. Nine-year wilkinson microwave anisotropy probe (wmap) observations: Cosmological parameter results. 2013.

[43] David W. Hogg. Distance measures in cosmology. 1999.

[44] Wayne Hu. Quadrupole types and polarization patterns, 2001.

[45] Renata Kallosh. More on Universal Superconformal Attractors. *Phys. Rev.*, D89(8):087703, 2014.

[46] Renata Kallosh and Andrei Linde. Superconformal generalization of the chaotic inflation model $\frac{\lambda}{4}\phi^4 - \frac{\xi}{2}\phi^2 R$. *JCAP*, 1306:027, 2013.

[47] Renata Kallosh and Andrei Linde. Superconformal generalizations of the Starobinsky model. *JCAP*, 1306:028, 2013.

[48] Renata Kallosh and Andrei Linde. Escher in the Sky. 2015.

[49] Renata Kallosh, Andrei Linde, and Diederik Roest. Superconformal Inflationary α-Attractors. *JHEP*, 11:198, 2013.

[50] Renata Kallosh, Andrei Linde, and Diederik Roest. Universal Attractor for Inflation at Strong Coupling. *Phys. Rev. Lett.*, 112(1):011303, 2014.

[51] William H. Kinney. Cosmology, inflation, and the physics of nothing. *NATO Sci.Ser.II*, 123:189–243, 2003.

[52] L.A. Kofman, A.D. Linde, and A.A. Starobinsky. Inflationary universe generated by the combined action of a scalar field and gravitational vacuum polarization. *Physics Letters B*, 157(5–6):361 – 367, 1985.

[53] Kuo Kovac. Bicep2 detection of b-mode polarization at degree angular scales. *arXiv:1403.3985v2*.

[54] L. D. Landau and E. M. Lifshitz. *Statistical Physics*. Pergamon Press, Ltd., 2nd impression edition, 1970.

[55] L. D. Landau and E. M. Lifshitz. *Mechanics*. Eslevier, Butterworth, Heinemann, 3rd edition, 1976.

[56] P. Langacker. Grand unification. 7(10):11419, 2012. revision 127918.

[57] Antony Lewis. CAMB Notes. `http://cosmologist.info/notes/CAMB.pdf`.

[58] Antony Lewis. Efficient sampling of fast and slow cosmological parameters. *Phys. Rev.*, D87:103529, 2013.

[59] Antony Lewis and Sarah Bridle. Cosmological parameters from CMB and other data: a Monte- Carlo approach. *Phys. Rev.*, D66:103511, 2002.

[60] Andrew R. Liddle. An Introduction to cosmological inflation. pages 260–295, 1999.

[61] Andrew R. Liddle and David H. Lyth. COBE, gravitational waves, inflation and extended inflation. *Phys. Lett.*, B291:391–398, 1992.

[62] Andrew R. Liddle, Paul Parsons, and John D. Barrow. Formalizing the slow roll approximation in inflation. *Phys.Rev.*, D50:7222–7232, 1994.

[63] James E. Lidsey, Andrew R. Liddle, Edward W. Kolb, Edmund J. Copeland, Tiago Barreiro, et al. Reconstructing the inflation potential : An overview. *Rev.Mod.Phys.*, 69:373–410, 1997.

[64] A. D. Linde. Chaotic inflation in supergravity after planck and bicep2. *arXiv:1405.0270v2*, 2014.

[65] A. D. Linde. Inflationary cosmology after planck 2013. *arXiv:1402.0526v2*, 2014.

[66] A.D. Linde. Scalar field fluctuations in the expanding universe and the new inflationary universe scenario. *Physics Letters B*, 116(5):335 – 339, 1982.

[67] A.D. Linde. Initial conditions for inflation. *Physics Letters B*, 162(4):281 – 286, 1985.

[68] Andrei Linde. Inflationary cosmology and creation of matter in the universe. *Classical and Quantum Gravity*, 18(16):3275, 2001.

[69] Andrei Linde. Current understanding of inflation. *New Astronomy Reviews*, 49(2–6):35 – 41, 2005. Sources and Detection of Dark Matter and Dark Energy in the Universe 6th {UCLA} Symposium on Sources and Detection of Dark Matter and Dark Energy in the Universe.

[70] Andrei Linde. Single-field α-attractors. *JCAP*, 1505(05):003, 2015.

[71] Andrei D. Linde. A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Phys. Lett. B*, 108:389–393, February 1982.

[72] Andrei D. Linde. Scalar field fluctuations in expanding universe and the new inflationary universe scenario. *Phys. Lett.*, B116:335, 1982.

[73] Andrei D. Linde. Chaotic inflation. *Phys.Lett.*, B129:177–181, 1983.

[74] Andrei D. Linde. *Particle Physics and Inflationary Cosmology*. Harwood, Chur, Switzerland, 1990.

[75] Malcom S Longair. *Galaxy Formation*. Springer, second edition, 2007.

[76] Robert E. Lopez and Michael S. Turner. Precision prediction for the big-bang abundance of primordial ^{4}He. *Phys. Rev. D*, 59:103502, Mar 1999.

[77] David H. Lyth and Andrew R. Liddle. *The Primordial Density Perturbation*. Cambridge University Press, 2009.

[78] Charles W. Misner, Kip S. Thorne, and John A. Wheeler. *Gravitation*. W. H. Freeman and Company, 1973.

[79] V. F. Mukhanov and G.C. Chibisov. Quantum fluctuations and a nonsingular universe. *JETP*, 33, 1981.

[80] K. A. Olive and et al. 2014 review of particle physics. Technical report, Particle Data Group pdg.lbl.gov, 2014.

[81] Keith A. Olive. TASI lectures on dark matter. pages 797–851, 2003.

[82] P. J. E. Peebles. Primordial helium abundance and the primordial fireball. ii. *Astrophysical Journal*, 146(8):542–552, 1966.

[83] P. J. E. Peebles. *Principles of Physical Cosmology*. Princeton University Press, 1993.

[84] Dennis V. Perepelista. Sakharov conditions for baryogenesis. Technical report, Columbia University Department of Physics, 2008.

[85] Planck Collaboration, Ade, P. A. R., and et al. Planck 2013 results. i. overview of products and scientific results. *Astronomy and Astrophysics*, 571:A1, 2014.

[86] Planck Collaboration, Ade, P. A. R., and et al. Planck 2013 results. xv. cmb power spectra and likelihood. *Astronomy and Astrophysics*, 571:A15, 2014.

[87] Planck Collaboration, Ade, P. A. R., and et al. Planck 2013 results. xvi. cosmological parameters. *Astronomy and Astrophysics*, 571:A16, 2014.

[88] Planck Collaboration, Ade, P. A. R., and et al. Planck 2015 results. XIII. Cosmological parameters. 2015.

[89] Planck Collaboration, Ade, P. A. R., and et al. Planck 2015 results. XX. Constraints on inflation. 2015.

[90] Cristiano Porciani. *Cosmological Petrubations*. University of Bonn, 2009.

[91] Robertson and Noonan. *Relativity and Geometry*. Saunders Physics Books, 1968.

[92] Diederik Roest and Marco Scalisi. Cosmological attractors from α-scale supergravity. *Phys. Rev.*, D92(4):043525, 2015.

[93] Bernard Shutz. *Geometrical Methods of Mathematical Physics*. Cambridge University Press, 1980.

[94] Christel J. Smith, George M. Fuller, and Michael S. Smith. Big bang nucleosynthesis with independent neutrino distribution functions. *Phys. Rev. D*, 79:105001, May 2009.

[95] A.A. Starobinsky. A new type of isotropic cosmological models without singularity. *Physics Letters B*, 91(1):99 – 102, 1980.

[96] A.A. Starobinsky. Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations. *Physics Letters B*, 117(3):175 – 178, 1982.

[97] Gary Steigman. Neutrinos and big bang nucleosynthesis. Technical Report arXiv:1208.0032v1, High Energy Physics - Phenomenology, 2012.

[98] Paul J. Steinhardt and Neil Turok. Cosmic evolution in a cyclic universe. *Phys. Rev. D*, 65:126003, May 2002.

[99] M. Trodden and S. Carroll. Tasi lectures: Introduction to cosmology. *Cornell University Print Archives: Astrophysics*, (0401547 v1), January 2004.

[100] Robert M. Wald. *General Relativity*. The University of Chicago Press, 1984.

[101] Robert M. Wald. *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*. The University of Chicago Press, 1994.

[102] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[103] Steven Weinberg. *Gravitation and Cosmology*. John Wiley and Sons, 1972.

[104] Steven Weinberg. A new light boson? *Phys. Rev. Lett.*, 40:223–226, Jan 1978.

[105] Steven Weinberg. Baryon- and lepton-nonconserving processes. *Phys. Rev. Lett.*, 43:1566–1570, Nov 1979.

[106] Steven Weinberg. *Cosmology*. Oxford University Press, 2008.

[107] Brian Whitt. Fourth-order gravity as general relativity plus matter. *Physics Letters B*, 145(3–4):176 – 178, 1984.

[108] Frank Wilczek and A. Zee. Operator analysis of nucleon decay. *Phys. Rev. Lett.*, 43:1571–1573, Nov 1979.

Appendix A

Physical Constants, Units, Notation, and Definitions

A.1 Physical Constants and Units.

The cgs system of units is used to calculate values.

Table A.1: Physical Constants from *Mathematica 10*

Constant	**Symbol**	Value	Units
Speed of light	c	2.9979×10^{10}	$cm\ sec^{-1}$
Gravitation	G	6.7384×10^{-8}	$cm^3 sec^{-2}\ g^{-1}$
Reduced Planck	$\hbar$	$1.0545717 \times 10^{-27}$	$erg\ sec$
Boltzmann	k_B	1.38065×10^{-16}	$erg\ K^{-1}$
Electron volt	eV	$1.6021766 \times 10^{-12}$	erg
Electron mass	m_e	9.109383×10^{-25}	g
Proton mass	m_p	1.672622×10^{-21}	g
Solar mass	$\odot$	1.988435×10^{33}	g
Mega parsec	Mpc	3.08568×10^{24}	cm

In the exposition of theory, units are set so that $c = \hbar = 1$. The following definition of the reduced Planck Mass is used in Chapters 3 and 4:

$$M_{Pl} = \sqrt{\frac{1}{8\pi}}\, m_{Pl} \tag{A.1.1}$$

where the Planck mass and Planck energy in units with $c = \hbar = 1$ are written

$$m_{Pl} = \sqrt{\frac{1}{G}}\,. \tag{A.1.2}$$

For example, applying this notation, the Einstein field equation is written

$$R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} = M_{Pl}^{-2}\, T_{\mu\nu}$$

and the Hubble parameter is

$$H = M_{Pl}^{-1} \sqrt{\frac{\rho}{3}}$$

where ρ includes the fictitious energy density of curvature.

Planck length and Planck time are

$$t_{Pl} = m_{Pl} \tag{A.1.3}$$

and

$$\ell_{Pl} = m_{Pl}^{-1}\,. \tag{A.1.4}$$

Therefore Planck density and Planck energy density are

$$m_{Pl}/\ell_{Pl}^3 = m_{Pl}^4 \,. \tag{A.1.5}$$

M_{Pl} is set to unity in the last section of Chapter 4, for consistency with the papers referenced in that section.

A.2 Notation

- Lower case Greek letters generally refer to coordinate indices when the timelike coordinate index is one of the indices.
- Lower case Latin letters generally refer to spatial coordinate indices.
- The cotangent space is called the dual space.
- Forms are called dual vectors.
- Vectors, vector fields, tensor, and tensor fields are all represented by upper case Latin letters (almost always in component form like V^σ) with the notable exception of the metric tensor, $g_{\mu\nu}$. Upper case Latin letters are also used for sets where, again, the context should be clear.
- Dual vectors and dual vector fields are generally represented by lower case Greek letters (almost always in component form like ω_σ). The use of Greek letters for dual vectors (forms) should be apparent from the context. Upper case Latin letters with a downstairs index may also be used from time to time.
- Coordinates are generally represented by x^μ, y^μ or some other lower case Latin letter with some other Greek index.

- ∂_μ is the partial derivative operator with respect to the coordinate x^μ and thus also represents a coordinate (holonomic) basis vector.

- ∇_μ is the covariant derivative along the basis vector e_μ(usually ∂_μ) thus $\nabla_\mu \equiv D_{e_\mu}$. It is not D_u. D_u would be represented as $U^\mu \nabla_{\mu.}$

- $\mathbb{R}$ is the set of real numbers. Members of $\mathbb{R}$ will generally be a, b, c, d and other Latin letters if needed.

- $\mathbb{C}$ is the set of complex numbers. Members of $\mathbb{C}$ will generally be a, b, c, d and other Latin letters if needed.

- $\mathfrak{M}$ is a differentiable manifold and generally identifies the four dimensional spacetime manifold with $(- + ++)$ metric signature since GR and its application to cosmology is the focus of interest rather than differentiable manifolds in general.

- $\mathfrak{F}$ is a set of scalar, real-valued functions or functionals depending on the domain. Members of $\mathfrak{F}$ will generally be f, g.

- The lower case letters ψ, ϕ , their inverses and perhaps some others will be used for maps between manifolds (e.g., $\psi : \mathfrak{M} \to \Re^4$) and in composite maps together with members of $\mathfrak{F}$.

- P, Q represent points in the manifold. T_P is the tangent vector space at point P in the manifold. Likewise T_P^* is dual vector space at point P in the manifold.

- Except as otherwise noted, we assume torsion-free, metric compatible covariant derivatives and use coordinate (holonomic) basis vectors $\{\partial_\mu\}$ and coordinate (holonomic) basis dual vectors $\{dx^\mu\}$.

- $\Gamma^\rho{}_{\mu\nu}$ is the connection coefficient symbol and unless otherwise noted it is the Christoffel symbol uniquely derived from the metric compatible covariant derivative.

This is a slight change from the notation in [21] where the author uses $\Gamma^{\rho}_{\mu\nu}$ to underscore that the symbol is not a tensor. Note that the first downstairs index represents the differentiating index for consistency with [21, 100].

A.3 Definitions

A.3.1 Metric Tensor

The metric tensor, $g_{\mu\nu}$, is a generalization of the dot product of ordinary three dimensional Euclidian space to the four dimensional spacetime manifold. It is a non degenerate symmetric rank $(0, 2)$ tensor:

$$g \equiv g_{\mu\nu}\, dx^{\mu} \otimes dx^{\nu}\,. \tag{A.3.1}$$

Since the metric tensor is non degenerate, its determinant does not vanish and the metric has an inverse:

$$g^{-1} \equiv g^{\mu\nu}\, \partial_{\mu} \otimes \partial_{\nu} \tag{A.3.2}$$

that is a rank $(2, 0)$ tensor.

The metric is often represented as the line element (interval) ds^2 where ds^2 is the sum of sixteen terms below representing an "infinitesimal squared distance". $\{x^{\mu}\}$ represents the coordinate functions and $\{dx^{\mu}\}$ represents basis dual vectors that are the gradients of the coordinate functions. dx^{μ}, dx^{ν} may be informally regarded as infinitesimal coordinate displacements.

The spacetime interval between vectors U and V is $g(U, V)$ and is usually represented as

$$ds^2 = g_{\mu\nu}\, dx^\mu \otimes dx^\nu \quad \text{or} \quad ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu\,.$$

ds^2 many be negative, positive, or zero, depending on whether the interval is timelike, spacelike, or lightlike. A massive object can travel from U to W in spacetime if W is inside the lightcone with the initial position U at the center of the light cone. This is a timelike interval and is negative if we use the metric signature $(-+++)$. For a timelike interval, $-ds^2$ is infinitesimal squared proper time (time measured by a clock moving with the massive object) and is usually given by $d\tau^2$ as

$$d\tau^2 = -g_{\mu\nu}\, dx^\mu \otimes dx^\nu \quad \text{or} \quad d\tau^2 = -g_{\mu\nu}\, dx^\mu\, dx^\nu\,.$$

The metric tensor and its inverse can be used to lower or raise indices on other tensors. The order of the operation is right to left and the raised or lowered index goes into the same slot as the original index. The same letter is frequently used to denote the new tensor but it is not really the same tensor. For example,

$$g^{\mu\gamma} T^{\alpha\beta}{}_{\gamma\delta} = T^{\alpha\beta\mu} \quad \text{and} \quad g_{\mu\alpha} T^{\alpha\beta}{}_{\gamma\delta} = T_\mu{}^\beta{}_\delta\,.$$

A.3.2 Covariant Derivative

A covariant derivative, ∇, is a tensorial extension of an ordinary derivative. The covariant derivative obeys the product rule. It is linear and commutes with tensor contractions. When applied to scalar fields, it is equivalent to an ordinary derivative. When applied to a

tensor of arbitrary rank, it operates as follows:

$$\begin{aligned}\nabla_\sigma T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} &= \partial_\sigma T^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} \\ &+ \Gamma^{\mu_1}{}_{\sigma\lambda} T^{\lambda\ldots\mu_k}{}_{\nu_1\ldots\nu_l} + \Gamma^{\mu_2}{}_{\sigma\lambda} T^{\mu_1\lambda\ldots\mu_k}{}_{\nu_1\ldots\nu_l} + \ldots \\ &- \Gamma^{\lambda}{}_{\sigma\nu_1} T^{\mu_1\ldots\mu_k}{}_{\lambda\ldots\nu_l} - \Gamma^{\lambda}{}_{\sigma\nu_2} T^{\mu_1\ldots\mu_k}{}_{\nu_1\lambda\ldots\nu_l} - \ldots \end{aligned} \tag{A.3.3}$$

[21, Eq. (3.17)] where, for purposes of this paper, $\Gamma^{\lambda}{}_{\mu\nu}$ is the Christoffel symbol connection equal to

$$\Gamma^{\lambda}{}_{\mu\nu} = \frac{1}{2} g^{\lambda\sigma}(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}) \tag{A.3.4}$$

[21, Eq. (3.1)]. From this definition it is evident that the Christoffel symbol is symmetric in its lower indices. The Christoffel symbol with all three indices lowered is

$$\Gamma_{\lambda\mu\nu} = \frac{1}{2}(\partial_\mu g_{\nu\lambda} + \partial_\nu g_{\lambda\mu} - \partial_\lambda g_{\mu\nu}) \tag{A.3.5}$$

[21, Eq. (3.27)] which is symmetric in the second and third lower indices.

A.3.3 Riemann Curvature Tensor

The change in the direction of a vector field during parallel transport around an infinitesimal loop beginning and ending at some point in a manifold, is the commutator of covariant derivatives applied to a vector field:

$$\begin{aligned}[\nabla_\mu, \nabla_\nu]V^\rho &= \nabla_\mu\nabla_\nu V^\rho - \nabla_\nu\nabla_\mu V^\rho \\ &= \partial_\mu(\nabla_\nu V^\rho) - \Gamma^{\lambda}{}_{\mu\nu}\nabla_\lambda V^\rho + \Gamma^{\rho}{}_{\mu\lambda}\nabla_\nu V^\lambda - \partial_\nu(\nabla_\mu V^\rho) + \Gamma^{\lambda}{}_{\nu\mu}\nabla_\lambda V^\rho - \Gamma^{\rho}{}_{\nu\lambda}\nabla_\mu V^\lambda \\ &\ldots \\ &= (\partial_\mu\Gamma^{\rho}{}_{\nu\sigma} - \partial_\nu\Gamma^{\rho}{}_{\mu\sigma} + \Gamma^{\rho}{}_{\mu\lambda}\Gamma^{\lambda}{}_{\nu\sigma} - \Gamma^{\rho}{}_{\nu\lambda}\Gamma^{\lambda}{}_{\mu\sigma})V^\sigma - (\Gamma^{\lambda}{}_{\mu\nu} - \Gamma^{\lambda}{}_{\nu\mu})\nabla_\lambda V^\rho \\ &= R^{\rho}{}_{\sigma\mu\nu}V^\sigma - T^{\lambda}{}_{\mu\nu}\nabla_\lambda V^\rho\end{aligned}$$

[21, Eqs. (3.112) and (3.113)].

The Riemann curvature tensor, $R^{\rho}{}_{\sigma\mu\nu}$, measures the portion of the change in the direction of the vector field during parallel transport around an infinitesimal loop beginning and ending at some point in a manifold that is proportional to vector field, V^{σ}. The torsion tensor, $T^{\lambda}{}_{\mu\nu}$. measures that part of the change that is proportional to the covariant derivative of the vector field, $\nabla_{\lambda}V^{\rho}$. Therefore,

$$R^{\rho}{}_{\sigma\mu\nu} = \partial_{\mu}\Gamma^{\rho}{}_{\nu\sigma} - \partial_{\nu}\Gamma^{\rho}{}_{\mu\sigma} + \Gamma^{\rho}{}_{\mu\lambda}\Gamma^{\lambda}{}_{\nu\sigma} - \Gamma^{\rho}{}_{\nu\lambda}\Gamma^{\lambda}{}_{\mu\sigma} \tag{A.3.6}$$

and

$$T^{\lambda}{}_{\mu\nu} = \Gamma^{\lambda}{}_{\mu\nu} - \Gamma^{\lambda}{}_{\nu\mu}\,.$$

Since for the purposes of this paper, the connection is the symmetric Christoffel symbol, it is evident that the torsion tensor, $T^{\lambda}{}_{\mu\nu} = 0$. Therefore

$$[\nabla_{\mu}, \nabla_{\nu}]V^{\rho} = R^{\rho}{}_{\sigma\mu\nu}V^{\sigma}\,. \tag{A.3.7}$$

A.3.4 Ricci Tensor

The Ricci tensor, $R_{\mu\nu}$, is the contraction of the Riemann curvature tensor on its first and third indices. Therefore,

$$R_{\mu\nu} = R^{\xi}{}_{\mu\xi\nu} = \partial_{\xi}\Gamma^{\xi}{}_{\nu\mu} - \partial_{\nu}\Gamma^{\xi}{}_{\xi\mu} + \Gamma^{\xi}{}_{\xi\lambda}\Gamma^{\lambda}{}_{\nu\mu} - \Gamma^{\xi}{}_{\nu\lambda}\Gamma^{\lambda}{}_{\xi\mu}. \tag{A.3.8}$$

A.3.5 Ricci Scalar

The Ricci scalar, R, is the contraction of the Ricci tensor. Therefore,

$$\begin{aligned} R &= g^{\nu\mu} R_{\mu\nu} \\ &= R^{\nu}{}_{\nu} \,. \end{aligned} \tag{A.3.9}$$

A.3.6 Conformal Transformation of the Ricci Scalar

This description of conformal transformation is based on Appendix D in Robert M. Wald's *General Relativity* [100].

The transformation of a metric

$$\bar{g}_{\mu\nu} = \omega^2 \, g_{\mu\nu} \,, \tag{A.3.10}$$

where ω is a smooth positive function, is called a conformal transformation. It follows that

$$\bar{g}^{\mu\nu} = \omega^{-2} g^{\mu\nu} \,. \tag{A.3.11}$$

The conformal transformation of the Ricci scalar is

$$\bar{R} = \omega^{-2} \left\{ R - 2(n-1) g^{\mu\nu} \nabla_\mu \nabla_\nu \ln \omega - (n-2)(n-1) g^{\mu\nu} \left(\nabla_\mu \ln \omega \right) \left(\nabla_\nu \ln \omega \right) \right\} \tag{A.3.12}$$

[100, Eq. (D.9)].

A.3.7 The Robertson-Walker Metric

The definitions of the Robertson-Walker metric presented here are based on §8.2 of Sean M. Carroll's *Spacetime and Geometry* [21] and §5.1 of Robert M. Wald's *General Relativity* [100]. See the original sources for additional information.

The value of $\mathcal{K}$ sets the constant curvature of the space at a moment in time. It is equal to one-sixth of the spatial Ricci scalar:

$$\mathcal{K} = \frac{^{(3)}R}{6} . \tag{A.3.13}$$

Space is positively curved, flat, or negatively curved depending on whether $\mathcal{K}$ is greater than, equal to, or less than zero. In curved space, the "radius" of curvature is $\left|\mathcal{K}\right|^{-\frac{1}{2}}$. Therefore, in the current era, the radius of curvature may be written

$$R_0 = \left|\mathcal{K}_0\right|^{-\frac{1}{2}} . \tag{A.3.14}$$

The Robertson-Walker metric in the $-+++$ metric for the homogeneous and isotropic Friedmann universe may be written as

$$ds^2 = -dt^2 + a^2(t)R_0^2\left(d\chi^2 + S_k^2(\chi)d\Omega^2\right) \tag{A.3.15}$$

where

$$S_k(\chi) = \begin{cases} \sin\chi & : k = \frac{\mathcal{K}}{|\mathcal{K}|} = 1 \\ \chi & : k = 0 \text{ when } \mathcal{K} = 0 \\ \sinh\chi & : k = \frac{\mathcal{K}}{|\mathcal{K}|} = -1 \end{cases} .$$

The metric in Eq. (A.3.15) may be restated using a transformation of coordinates into

$$ds^2 = -dt^2 + a^2(t)R_0^2\left(\frac{dr^2}{1 - k\,r^2} + r^2 d\Omega^2\right) \tag{A.3.16}$$

where r is a dimensionless radial coordinate defined as

$$r = \begin{cases} \sin(\chi) & : k = \frac{\mathcal{K}}{|\mathcal{K}|} = 1 \\ \chi & : k = 0 \text{ when } \mathcal{K} = 0 \\ \sinh(\chi) & : k = \frac{\mathcal{K}}{|\mathcal{K}|} = -1 \,. \end{cases}$$

In flat space and Cartesian coordinates, the metric is

$$ds^2 = -dt^2 + a^2(t) R_0^2 \left(dx^2 + dy^2 + dz^2 \right) . \tag{A.3.17}$$

Appendix B

Derivations

B.1 Einstein's Equation for Gravitation

The Einstein field equation (or simply the Einstein equation) will be derived from the principle of least action based on the Lagrangian for gravity in empty space,

$$\frac{\mathcal{L}}{\sqrt{-g}} = \frac{1}{2\kappa}R \tag{B.1.1}$$

[100, Eq. (E.12)], and an unspecified Lagrangian corresponding to matter,

$$\mathcal{L}_m\,. \tag{B.1.2}$$

The Einstein-Hilbert action for gravity in empty space is

$$S = \int_{\mathfrak{M}} \sqrt{-g}dx^4 \ \frac{1}{2\kappa}R$$

[100, Eq. (E.13)] where R is the Ricci scalar, κ = is the Einstein constant described below and $\mathfrak{M}$ is the spacetime manifold. The action of empty space and matter is given by

$$S = \int_{\mathfrak{M}} \sqrt{-g} dx^4 \left[\frac{1}{2\kappa} R + \mathcal{L}_m \right].$$

The variation of the action integral of S with respect to the inverse metric vanishes, yielding the following expression:

$$\begin{aligned}
0 &= \delta S \\
&= \delta \int_{\mathfrak{M}} \sqrt{-g} dx^4 \left[\frac{1}{2\kappa} R + \mathcal{L}_m \right] \\
&= \int_{\mathfrak{M}} dx^4 \left[\frac{1}{2\kappa} \delta\left(\sqrt{-g} R \right) + \delta\left(\sqrt{-g} \mathcal{L}_m \right) \right] \\
&= \int_{\mathfrak{M}} dx^4 \left[\frac{1}{2\kappa} \left(R\delta \sqrt{-g} + \sqrt{-g} \delta R \right) + \delta\left(\sqrt{-g} \mathcal{L}_m \right) \right].
\end{aligned}$$

The variation of the Ricci scalar is $\delta R = \delta\left(g^{\mu\nu} R_{\mu\nu}\right) = R_{\mu\nu} \delta g^{\mu\nu} + g^{\mu\nu} \delta R_{\mu\nu}$ and $\delta \sqrt{-g} = -\frac{1}{2} \sqrt{-g}\, g_{\mu\nu} \delta g^{\mu\nu}$ (see Eq. (B.1.10)). Thus the variation of the action equation may be re-stated as:

$$\begin{aligned}
0 &= \delta S \\
&= \int_{\mathfrak{M}} dx^4 \left[\frac{1}{2\kappa} \left(R(-\frac{1}{2} \sqrt{-g}\, g_{\mu\nu} \delta g^{\mu\nu}) + \sqrt{-g} R_{\mu\nu} \delta g^{\mu\nu} + \sqrt{-g}\, g^{\mu\nu} \delta R_{\mu\nu} \right) + \delta\left(\sqrt{-g} \mathcal{L}_m \right) \right] \\
&= \int_{\mathfrak{M}} dx^4 \left[\frac{1}{2\kappa} \left(\sqrt{-g} \delta g^{\mu\nu} R_{\mu\nu} - \sqrt{-g} \delta g^{\mu\nu} \frac{1}{2} R\, g_{\mu\nu} + \sqrt{-g}\, g^{\mu\nu} \delta R_{\mu\nu} \right) + \delta\left(\sqrt{-g} \mathcal{L}_m \right) \right] \\
&= \int_{\mathfrak{M}} dx^4\, \delta g^{\mu\nu}\, \sqrt{-g} \left[\frac{1}{2\kappa} \left(R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} + \frac{g^{\mu\nu} \delta R_{\mu\nu}}{\delta g^{\mu\nu}} \right) + \frac{\delta\left(\sqrt{-g} \mathcal{L}_m \right)}{\sqrt{-g}\, \delta g^{\mu\nu}} \right].
\end{aligned}$$

By Stoke's Theorem the third term above, $\int_{\mathfrak{M}} d^4x\ \sqrt{-g}\, g^{\mu\nu} \delta R_{\mu\nu}$, is a surface term over the boundary of spacetime. Since variations of the inverse metric and its first derivative may be set to zero at infinity (effectively the boundary of spacetime), this term does not contribute

to δS (see Eq. (B.1.15)) and the variation of the action equation may be restated as

$$\int_{\mathfrak{M}} d^4x\, \delta g^{\mu\nu} \sqrt{-g} \left[\frac{1}{2\kappa} \left(R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} \right) + \frac{\delta\left(\mathcal{L}_m \sqrt{-g} \right)}{\sqrt{-g}\, \delta g^{\mu\nu}} \right] = 0. \tag{B.1.3}$$

Since the preceding equation is true for all $\delta g^{\mu\nu}$, the expression in brackets must always vanish. Therefore, it follows that

$$\begin{aligned} R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} &= -2\kappa \left(\frac{\delta\left(\mathcal{L}_m \sqrt{-g} \right)}{\sqrt{-g}\, \delta g^{\mu\nu}} \right) \\ &= \kappa \left(-2 \frac{\delta\left(\mathcal{L}_m \sqrt{-g} \right)}{\sqrt{-g}\, \delta g^{\mu\nu}} \right) \end{aligned} \tag{B.1.4}$$

where the stress-energy tensor for matter or any source of energy, α, is

$$T_{\mu\nu} = -2 \frac{\delta\left(\mathcal{L}_\alpha \sqrt{-g} \right)}{\sqrt{-g}\, \delta g^{\mu\nu}} \tag{B.1.5}$$

[21, Eq. (4.76)]. Although matter was used above as the sole source of stress-energy, it is evident that the stress-energy tensor may be adapted to any general relativity problem by using appropriate Lagrangians, such as $\mathcal{L}_\Lambda$ for dark energy. See, for example, Eq. (B.4.5), for the development of a stress-energy tensor from the energy of a scalar field.

Applying Poisson's equation, $\nabla^2 \Phi = 4\pi G \rho$, the Einstein's constant is determined to have the value $\kappa = 8\pi G$ [21]. Thus, the Einstein field equation takes the following form:

$$R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} = 8\, \pi\, G\, T_{\mu\nu} \tag{B.1.6}$$

which is equivalent to

$$R_{\mu\nu} = 8\, \pi\, G \left(T_{\mu\nu} - \frac{1}{2} T\, g_{\mu\nu} \right) \tag{B.1.7}$$

after contracting both sides to find that $R = -8\pi G T$ and then substituting this result into Eq. (B.1.6).

The Einstein tensor is

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} \,. \tag{B.1.8}$$

leading to a simplified restatement of Eq. (B.1.6) as

$$G_{\mu\nu} = 8\pi G\, T_{\mu\nu} \,. \tag{B.1.9}$$

Supporting derivation: $\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}\, g_{\mu\nu}\delta g^{\mu\nu}$**.** Applying the product rule yields

$$g^{\mu\nu}\delta\, g_{\mu\nu} + \delta\, g^{\mu\nu}\, g_{\mu\nu} = \delta(g^{\mu\nu}\, g_{\mu\nu}) = \delta\left(g^{\mu}{}_{\nu}\right) = \delta\left(\delta^{\mu}{}_{\nu}\right) = 0$$

and

$$g^{\mu\nu}\delta\, g_{\mu\nu} = -\delta\, g^{\mu\nu}\, g_{\mu\nu} \,.$$

The metric, $g_{\mu\nu}$, can be diagonalized by a coordinate transformation to $\tilde{g}$. Therefore, the determinant $g = \tilde{g}_{00} \cdot \tilde{g}_{11} \cdot \tilde{g}_{22} \cdot \tilde{g}_{33}$. The variation of g determined by applying the product rule and the previous result yields

$$\begin{aligned}
\delta g &= \delta(\tilde{g}_{00} \cdot \tilde{g}_{11} \cdot \tilde{g}_{22} \cdot \tilde{g}_{33}) \\
&= \delta\tilde{g}_{00}\left(\tilde{g}_{11} \cdot \tilde{g}_{22} \cdot \tilde{g}_{33}\right) + \delta\tilde{g}_{11}\left(\tilde{g}_{00} \cdot \tilde{g}_{22} \cdot \tilde{g}_{33}\right) + \delta\tilde{g}_{22}\left(\tilde{g}_{00} \cdot \tilde{g}_{11} \cdot \tilde{g}_{33}\right) + \delta\tilde{g}_{33}\left(\tilde{g}_{00} \cdot \tilde{g}_{11} \cdot g_{22}\right) \\
&= \delta\tilde{g}_{00}\tilde{g}^{00} g + \delta\tilde{g}_{11}\tilde{g}^{11} g + \delta\tilde{g}_{22}\tilde{g}^{22} g + \delta\tilde{g}_{33}\tilde{g}^{33} g \\
&= g\left(\tilde{g}^{\sigma\sigma}\delta\tilde{g}_{\sigma\sigma}\right) \\
&= g\left(g^{\mu\nu}\delta g_{\mu\nu}\right) \\
&= -g\left(g_{\mu\nu}\delta g^{\mu\nu}\right) \,.
\end{aligned}$$

The variation of $\delta\ \sqrt{u}$ is

$$\delta\sqrt{u} = \frac{\partial u^{\frac{1}{2}}}{\partial u}\delta u = \frac{1}{2}u^{-\frac{1}{2}}\delta u\,.$$

Substituting $-g$ for u and applying the derivations above yields the following expression:

$$\begin{aligned}
\delta\sqrt{-g} &= \frac{1}{2\sqrt{-g}}\delta(-g) \\
&= -\frac{1}{2\sqrt{-g}}\delta g \\
&= -\frac{1}{2\sqrt{-g}}(-g)\left(g_{\mu\nu}\delta g^{\mu\nu}\right) \\
&= -\frac{1}{2}\sqrt{-g}\,g_{\mu\nu}\delta g^{\mu\nu}\,. \qquad \text{(B.1.10)}
\end{aligned}$$

Supporting derivation: $\int_{\mathfrak{M}} d^4x\ \sqrt{-g}g^{\mu\nu}\delta R_{\mu\nu} = 0$**.** The Riemann curvature tensor in the tensor space over the spacetime manifold (assumed torsion free) is

$$R^{\rho}{}_{\sigma\mu\nu} = \partial_\mu\Gamma^{\rho}{}_{\nu\sigma} - \partial_\nu\Gamma^{\rho}{}_{\mu\sigma} + \Gamma^{\rho}{}_{\mu\lambda}\Gamma^{\lambda}{}_{\nu\sigma} - \Gamma^{\rho}{}_{\nu\lambda}\Gamma^{\lambda}{}_{\mu\sigma}\,.$$

The variation of the Ricci tensor is

$$\begin{aligned}
\delta R_{\mu\nu} &= \delta R^{\rho}{}_{\mu\rho\nu} \\
&= \delta\left(\partial_\rho\Gamma^{\rho}{}_{\nu\mu} - \partial_\nu\Gamma^{\rho}{}_{\rho\mu} + \Gamma^{\rho}{}_{\rho\xi}\Gamma^{\xi}{}_{\nu\mu} - \Gamma^{\rho}{}_{\nu\xi}\Gamma^{\xi}{}_{\rho\mu}\right) \\
&= \partial_\rho\left(\delta\Gamma^{\rho}{}_{\nu\mu}\right) - \partial_\nu\left(\delta\Gamma^{\rho}{}_{\rho\mu}\right) + \Gamma^{\rho}{}_{\rho\xi}\delta\Gamma^{\xi}{}_{\nu\mu} + \Gamma^{\xi}{}_{\nu\mu}\delta\Gamma^{\rho}{}_{\rho\xi} - \Gamma^{\rho}{}_{\nu\xi}\delta\Gamma^{\xi}{}_{\rho\mu} - \Gamma^{\xi}{}_{\rho\mu}\delta\Gamma^{\rho}{}_{\nu\xi}\,.
\end{aligned} \qquad \text{(B.1.11)}$$

The variation of a Christoffel symbol is a tensor as seen below. The covariant derivative of the metric tensor is

$$\nabla_\lambda\, g_{\mu\nu} = \partial_\lambda\, g_{\mu\nu} - \Gamma^{\xi}{}_{\lambda\mu}\, g_{\xi\nu} - \Gamma^{\xi}{}_{\lambda\nu}\, g_{\mu\xi}\,.$$

Variation of this expression yields the following result:

$$0 = \delta(\nabla_\lambda g_{\mu\nu}) = \underbrace{\partial_\lambda (\delta g)_{\mu\nu} - \Gamma^\xi{}_{\lambda\mu}(\delta g)_{\xi\nu} - \Gamma^\xi{}_{\lambda\nu}(\delta g)_{\mu\xi}}_{\nabla_\lambda (\delta g)_{\mu\nu}} - g_{\xi\nu}\delta\Gamma^\xi{}_{\lambda\mu} - g_{\mu\xi}\delta\Gamma^\xi{}_{\lambda\nu}$$
$$= \nabla_\lambda (\delta g)_{\mu\nu} - g_{\xi\nu}\delta\Gamma^\xi{}_{\lambda\mu} - g_{\mu\xi}\delta\Gamma^\xi{}_{\lambda\nu} \,.$$

Cycling the indices yields the following three equations:

$$\begin{aligned}
\nabla_\lambda (\delta g)_{\mu\nu} &= g_{\xi\nu}\delta\Gamma^\xi{}_{\lambda\mu} + g_{\mu\xi}\delta\Gamma^\xi{}_{\lambda\nu} \,, \\
\nabla_\mu (\delta g)_{\nu\lambda} &= g_{\xi\lambda}\delta\Gamma^\xi{}_{\mu\nu} + g_{\nu\xi}\delta\Gamma^\xi{}_{\mu\lambda} \,, \text{ and} \\
\nabla_\nu (\delta g)_{\lambda\mu} &= g_{\xi\mu}\delta\Gamma^\xi{}_{\nu\lambda} + g_{\lambda\xi}\delta\Gamma^\xi{}_{\nu\mu} \,.
\end{aligned}$$

Adding the first two and subtracting the third yields the following expression after grouping like terms:

$$\begin{aligned}
\nabla_\lambda (\delta g)_{\mu\nu} + \nabla_\mu (\delta g)_{\nu\lambda} - \nabla_\nu (\delta g)_{\lambda\mu} &= g_{\mu\xi}\delta\Gamma^\xi{}_{\lambda\nu} - g_{\xi\mu}\delta\Gamma^\xi{}_{\nu\lambda} \\
&\quad + g_{\xi\lambda}\delta\Gamma^\xi{}_{\mu\nu} - g_{\lambda\xi}\delta\Gamma^\xi{}_{\nu\mu} \\
&\quad + g_{\nu\xi}\delta\Gamma^\xi{}_{\mu\lambda} + g_{\xi\nu}\delta\Gamma^\xi{}_{\lambda\mu} \,.
\end{aligned}$$

Using the symmetry of the metric tensor and the second form of the Christoffel symbol and multiplying through by $\frac{1}{2}g^{\xi\nu}$yields the following expression:

$$\delta\Gamma^\xi{}_{\lambda\mu} = \frac{1}{2}g^{\xi\nu}\left(\nabla_\nu (\delta g)_{\lambda\mu} - \nabla_\lambda (\delta g)_{\mu\nu} - \nabla_\mu (\delta g)_{\nu\lambda}\right) .$$

Since the term on the right side is a rank $(1,2)$ tensor, $T^\xi{}_{\lambda\mu}$, it follows that the variation of a Christoffel connection is a rank $(1,2)$ tensor. The covariant derivative of $\delta\Gamma^\rho{}_{\nu\mu}$ is

$$\nabla_\rho\left(\delta\Gamma^\rho{}_{\nu\mu}\right) = \partial_\rho\left(\delta\Gamma^\rho{}_{\nu\mu}\right) + \Gamma^\rho{}_{\rho\xi}\delta\Gamma^\xi{}_{\nu\mu} - \Gamma^\xi{}_{\rho\nu}\delta\Gamma^\rho{}_{\xi\mu} - \Gamma^\xi{}_{\rho\mu}\delta\Gamma^\rho{}_{\nu\xi} \,. \tag{B.1.12}$$

Similarly,

$$\nabla_\nu \left(\delta\Gamma^\rho{}_{\rho\mu}\right) = \partial_\nu \left(\delta\Gamma^\rho{}_{\rho\mu}\right) + \Gamma^\rho{}_{\nu\xi}\delta\Gamma^\xi{}_{\rho\mu} - \Gamma^\xi{}_{\nu\rho}\delta\Gamma^\rho{}_{\xi\mu} - \Gamma^\xi{}_{\nu\mu}\delta\Gamma^\rho{}_{\rho\xi}\,. \tag{B.1.13}$$

The difference between these covariant derivatives is the variation of the Ricci tensor in Eq. (B.1.11):

$$\begin{aligned}
\nabla_\rho \left(\delta\Gamma^\rho{}_{\nu\mu}\right) - \nabla_\nu \left(\delta\Gamma^\rho{}_{\rho\mu}\right) &= \partial_\rho \left(\delta\Gamma^\rho{}_{\nu\mu}\right) + \Gamma^\rho{}_{\rho\xi}\delta\Gamma^\xi{}_{\nu\mu} - \Gamma^\xi{}_{\rho\nu}\delta\Gamma^\rho{}_{\xi\mu} - \Gamma^\xi{}_{\rho\mu}\delta\Gamma^\rho{}_{\nu\xi} \\
&\quad - \left(\partial_\nu \left(\delta\Gamma^\rho{}_{\rho\mu}\right) + \Gamma^\rho{}_{\nu\xi}\delta\Gamma^\xi{}_{\rho\mu} - \Gamma^\xi{}_{\nu\rho}\delta\Gamma^\rho{}_{\xi\mu} - \Gamma^\xi{}_{\nu\mu}\delta\Gamma^\rho{}_{\rho\xi}\right) \\
&= \partial_\rho \left(\delta\Gamma^\rho{}_{\nu\mu}\right) - \partial_\nu \left(\delta\Gamma^\rho{}_{\rho\mu}\right) \\
&\quad + \Gamma^\rho{}_{\rho\xi}\delta\Gamma^\xi{}_{\nu\mu} + \Gamma^\xi{}_{\nu\mu}\delta\Gamma^\rho{}_{\rho\xi} - \Gamma^\rho{}_{\nu\xi}\delta\Gamma^\xi{}_{\rho\mu} - \Gamma^\xi{}_{\rho\mu}\delta\Gamma^\rho{}_{\nu\xi} \\
&= \delta R_{\mu\nu}\,.
\end{aligned} \tag{B.1.14}$$

$\int_{\mathfrak{M}} d^4x\ \sqrt{-g}g^{\mu\nu}\delta R_{\mu\nu}$ may be restated as follows using the results above and metric compatibility (i.e., $\nabla_\rho\, g^{\mu\nu} = 0$):

$$\begin{aligned}
\int_{\mathfrak{M}} d^4x\ \sqrt{-g}\, g^{\mu\nu}\, \delta R_{\mu\nu} &= \int_{\mathfrak{M}} d^4x\ \sqrt{-g}\, g^{\mu\nu} \left(\nabla_\rho \left(\delta\Gamma^\rho{}_{\nu\mu}\right) - \nabla_\nu \left(\delta\Gamma^\rho{}_{\rho\mu}\right)\right) \\
&= \int_{\mathfrak{M}} d^4x\ \sqrt{-g} \left(\nabla_\rho \left(\, g^{\mu\nu}\delta\Gamma^\rho{}_{\nu\mu}\right) - \nabla_\nu \left(\, g^{\mu\nu}\delta\Gamma^\rho{}_{\rho\mu}\right)\right) \\
&= \int_{\mathfrak{M}} d^4x\ \sqrt{-g} \left(\nabla_\xi \left(\, g^{\mu\nu}\delta\Gamma^\xi{}_{\nu\mu}\right) - \nabla_\xi \left(\, g^{\mu\xi}\delta\Gamma^\rho{}_{\rho\mu}\right)\right) \\
&= \int_{\mathfrak{M}} d^4x\ \sqrt{-g}\, \nabla_\xi \left(\underbrace{g^{\mu\nu}\delta\Gamma^\xi{}_{\nu\mu} - \, g^{\mu\xi}\delta\Gamma^\rho{}_{\rho\mu}}_{V^\xi}\right) \\
&= \int_{\mathfrak{M}} d^4x\ \sqrt{-g}\, \nabla_\xi\, V^\xi\,.
\end{aligned}$$

Note that the sum of the two terms above the under-brace contracts into a vector field and that $d^4x\ \sqrt{-g}\, \nabla_\xi\, V^\xi$ is the natural volume element of the covariant divergence of a vector

field. Stoke's Theorem generalized to differential geometry the may be stated as:

$$\int_{\partial M} \omega = \int_{M} d\omega$$

where M is region in an n-dimensional manifold (or the entire manifold), ω is an $(n-1)-$form and $d\omega$ is an $n-$form. Thus, using the entire four dimensional spacetime manifold and $d^4x\ \sqrt{-g}\,\nabla_\xi V^\xi$ as the 4−form, by Stoke's Theorem [21, Eq E.14]:

$$\int_{\mathfrak{M}} d^4x\ \sqrt{-g}\,\nabla_\xi V^\xi = \int_{\partial\mathfrak{M}} d^3x\ \sqrt{-\gamma}\,V^\xi$$

where γ_{ij} is the metric of the spacelike hypersurface at the boundary of the spacetime manifold, $\sqrt{|\gamma|}\,d^3x$ is the surface element of the hypersurface of the boundary and n^μ is the unit normal vector to the boundary. Variations of $g^{\mu\nu}$ and its first derivative may be set to zero at infinity (effectively the boundary of spacetime). Therefore, the boundary term vanishes, and

$$\begin{aligned}\int_{\mathfrak{M}} d^4x\ \sqrt{-g}g^{\mu\nu}\delta R_{\mu\nu} &= \int_{\partial\mathfrak{M}} d^3x\ \sqrt{-\gamma}\,V^\xi \\ &= 0\,. \end{aligned} \tag{B.1.15}$$

B.2 Local Energy-Momentum Conservation in Spacetime

Overview. In classical field theory, time is absolute and fields evolve in static space. The theory of special relativity replaced the axiom of absolute or invariant time with two postulates: (a) the invariant speed of light in all inertial reference frames and (b) the invariance of spacetime intervals on transformations between inertial reference frames. This is called Lorentz invariance. Quantum field theory may be described as relativistic quantum mechanics consistent with special relativity and Minkowski spacetime. The operators and expressions in quantum field theory are required to be Lorentz invariant. Space is still fixed

even though time is not absolute. If space does not evolve with time, then time-translation invariance holds and energy is the corresponding conserved quantity.

In general relativity, spacetime is dynamic. Spacetime interacts with matter and radiation fields. To paraphrase John Wheeler, when particles move through spacetime they change spacetime and spacetime defines how the particles move.

A dramatic example is the theoretical expansion of space during inflation that increased total energy from about a milligram to more than 10^{57} grams in the observable universe alone (see Eq. §4.2). A less extreme example but one that is widely accepted and well supported by observations is based on the accelerated expansion of space driven by the constant (or near constant) energy density of dark energy. Total energy in the universe is increasing constantly since dark energy increases in proportion to the expanding volume of space.

The CMB provides a simple example of energy lost due to the dynamics of space. Photons decoupled from matter when the average temperature of the CMB was more than one thousand times greater than it is today (see §2.4). Therefore, the energy of each CMB photon observed today has lost more than 99.9% of its energy since decoupling. Yet the CMB photons that are measured today had no interaction with other particles during the intervening 13.4 billion years. The expansion of space stretched the wavelengths of CMB photons, decreasing their frequencies, resulting in decreasing energy: $E = h\nu$.

A commonly held counter argument to the loss of energy conservation in general relativity is based on the energy of the classical "gravitational field". By setting gravitational potential to zero at infinity, the energy of the gravitational field is negative throughout space. Therefore, as space expands, the total negative energy of the gravitational field increases, offsetting the increases (or decreases) in the energy of other fields. Based on this argument, during inflation, the energy of the *inflaton field* increased enormously and was offset by the corresponding increase in the negative energy of the gravitational field. During re-heating, energy was transferred from the *inflaton field* to matter and radiation fields with or without

expansion and gravitation field offsets. After the beginning of the radiation era and through to the current era, the increase in the energy of dark energy and the decrease in the energy of radiation fields due to redshift are offset by corresponding increases in the negative energy of the gravitational field.

However, energy in general relativity is based on energy density. The gravitational "field" has no energy density that allows it to be introduced in this manner into the theory. Since modern cosmology is based largely on the highly successful theory of general relativity, most cosmologists take the view that spacetime can add or remove energy from matter fields and that, therefore, energy is not conserved [21].

In general relativity, the stress-energy tensor is *locally* conserved at every point in the curved spacetime manifold, matching the conservation of the Einstein tensor in Eq. (B.1.8). The evolution of the energy and momentum of particles and fields in tandem with the evolution of spacetime is precisely defined by Einstein's equation without requiring conservation of energy.

Perhaps advances in quantum theories of gravity and experiments will settle the issue by identifying the scalar fields thought to be the source of early inflation and expansion driven by dark energy. The remainder of this section describes the local conservation of the stress-energy tensor in curved spacetime. For additional information see Chapters 4, 5, and 20 in *Gravitation* [78].

Additional information on the incompatibility of the gravitational field with general relativity is provided in *Graviational Field Pseudo Tensor* below.

Special Relativity and the Poincaré Group. The spacetime manifold of special relativity is called the Minkowski spacetime. Spacetime is flat and static. It is neither expanding nor contracting.

An "inertial reference frame" or coordinate system for the manifold may be constructed by selecting an arbitrary point as the origin and establishing a spatial coordinate system

with arbitrary orientation of the three mutually orthogonal axes. Ideal clocks are placed at all locations in space including the origin. The clocks are synchronized so that the clock at position $\mathbf{x}'$ is set to $t' = \frac{1}{2}(t + t_r)$ where t is the time at the origin when a light signal is sent to $\mathbf{x}'$ and t_r is the time the signal is received back from $\mathbf{x}'$ assuming that it was returned immediately upon receipt. The inertial coordinate system covers the entire Minkowski spacetime manifold.

Spacetime intervals are based on the covariant metric tensor, represented in line element form as

$$\begin{aligned} ds^2 &= \eta_{\mu\nu} dx^\mu \, dx^\nu \\ &= -dt^2 + dx^2 + dy^2 + dz^2 \, . \end{aligned} \tag{B.2.1}$$

The complete set of transformations between inertial reference frames is defined by the Poincaré group of transformation [37]. The Poincaré group of transformations includes the following ten elements consisting of four translations, three spatial rotations, and three "boosts":

- Four translations are defined by $x^{\mu'} = x^\mu + a^\mu$ where a^μ is a constant displacement in time and/or space.

- Spatial rotations are defined by matrices, $\Lambda^{\mu'}{}_{\nu}$, from the unprimed to the primed frame, and $\Lambda^{\nu}{}_{\mu'}$, from the primed to the unprimed frame. For example, a spatial rotation in the x, y plane is:

$$\Lambda^{\mu'}{}_{\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \Lambda^{\nu}{}_{\mu'} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{B.2.2}$$

parameterized by θ. It is evident that $\Lambda^{\mu'}{}_{\nu}\Lambda^{\nu}{}_{\mu'} = I$. There are three spatial rotations: one for each plane.

- Rotations between time and a spatial coordinate are called "boosts". They are also defined by matrices, $\Lambda^{\mu'}{}_{\nu}$, from the unprimed to the primed frame, and $\Lambda^{\nu}{}_{\mu'}$, from the primed to the unprimed frame. The boost in the x direction from the unprimed to the primed frame is

$$\Lambda^{\mu'}{}_{\nu} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \equiv \begin{pmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{B.2.3}$$

The boost in the x direction from the primed to the unprimed frame is

$$\Lambda^{\nu}{}_{\mu'} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \equiv \begin{pmatrix} \cosh\phi & \sinh\phi & 0 & 0 \\ \sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{B.2.4}$$

The boost above (and its inverse) is parameterized by the relative velocity in the x direction between the frames, v, where $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ and $\beta = \frac{v}{c}$. There are three boosts: one for each spatial coordinate. As in the case of spatial rotations, $\Lambda^{\mu'}{}_{\nu}\Lambda^{\nu}{}_{\mu'} = I$.

Transformations may be performed in any sequence with all ten free parameters: $a_t, a_x, a_y, a_z, \theta_{xy}, \theta_{yz}, \theta_{zx}, v_x, v_y,$ and v_z. Inertial reference frames produced by transformations in the Poincaré group of transformations are called Lorentz frames. In special relativity, a Lorentz frame covers the entire manifold. Vectors in special relativity are vectors in the spacetime manifold.

The spatial rotations and boosts form the Lorentz group that is a subgroup of the Poincaré group. A general transformation between inertial reference frames formed from the elements of the Poincaré group may be written

$$\mathbf{x}' = \mathbf{\Lambda x} + \mathbf{a} \tag{B.2.5}$$

where the $\mathbf{\Lambda}$ is any product of the spatial rotations and boosts and $\mathbf{a}$ is a constant four-vector of displacement formed as the sum of any number of separate constant four-vector displacements.

Transformations from the Lorentz group can be applied to basis vectors, basis dual vectors, vector components, and dual vector components:

$$\begin{aligned}
\partial_{\mu'} &= \partial_\nu \Lambda^\nu{}_{\mu'}, & \partial_\nu &= \partial_{\mu'} \Lambda^{\mu'}{}_\nu \\
dx^{\mu'} &= \Lambda^{\mu'}{}_\nu \, dx^\nu, & dx^\nu &= \Lambda^\nu{}_{\mu'} \, dx^{\mu'} \\
U^{\mu'} &= \Lambda^{\mu'}{}_\nu \, U^\nu, & U^\nu &= \Lambda^\nu{}_{\mu'} \, U^{\mu'} \\
\sigma_{\mu'} &= \sigma_\nu \Lambda^\nu{}_{\mu'}, & \sigma_\nu &= \sigma_{\mu'} \, \Lambda^{\mu'}{}_\nu \, .
\end{aligned}$$

Therefore, the representation of vectors, dual vectors, and, by extension, tensors are independent of the frame of reference. For example, since $\Lambda^{\mu'}{}_\nu \Lambda^\nu{}_{\mu'} = I$,

$$\mathbf{A} = \partial_{\mu'} A^{\mu'} = \partial_{\mu'} \Lambda^{\mu'}{}_\nu \Lambda^\nu{}_{\mu'} A^{\mu'} = \partial_\nu A^\nu \, . \tag{B.2.6}$$

Therefore, the magnitude of an arbitrary vector, $\mathbf{A}$, defined by the metric as the square root of

$$\eta(\mathbf{A}, \mathbf{A}) \equiv \mathbf{A} \cdot \mathbf{A} = \mathbf{A}^2 \tag{B.2.7}$$

is invariant upon transformation between inertial reference frames formed from the elements of the Lorentz group.

For example, if $\mathbf{A}$ represents the four-vector between two events in spacetime, $\mathbf{A} = \mathbf{x_1} - \mathbf{x_2}$, then spacetime interval between two events, $|A| = |\mathbf{x_1} - \mathbf{x_2}|$ is invariant on these transformations. The interval between events in a reference frame transformed by translation is

$$x_1^{\mu'} - x_1^{\mu'} = x_1^{\mu} + a^{\mu} - x_2^{\mu} - a^{\mu} = x_1^{\mu} - x_2^{\mu} \,. \tag{B.2.8}$$

Therefore, the spacetime interval between two events is invariant between the two inertial reference frames based on any transformation formed from the Poincaré group as shown in Eq.(B.2.5). These are the transformations that satisfy the second postulate of special relativity.

The invariance of time translation, the three spatial translations, and the three spatial rotations are symmetries with the conserved quantities: energy, momentum and angular momentum. Boosts in a Lorentz frame are the relativistic counterparts to the Galilei transformation, $\mathbf{x}' = \mathbf{x} + \mathbf{v}t$. The conserved quantity of a boost symmetry is the relativistic center of mass [37]. The Poincaré group is the group of transformations consistent with the symmetries of Minkowski spacetime [11].

Therefore, the four-momentum, representing energy density and the components of the momentum, is conserved in the flat Minkowski spacetime of special relativity.

The Stress-Energy Tensor. The worldline (the evolving position in spacetime) of a particle as a function of the particle's proper time is

$$\mathscr{P}(\tau) = x^{\mu}(\tau)\partial_{\mu} \,. \tag{B.2.9}$$

The four-velocity vector of the particle is

$$\begin{aligned}\frac{\partial \mathscr{P}(\tau)}{d\tau} &= \frac{\partial x^{\mu}(\tau)}{\partial \tau}\partial_{\mu} \\ &= U^{\mu}\partial_{\mu}\end{aligned} \tag{B.2.10}$$

where $U^0 = \gamma$ and $U^i = \gamma V_i$. The term

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} = \frac{1}{\sqrt{1 - v^2}}\,,$$

and V^{μ} is the ordinary velocity vector. The four-momentum vector of the particle is

$$P^{\mu} = m\,U^{\mu} \tag{B.2.11}$$

where m is the mass of the particle.

The components of the stress-energy tensor, $T^{\mu\nu}$, may be described as follows [78] where P^{μ} represents a component of four-momentum, F^k represents a component of force, and ∂_k is a basis vector:

$$T^{\mu\nu} = \begin{cases} T^{00} & \text{Energy density: } P^0\text{, per unit volume} \\ T^{0k} & \text{Energy flux: } P^0 \text{ transfer per unit of time (power) through a unit surface } \perp \text{ to } \partial_k \\ T^{j0} & \text{Momentum density: } P^j\text{, per unit volume} \\ T^{kk} & \text{Pressure: } P^k \text{ transfer per unit of time } (F^k) \text{ through a unit surface } \perp \text{ to } \partial_k \\ T^{jk} & \text{Stress : } P^j \text{ transfer per unit of time } (F^j) \text{ through a unit surface } \perp \text{ to } \partial_k\,. \end{cases}$$

Global Conservation of the Stress-Energy Tensor in Minkowski Spacetime. Consider a closed region of Minkowski spacetime, V, enclosed by the three dimensional spatial

surface ∂V. From the definitions above it follows that

$$\Delta E = \oint_{\partial V} T^{00} d^3x \tag{B.2.12}$$

is the net change in energy transferred through the surface and that

$$\Delta P^j = \oint_{\partial V} T^{j0} d^3x \tag{B.2.13}$$

is the net change in j-momentum transferred through the surface. By conservation of energy and momentum, any energy or momentum transferred into the region must leave the region somewhere. Therefore, $\Delta E = \Delta P^j = 0$ [78].

More generally, consider a closed region of Minkowski spacetime, V, enclosed by the three dimensional surface ∂V where the three-dimension surface omits either the time dimension or one of the three spatial dimensions; $\mu = 0, 1, 2, \text{or } 3$. Particles and fields flow in and out of the volume through the surface, carrying energy and momentum with them. Since energy-momentum is conserved in the global Lorentz frame, the net flow of energy-momentum through the surface must be zero:

$$\oint_{\partial V} T^{\mu\nu} \Sigma_\mu = 0$$

for every $\nu = 0, 1, 2, 3$ and $\Sigma_\mu = d^3x$ omitting dx^μ.

By the integral form of Gauss' Theorem [78] expressed in four dimensional space,

$$\oint_{\partial V} T^{\mu\nu} d^3\Sigma_\mu = \int_V \partial_\mu T^{\mu\nu} d^4\, x = 0\,. \tag{B.2.14}$$

In order for the middle term to vanish for every four-dimensional volume, the integrand must vanish. Therefore,

$$\partial_\mu T^{\mu\nu} = 0 \tag{B.2.15}$$

in Minkowski spacetime.

Gravitational Field Pseudo Tensor. At any event in spacetime, a coordinate system may be selected and the action of the gravitational field within that frame may be defined based on a pseudo tensor, $t^{\mu\nu}$. It is called a pseudo tensor because it is not independent of the coordinate system. The energy tensor of the system excluding gravity, $T^{\mu\nu}$, is determined within the same coordinate system and is added to $t^{\mu\nu}$ to form an effective pseudo stress-energy tensor that includes the gravitational field:

$$T^{\mu\nu}_{eff} = T^{\mu\nu} + t^{\mu\nu}\,. \tag{B.2.16}$$

Based on the coordinate system and $T^{\mu\nu}_{eff}$, the integral form of Gauss' Theorem expressed in four dimensional space in Eq. (B.2.14) holds:

$$\oint_{\partial V} T^{\mu\nu} d^3\Sigma_\mu = \int_V \partial_\mu T^{\mu\nu}_{eff} d^4\,x = 0\,, \tag{B.2.17}$$

even if the volume of spacetime is increasing. Therefore, the integrand must vanish:

$$\partial_\mu T^{\mu\nu}_{eff} = 0 \tag{B.2.18}$$

establishing the global conservation of the pseudo stress energy tensor based on the ordinary derivative defined by the selected coordinate system.

There is, however, no well-defined local action of gravity. This is evident since it is possible to pick another reference frame in which $t^{\mu\nu}$ constructed in the original frame

vanishes, i.e., the frame of a free-falling object. If a change in frame causes a force to vanish, the local action of gravity is not well defined. The predictive strength of general relativity is derived from the theory that the action of gravity is described by the dynamic geometry of spacetime that is precisely defined locally (everywhere and at each moment) by co-evolution of the distribution of matter and energy in all of its forms. For additional information see §20.3 in Ref. [78].

The Minimum Coupling Principle. The Minimum Coupling Principle in general relativity states that in small enough regions of curved spacetime, the laws of physics reduce to those of special relativity. Broadly speaking, this means that any statement of physics that holds in the global Lorentz frame of special relativity also holds in the local Lorentz frame at each point in curved spacetime [21, p. 152]. The local Lorentz frame is established within a flat tangent vector space and dual vector space at each point in curved spacetime, not within the curved spacetime manifold itself. The local inertial coordinate system covers an infinitesimal region about the point in the curved spacetime manifold.

Conservation of the Stress-Energy Tensor in Curved Spacetime. Applying one or more coordinate transformations in the Poincaré group between the local Lorentz frames for points that lie within the intersection of the frames permits the entire curved spacetime manifold to be stitched together from local Lorentz frames. Therefore, if an expression of physics is written in coordinate-invariant form (i.e., tensorial form), then the statement of physics is true throughout curved spacetime. This includes replacing ordinary derivatives with covariant derivatives everywhere in the expression. The tensor expressions of physics are globally valid but the outcomes are local.

The law of *local* energy-momentum conservation in curved spacetime is written

$$\nabla_\mu T^{\mu\nu} = 0 \tag{B.2.19}$$

where $T^{\mu\nu}$ is the contravariant form of the stress-energy tensor and ∇ represents the covariant derivative defined in §A.3.2.

B.3 Evolution of the Copernican Spacetime Manifold

Overview. In this section, the time evolution of Copernican spacetime will be explored based on general relativity. This section introduces the energy constituents of the universe and the relationship between energy density, pressure, and the expansion of space. The Einstein equation, the Robertson-Walker metric, and a model of the universe as a perfect fluid will be used as the basis for the analysis.

B.3.1 Preliminary Derivations

The standard covariant form of the Einstein equation is

$$R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} = 8\pi G T_{\mu\nu}$$

and the trace reversed covariant form is

$$R_{\mu\nu} = 8\pi G \left(T_{\mu\nu} - \frac{1}{2} T\, g_{\mu\nu} \right).$$

See §B.1 for a derivation of these equations from the principle of least action.

The reduced circumference polar form of the Robertson-Walker metric (Eq. (A.3.16)) and its inverse in matrix formats are

$$g = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & \frac{a^2 R_0^2}{1 - k r^2} & & 0 \\ 0 & 0 & a^2 R_0^2 r^2 & 0 \\ 0 & 0 & 0 & a^2 R_0^2 r^2 \sin^2\theta \end{pmatrix} \text{ and } \tag{B.3.1}$$

$$g^{-1} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & \frac{1-k\,r^2}{a^2 R_0^2} & & 0 \\ 0 & 0 & \frac{1}{a^2 R_0^2\, r^2} & 0 \\ 0 & 0 & 0 & \frac{1}{a^2 R_0^2\, r^2 \sin^2\theta} \end{pmatrix}. \tag{B.3.2}$$

The Christoffel connection expressed in terms of the metric (Eq. (A.3.4)) is

$$\Gamma^{\lambda}{}_{\mu\nu} = \frac{1}{2} g^{\lambda\sigma}(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu})\,. \tag{B.3.3}$$

Applying this equation to the metric and its inverse, the Christoffel connections may be represented in matrix form as:

$$\Gamma^0_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{a\dot{a}R_0^2}{1-kr^2} & 0 & 0 \\ 0 & 0 & r^2 a\dot{a}R_0^2 & 0 \\ 0 & 0 & 0 & r^2 a \sin^2(\theta)\dot{a}R_0^2 \end{pmatrix}, \tag{B.3.4}$$

$$\Gamma^1_{\mu\nu} = \begin{pmatrix} 0 & \frac{\dot{a}}{a} & 0 & 0 \\ \frac{\dot{a}}{a} & \frac{kr}{1-kr^2} & 0 & 0 \\ 0 & 0 & -r\left(1-kr^2\right) & 0 \\ 0 & 0 & 0 & -r\left(1-kr^2\right)\sin^2(\theta) \end{pmatrix}, \tag{B.3.5}$$

$$\Gamma^2_{\mu\nu} = \begin{pmatrix} 0 & 0 & \frac{\dot{a}}{a} & 0 \\ 0 & 0 & \frac{1}{r} & 0 \\ \frac{\dot{a}}{a} & \frac{1}{r} & 0 & 0 \\ 0 & 0 & 0 & -\cos(\theta)\sin(\theta) \end{pmatrix}, \text{ and} \tag{B.3.6}$$

$$\Gamma^3_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & \frac{\dot{a}}{a} \\ 0 & 0 & 0 & \frac{1}{r} \\ 0 & 0 & 0 & \cot(\theta) \\ \frac{\dot{a}}{a} & \frac{1}{r} & \cot(\theta) & 0 \end{pmatrix}. \tag{B.3.7}$$

Per Eq. (A.3.6) the Riemann curvature tensor may be expressed in terms of the Christoffel connections as:

$$R^{\rho}{}_{\sigma\mu\nu} = \partial_\mu \Gamma^{\rho}{}_{\nu\sigma} - \partial_\nu \Gamma^{\rho}{}_{\mu\sigma} + \Gamma^{\rho}{}_{\mu\lambda}\Gamma^{\lambda}{}_{\nu\sigma} - \Gamma^{\rho}{}_{\nu\lambda}\Gamma^{\lambda}{}_{\mu\sigma}\,.$$

The Ricci tensor, $R_{\mu\nu}$, is the contraction of the Riemann curvature tensor on the first and third indices. Therefore,

$$\begin{aligned} R_{\mu\nu} &= R^{\xi}{}_{\mu\xi\nu} \\ &= \partial_\xi \Gamma^{\xi}{}_{\nu\mu} - \partial_\nu \Gamma^{\xi}{}_{\xi\mu} + \Gamma^{\xi}{}_{\xi\lambda}\Gamma^{\lambda}{}_{\nu\mu} - \Gamma^{\xi}{}_{\nu\lambda}\Gamma^{\lambda}{}_{\xi\mu}\,. \end{aligned} \tag{B.3.8}$$

Using this equation and the Christoffel connections in the matrices above, the matrix representation of the Ricci tensor is

$$R_{\mu\nu} = \begin{pmatrix} -\frac{3\ddot{a}}{a} & 0 & 0 & 0 \\ 0 & \frac{2\left(\dot{a}^2R_0^2+k\right)+a\ddot{a}R_0^2}{1-kr^2} & 0 & 0 \\ 0 & 0 & r^2\left(2\left(\dot{a}^2R_0^2+k\right)+a\ddot{a}R_0^2\right) & 0 \\ 0 & 0 & 0 & r^2\sin^2(\theta)\left(2\left(\dot{a}^2R_0^2+k\right)+a\ddot{a}R_0^2\right) \end{pmatrix}. \tag{B.3.9}$$

The Ricci scalar is the trace of the Ricci tensor with one raised index. Therefore the Ricci scalar is:

$$\begin{aligned} R &= R^{\nu}{}_{\nu} \\ &= g^{\nu\mu} R_{\mu\nu} \\ &= \frac{6\left[\left(a\ddot{a} + \dot{a}^2\right) R_0^2 + k\right]}{a^2 R_0^2} \\ &= 6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2 R_0^2}\right]. \end{aligned} \tag{B.3.10}$$

The stress energy tensor of a perfect fluid in arbitrary spacetime is

$$T_{\mu\nu} = (\rho + p)\, U_\mu U_\nu + p g_{\mu\nu} \tag{B.3.11}$$

where U_μ, U_ν are the four-velocity vectors in rest state with lowered indices, ρ represents energy density and p represents pressure that is equal in all spatial dimensions for a perfect fluid. Since $U^\mu = (1,0,0,0)^T$, it follows that $U_\mu = g_{\mu\nu} U^\mu = (-1,0,0,0)$. Therefore, the

stress-energy tensor and its trace are as follows:

$$
\begin{aligned}
T_{\mu\nu} &= (\rho + p)\, U_\mu U_\nu + p g_{\mu\nu} \\
&= (\rho + p) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + p \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & \frac{a^2 R_0^2}{1-k\,r^2} & & 0 \\ 0 & 0 & a^2 R_0^2\, r^2 & 0 \\ 0 & 0 & 0 & a^2 R_0^2\, r^2 \sin^2\theta \end{pmatrix} \\
&= \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & \frac{a^2 R_0^2}{1-k\,r^2} p & & 0 \\ 0 & 0 & a^2 R_0^2\, r^2 p & 0 \\ 0 & 0 & 0 & a^2 R_0^2\, r^2 \sin^2\theta\, p \end{pmatrix}; \qquad \text{(B.3.12)} \\
T^\mu{}_\nu &= \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & \frac{1-k\,r^2}{a^2 R_0^2} & & 0 \\ 0 & 0 & \frac{1}{a^2 R_0^2\, r^2} & 0 \\ 0 & 0 & 0 & \frac{1}{a^2 R_0^2\, r^2 \sin^2\theta} \end{pmatrix} \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & \frac{a^2 R_0^2}{1-k\,r^2} p & & 0 \\ 0 & 0 & a^2 R_0^2\, r^2 p & 0 \\ 0 & 0 & 0 & a^2 R_0^2\, r^2 \sin^2\theta\, p \end{pmatrix} \\
&= \begin{pmatrix} -\rho & 0 & 0 & 0 \\ 0 & p & & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}; \text{ and} \qquad \text{(B.3.13)} \\
T &= T^\nu{}_\nu = -\rho + 3p\,. \qquad \text{(B.3.14)}
\end{aligned}
$$

The conservation of energy-momentum in curved spacetime requires that for each value of ν and stress-energy tensor, $T^\mu{}_\nu$,

$$\nabla_\mu T^\mu{}_\nu = 0\,. \qquad \text{(B.3.15)}$$

(See §B.2 *Local Energy-Momentum Conservation* for a derivation of this expression.) Thus, for $\nu = 0$:

$$\nabla_\mu T^\mu{}_0 = 0\,.$$

From the general rule for covariant derivatives in Eq. (A.3.3),

$$\nabla_\mu T^\mu{}_0 = \partial_\mu T^\mu{}_0 + \Gamma^\mu{}_{\mu\lambda} T^\lambda{}_0 - \Gamma^\lambda{}_{\mu 0} T^\mu{}_\lambda\,.$$

Eq. (B.3.13) $T^\mu{}_\nu$ is diagonal. Therefore, on the right side $\mu = 0$ in the first term, $\lambda = 0$ in the second term, and $\lambda = \mu$ in the third term. Therefore,

$$\nabla_\mu T^\mu{}_0 = \partial_0 T^0{}_0 + \Gamma^\mu{}_{\mu 0} T^0{}_0 - \Gamma^\mu{}_{\mu 0} T^\mu{}_\mu\,.$$

Substituting the elements of $T^\mu{}_\nu$ from Eq. (B.3.13) and eliminating all vanishing Christoffel symbols from Eqs. (B.3.4), (B.3.5), (B.3.6), and (B.3.7) yields

$$\begin{aligned}
\nabla_\mu T^\mu{}_0 &= \partial_0(-\rho) + \Gamma^1{}_{10}(-\rho) + \Gamma^2{}_{20}(-\rho) + \Gamma^3{}_{30}(-\rho) - \Gamma^\mu{}_{\mu 0} T^\mu{}_\mu \\
&= \partial_0(-\rho) + 3\frac{\dot{a}}{a}(-\rho) - \Gamma^\mu{}_{\mu 0} T^\mu{}_\mu \\
&= \partial_0(-\rho) + 3\frac{\dot{a}}{a}(-\rho) - \Gamma^0{}_{00} T^0{}_0 - \Gamma^1{}_{10} T^1{}_1 - \Gamma^2{}_{20} T^2{}_2 - \Gamma^3{}_{30} T^3{}_3 \\
&= \partial_0(-\rho) + 3\frac{\dot{a}}{a}(-\rho) - 0 \cdot (-\rho) - 3p\frac{\dot{a}}{a} \\
&= -\dot{\rho} - 3\frac{\dot{a}}{a}(\rho + p).
\end{aligned}$$

This yields the continuity equation:

$$0 = \nabla_\mu T^\mu{}_0 = \dot{\rho} = -3\frac{\dot{a}}{a}(\rho + p)$$

or

$$\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + p) \ . \tag{B.3.16}$$

Dividing both sides by ρ and applying the equation of state, $w = \frac{p}{\rho}$, yields

$$\frac{\dot{\rho}}{\rho} = -3\frac{\dot{a}}{a}(1 + w) \ .$$

Integrating both sides yields an expression of the time evolution of energy density as a function of the conserved parameter, w, and the time dependent scale factor, $a(t)$:

$$\int \frac{\dot{\rho}}{\rho} d\rho = -3(1 + w) \int \frac{\dot{a}}{a} da$$

$$ln\,\rho \propto -3(1 + w)\, ln\, a$$

$$\rho \propto a^{-3(1+w)} \ . \tag{B.3.17}$$

Therefore, if the observed current energy density of source i is ρ_{i0}, then the energy density at time t is

$$\rho_i(t) = \rho_{i0} a(t)^{-n_i} \tag{B.3.18}$$

where $n_i = 3(1 + w_i)$.

By Eq. (B.3.17), the energy density of source i with state of equation parameter w_i is proportional to the scale factor: $\rho_i(t) \propto a(t)^{-3(1+w_i)}$. Therefore, if the observed current energy density of source i is ρ_{i0}, then $\rho_i = \rho_{i0}\, a^{-n_i}$ where $n_i = 3(1 + w_i)$.

If the cosmological fluid consists of co-moving non relativistic (i.e., cold) particles of matter (e.g., dust or galaxies), then the pressure between the matter particles is approximately zero. This is an ideal gas of non relativistic particles of matter:

$$pV = nRT \ .$$

Therefore,

$$p = \frac{n}{V} RT = \rho_m RT$$

where R is the weighted average gas constant for the molecules in the dust and ρ_m is the mass density of the gas. Thermal speed of the molecules in the gas is $\sqrt{RT}$. Since the dust is cold, $\sqrt{RT} \ll c$. By $E = m\,c^2$, energy density of the dust is $\rho_m\,c^2$. Therefore,

$$w = \frac{p}{\rho} = \frac{\rho_m RT}{\rho_m\,c^2} = \frac{RT}{c^2} \simeq 0\,.$$

Therefore, the equation of state is

$$\rho_m(t) \propto \frac{1}{a^3(t)}$$

since $p_m = 0$ and $w = 0$. Therefore, the time evolution of the energy density of the universe dominated by non relativistic particles of matter is inversely proportion to the third power of the scale factor.

Since the trace of the radiation stress energy tensor is zero [21, Eq. (8.59], it follows from Eq. (B.3.14), that $-\rho + 3p = 0$. Therefore, $w = \frac{p}{\rho} = \frac{1}{3}$ and

$$\rho_r(t) \propto \frac{1}{a^4(t)}.$$

Therefore, the time evolution of the energy density of the universe dominated by radiation is inversely proportional to the fourth power of the scale factor.

B.3.2 Friedmann Equations

Overview. The dynamics of the scale factor in the Robertson-Walker metric are developed using the Einstein Equation. The results are the acceleration equation, the Friedmann equation, and a restatement of the continuity equation.

The Acceleration Equation. Applying the components from the expressions for $g_{\mu\nu}$, $R_{\mu\nu}$, $T_{\mu\nu}$ and T (see Eqs. (B.3.1), (B.3.9), (B.3.12), (B.3.14) above), the trace reversed form of the Einstein equation for R_{00} is

$$\begin{aligned} R_{00} &= 8\pi G\left(T_{00} - \frac{1}{2}T g_{00}\right) \\ -\frac{3\ddot{a}}{a} &= 8\pi G\left(\rho - \frac{1}{2}(-\rho + 3p)(-1)\right) \\ &= 8\pi G\left(\frac{1}{2}\rho + \frac{3}{2}p\right) \\ &= 4\pi G(\rho + 3p) \end{aligned}$$

which simplifies to the acceleration equation:

$$\frac{\ddot{a}}{a} = -\frac{4}{3}\pi G(\rho + 3p)\ . \tag{B.3.19}$$

The Friedmann Equation. The expression for R_{11} yields

$$\frac{2\left(\dot{a}^2R_0^2 + k\right) + a\ddot{a}R_0^2}{1 - kr^2} = 8\pi G\left(\frac{a^2R_0^2}{1 - k\,r^2}p - \frac{1}{2}(-\rho + 3p)\left(\frac{a^2R_0^2}{1 - k\,r^2}\right)\right).$$

Multiplying both sides by $\frac{1-kr^2}{a^2R_0^2}$ and simplifying the parenthetical expression on the right side yields

$$\frac{2\left(\dot{a}^2 + k/R_0^2\right) + a\ddot{a}}{a^2} = 4\pi G(\rho - p)\ .$$

Substituting from Eq. (B.3.19) into this expression and solving for $\left(\frac{\dot{a}}{a}\right)^2$ yields the Friedmann equation:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3} - \frac{k}{a^2 R_0^2}\,. \tag{B.3.20}$$

As expected from the isotropy of space, the corresponding expressions for R_{22} and R_{33} both simplify to the Friedmann equation in Eq. (B.3.20).

Restatement of the Continuity Equation. The Hubble parameter is time evolution of the rate of expansion of space. It is defined as

$$H = \frac{\dot{a}}{a}\,. \tag{B.3.21}$$

Substituting definition of the Hubble parameter into the continuity equation in Eq. (B.3.16) yields the following restatement of the continuity equation:

$$\dot{\rho} = -3H\left(\rho + p\right)\,. \tag{B.3.22}$$

B.3.3 Evolution of Energy Density, Scale, and the Hubble Parameter

The scale factor, $a(t)$, has been normalized to $a_0 = 1$ where t_0 represents current time. The observed value of the Hubble parameter in the current era is designated as H_0. The Hubble parameter is defined as $H(t) = \frac{\dot{a}}{a}$. Since H_0 is approximately constant in the current era, $\dot{a} \approx H_0\, a$, which yields $a \approx e^{H_0\, t}$.

From Eq. (B.3.17), the energy density of energy source, i, with state of equation parameter, $w_i\, \rho_i$, is proportional to the scale factor:

$$\rho_i(t) \propto a(t)^{-3(1+w_i)}\,.$$

Therefore, if the observed current energy density of source i is ρ_{i0} , then

$$\rho_i = \rho_{i0} a^{-n_i} \tag{B.3.23}$$

where $n_i = 3\,(1 + w_i)$.

The Friedmann equation for a universe with zero spatial curvature that is strongly dominated by a single energy source is $\left(\frac{\dot{a}}{a}\right)^2 \approx \frac{8\pi G}{3}\rho_i$. For a source with $n_i > 0$, taking the square root of both sides and substituting $\rho_{i0} a^{-n_i}$ for ρ_i yields

$$\begin{aligned} \frac{\dot{a}}{a} &\approx \sqrt{\frac{8\pi G}{3}}\sqrt{\rho_i} \\ &\propto \sqrt{\rho_{i0}}\sqrt{a^{-n_i}} \\ &\propto a^{-\frac{n_i}{2}} \,. \end{aligned}$$

Substituting $\frac{da}{dt}$ for $\dot{a}$, multiplying both sides of the result by $a^{\frac{n_i}{2}} dt$, and integrating both sides of the resulting equation yields

$$\frac{a^{\frac{n_i}{2}}}{\frac{n_i}{2}} \propto t \,.$$

Absorbing the denominator on the left side into the constant of proportionality and solving for the scale factor yields an expression for the scale factor in terms of time:

$$a \propto t^{\frac{2}{n_i}} \,. \tag{B.3.24}$$

It has been established above that $w = 0, \frac{1}{3}, -\frac{1}{3}, -1$ for non relativistic matter (including dark matter), radiation (including relativistic neutrinos), curvature, and dark energy respectively. The energy density, scale factor, and Hubble parameter corresponding to times when energy is dominated by a single source are summarized in Table 2.1.

B.4 The Stress Energy Tensor of a Scalar Field

The derivations in this section are based on [21, 77]. The Lagrangian of a scalar field in curved space is

$$\begin{aligned}\mathcal{L} &= -\frac{1}{2} g^{\mu\nu} \nabla_\mu \phi \nabla_\nu \phi - V(\phi) \\ &= -\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) \\ &= -\frac{1}{2} g^{00} \partial_0 \phi \partial_0 \phi - \frac{1}{2} g^{ij} \partial_i \phi \partial_i \phi - V(\phi) \\ &= \frac{1}{2} \dot{\phi}^2 - \frac{1}{2} (\nabla \phi)^2 - V(\phi) \end{aligned} \tag{B.4.1}$$

where the second step is justified because the covariant derivative of a scalar function is equivalent to the ordinary derivative of the function. The first term is the kinetic energy of the field, the second term is the gradient energy of the field, and the final term is the potential energy.

The action integral of the field is

$$S = \int_{\mathfrak{M}} \sqrt{-g} dx^4 \left[-\frac{1}{2} g^{\mu\nu} \nabla_\mu \phi \nabla_\nu \phi - V(\phi) \right].$$

The variation of the action integral with respect to the variation of the inverse metric vanishes. The resulting equation of motion is

$$g^{\mu\nu} \nabla_\mu \nabla_\nu \phi - \frac{d V}{d \phi} = 0 .$$

The first term is the d'Almbertian in curved spacetime denoted by the symbol $\Box$:

$$\begin{aligned}\Box &\equiv g^{\mu\nu} \nabla_\mu \nabla_\nu \\ &= -\ddot{\phi} + \nabla^2 \phi \end{aligned} \tag{B.4.2}$$

in the $-+++$ metric signature. Therefore, the equation of motion may be written

$$\ddot{\phi} - \nabla^2\phi + \frac{d\,V}{d\,\phi} = 0\,. \tag{B.4.3}$$

If the field is homogeneous, then $\partial_i\,\phi = 0$ for $i = 1, 2, 3$. Using the Robertson-Walker metric in Eq. (A.3.16) and the Christoffel symbols developed in Eqs. (B.3.4)-(B.3.7), the equation of motion of a homogeneous scalar field becomes

$$\ddot{\phi} + 3\,H\,\dot{\phi} + \frac{d\,V}{d\,\phi} = 0 \tag{B.4.4}$$

where H is the Hubble parameter. The Hubble parameter is a friction term in the equation of motion causing the value of $V(\phi)$ to change slowly.

Applying Eqs. (B.1.5) and (B.1.10), the components of the stress energy tensor corresponding to the scalar field are determined as follows:

$$\begin{aligned}
T_{\mu\nu} &= -2\frac{\delta\Big(\mathcal{L}_\phi\,\sqrt{-g}\Big)}{\sqrt{-g}\,\delta g^{\mu\nu}} \\
&= -2\frac{\delta\Big(\Big(-\frac{1}{2}g^{\mu\nu}\nabla_\mu\,\phi\,\nabla_\nu\,\phi - V(\phi)\Big)\,\sqrt{-g}\Big)}{\sqrt{-g}\,\delta g^{\mu\nu}} \\
&= -2\frac{\delta\Big(-\frac{1}{2}g^{\mu\nu}\nabla_\mu\,\phi\,\nabla_\nu\,\phi - V(\phi)\Big)\sqrt{-g} + \Big(-\frac{1}{2}g^{\mu\nu}\nabla_\mu\,\phi\,\nabla_\nu\,\phi - V(\phi)\Big)\delta\,\sqrt{-g}}{\sqrt{-g}\,\delta g^{\mu\nu}} \\
&= -2\frac{\Big(-\frac{1}{2}\delta g^{\mu\nu}\nabla_\mu\,\phi\,\nabla_\nu\,\phi\Big)\,\sqrt{-g} + \Big(-\frac{1}{2}g^{\mu\nu}\nabla_\mu\,\phi\,\nabla_\nu\,\phi - V(\phi)\Big)\Big(-\frac{1}{2}\,\sqrt{-g}\,g_{\mu\nu}\delta g^{\mu\nu}\Big)}{\sqrt{-g}\,\delta g^{\mu\nu}} \\
&= -2\left(\left(-\frac{1}{2}\nabla_\mu\,\phi\,\nabla_\nu\,\phi\right) + \left(-\frac{1}{2}g^{\rho\sigma}\nabla_\rho\,\phi\,\nabla_\sigma\,\phi - V(\phi)\right)\left(-\frac{1}{2}\,g_{\mu\nu}\right)\right) \\
&= \nabla_\mu\,\phi\,\nabla_\nu\,\phi - \frac{1}{2}g_{\mu\nu}g^{\rho\sigma}\nabla_\rho\,\phi\,\nabla_\sigma\,\phi - V(\phi)g_{\mu\nu}\,.
\end{aligned} \tag{B.4.5}$$

Since the covariant derivative of a scalar function is equivalent to the ordinary derivative, the previous expression may be written

$$T_{\mu\nu} = \partial_\mu \phi \, \partial_\nu \phi - \frac{1}{2} g_{\mu\nu} g^{\rho\sigma} \partial_\rho \phi \, \partial_\sigma \phi - V(\phi) g_{\mu\nu} \, . \tag{B.4.6}$$

The energy density of the scalar field is

$$\begin{aligned} \rho &= \partial_0 \phi \, \partial_0 \phi - \frac{1}{2} g_{00} g^{\rho\sigma} \partial_\rho \phi \, \partial_\sigma \phi - V(\phi) g_{00} \\ &= \dot{\phi}^2 + \frac{1}{2} g^{\rho\sigma} \partial_\rho \phi \, \partial_\sigma \phi + V(\phi) \\ &= \frac{1}{2} \dot{\phi}^2 + \frac{1}{2} g^{ij} \partial_i \phi \, \partial_j \phi + V(\phi) \, . \end{aligned} \tag{B.4.7}$$

In the case of a homogeneous scalar field, $\partial_i \phi = 0$; therefore, the energy density reduces to

$$\rho = \frac{1}{2} \dot{\phi}^2 + V(\phi) \, . \tag{B.4.8}$$

Additionally, the isotropic pressure of the scalar field for each $i = 1, 2, 3$ is

$$\begin{aligned} p &= \partial_i \phi \, \partial_i \phi - \frac{1}{2} g_{ii} g^{\rho\sigma} \partial_\rho \phi \, \partial_\sigma \phi - V(\phi) g_{ii} \\ &= 0 - \frac{1}{2} g^{\rho\sigma} \partial_\rho \phi \, \partial_\sigma \phi - V(\phi) \\ &= -\frac{1}{2} g^{00} \partial_0 \phi \, \partial_0 \phi - V(\phi) \\ &= \frac{1}{2} \dot{\phi}^2 - V(\phi) \, . \end{aligned} \tag{B.4.9}$$

B.5 Threshold Temperature

Continuous creation of a particle/antiparticle pair occurs in a plasma of particles in thermal equilibrium if the temperature of the plasma is high enough so that the combined energy of two directly colliding photons equals the rest energy of the particle pair. Since

photons have nearly equal energy when the plasma is in thermal equilibrium, $2\,h\nu = 2\,m\,c^2$ where m is the rest mass of the particle. The average energy of a photon at temperature, T, is equal to the ratio of energy density for Eq. (B.13.5) to number density in Eq. (B.13.8). Therefore, the average energy per photon in a gas is

$$h\bar{\nu} = \frac{\rho(T)}{n(T)} = \frac{\pi^4\, k_B\, T}{30\zeta(3)} \simeq 2.7 \times k_B\, T \tag{B.5.1}$$

and the threshold or formation temperature is

$$\begin{aligned} T_{threshold} &\simeq \frac{m\,c^2}{2.7 \times k_B} \\ &\approx m\,c^2\,{k_B}^{-1}\,. \end{aligned} \tag{B.5.2}$$

B.6 Thomson Scattering and Optical Depth

Since the energy of photons is less than the rest mass of electrons and protons, the interactions between charged particles and photons are in the form of Thomson scattering. The electromagnetic wave incident on a charged particle accelerates the particle in the direction of the oscillating electric field. The radiation emitted by the accelerated particle (the scattered wave) has the same frequency as the incident wave. Since the reverse process is the same as the original process, the probability of each process is the same. If the reaction rates are greater than the expansion rate given by the Hubble parameter, then the charged particles and photons remain in thermal equilibrium. The interaction rates are based on the number density of the charged particles and the Thomson cross-section given by [33]

$$\sigma_T = \frac{8\pi}{3}\left(\frac{\alpha\hbar c}{m\,c^2}\right)^2\,. \tag{B.6.1}$$

The mass of a proton is nearly 2,000 times greater than the mass of an electron. Consequently, the proton scattering cross-section is more than six orders of magnitude smaller than the electron scattering cross-section. Given the comparatively small scattering cross-section of protons and the even smaller cross-section of heavier ions, the probability of Thomson scattering through the plasma is effectively based on the number density of electrons.

If I_{source} is the intensity of light at the source and I_{obs} is the observed intensity of light after it has traveled a path through some medium, then the optical depth, τ, is given by [106]

$$I_{obs}/I_{source} = e^{-\tau(t)} \ .$$

The optical depth of light traveling through a medium of free electrons from time t to the present time, t_0, is

$$\tau(t) = \sigma_T \int_t^{t_0} n_e(t) dt \tag{B.6.2}$$

where n_e is number density of electrons, σ_{T_t} is the Thomson electron scattering cross-section, and $n_e\sigma_{T_t}$ is the probability that a photon will scatter per unit of time [77]. This equation may be written as a function of z:

$$\begin{aligned} \tau(z) &= \sigma_T H_0^{-1} \int_0^z n_e(z') \frac{dz'}{E(z')} \\ &= n_{e0}\sigma_T H_0^{-1} \int_0^z (1+z')^2 \frac{dz'}{E(z')} \end{aligned} \tag{B.6.3}$$

using Eq. (2.3.13) to obtain the radial path of light, $dt = -\frac{1}{1+z}dz$ from Eqs. (2.3.1) and (2.3.5), and $n_e(z) = (1+z)^3 n_{e0}$. Therefore, the probability that light emitted at redshift z

traveled without scattering through to the current era is

$$Pr(z) = e^{-\tau(z)} \tag{B.6.4}$$

where $\tau(z)$ is given by Eq. (B.6.3).

B.7 Standard Model Elementary Particles Table

The following table is based on information published in the 2014 Review of Particle Physics prepared by the Particle Data Group. For more information see [80]. Information from the Particle Data Group was used in Eq. (B.5.2) to determine the approximate threshold temperature for particle pair creation presented in the table. Data in column g are used in Eqs. (B.13.12) to calculate the time of phase transitions and other events in early cosmological history. The symbol g represents the total number of spin states times the number of color states (if any) for the particle and its anti-particle (if any).

The effective degrees of freedom, $\mathcal{N}$, at a particular temperature, T, is determined based on elementary particles with threshold temperature less than T as follows:

$$\mathcal{N} = \sum_{i,bosons} g_i + \frac{7}{8} \sum_{i,fermions} g_i \,. \tag{B.7.1}$$

Table B.1: Standard Model Particles

Particle	**Type**	**Spin States**	**Color States**	**Rest Energy**	**Threshold**	**g**
Higgs	B	1		$126\,GeV$	$10^{15}K$	1
Photons	B	2				2
Gluons	B	2	8			16
$W^{\pm}$ pair	B	3		$80.4\,GeV$	$10^{15}K$	6
Z^0	B	3		$91.2\,GeV$	$10^{15}K$	3
Electron-Neutrino	F	1		$< 2\,eV$	$10^{4}K$	2
Muon-Neutrino	F	1		$< 0.19\,MeV$	$10^{9}K$	2
Tau-Neutrino	F	1		$< 18.2\,MeV$	$10^{11}K$	2
Electron/positron	F	2		$0.511\,MeV$	$10^{10}K$	4
Muon	F	2		$105.7\,MeV$	$10^{12}K$	4
Tau	F	2		$1.777\,GeV$	$10^{13}K$	4
Up	F	2	3	$2.3^{+0.7}_{-0.5}\,MeV$	$10^{10}K$	12
Down	F	2	3	$4.8^{+0.5}_{-0.3}\,MeV$	$10^{11}K$	12
Strange	F	2	3	$95^{+5}_{-5}\,MeV$	$10^{12}K$	12
Charm	F	2	3	$1.275^{+0.025}_{-0.0255}\,GeV$	$10^{13}K$	12
Bottom	F	2	3	$4.18^{+0.03}_{-0.03}\,GeV$	$10^{13}K$	12
Top	F	2	3	$173.21^{+0.51}_{-0.51}\,GeV$	$10^{15}K$	12

B.8 Expansion: Flatness

In this section, an expression is derived for the quantity of expansion during the inflation epoch that produces the current magnitude of the curvature density parameter. In this section, the subscripts *i*, *e*, and 1 represent values at the beginning of inflation, the end of

inflation, and the beginning of the radiation era, respectively. The derivation is based on the following three assumptions:

- Space expanded approximately exponentially during the inflation epoch (hence $a \propto e^{\lambda t}$ for some constant λ and $H = \dot{a}/a \approx \lambda$),
- The magnitude of the curvature density parameter at the beginning of inflation was unity, and
- The rate of expansion, $\dot{a} = aH$, did not change much from the end of inflation at t_e to the beginning of the radiation era at t_1 (hence $a_e H_e \approx a_1 H_1$).

If the magnitude of the curvature density parameter was unity at the beginning of inflation (meaning that the energy density was either zero or twice the critical density), then $a_i^2 = |k| / \lambda^2 R_0^2$ and $a_e^2 H_e^2 = |k| / \lambda^2 R_0^2 e^{2N} \lambda^2$. Therefore, the magnitude of the curvature density parameter at the end of inflation may be written [106, Section 4.1]

$$\begin{aligned} \left|\Omega_{ke}\right| &= \frac{|k|}{a_e^2 H_e^2 R_0^2} \\ &= \frac{|k|}{\frac{|k|}{\lambda^2 R_0^2} e^{2N} \lambda^2 R_0^2} \\ &= e^{-2N} . \end{aligned} \tag{B.8.1}$$

Since $e^{2N} = \frac{a_e^2 H_e^2 R_0^2}{|k|} \rightarrow e^{-2N} \frac{a_e^2 H_e^2 R_0^2}{|k|} = 1$, the current magnitude of the curvature density parameter may be written

$$\begin{aligned} \left|\Omega_{k0}\right| &= \frac{|k|}{a_0^2 H_0^2 R_0^2} \\ &= e^{-2N} \frac{a_e^2 H_e^2 R_0^2}{|k|} \frac{|k|}{a_0^2 H_0^2 R_0^2} \\ &= e^{-2N} \left(\frac{a_e H_e}{a_0 H_0}\right)^2 . \end{aligned} \tag{B.8.2}$$

Referring to Table 2.1, the scale factor at the time that the energy density of radiation equaled the energy density of matter is determined as follows:

$$\rho_{m0} a_{eq}^{-3} = \rho_{r0} a_{eq}^{-4} \rightarrow \; a_{eq} = \rho_{r0}/\rho_{m0} = \Omega_{r0}/\Omega_{m0} \, . \tag{B.8.3}$$

Substituting $\rho_i = \rho_{i0} a^{n_i}$ from Table 2.1 yields the Hubble parameter at radiation-matter equality as follows:

$$\begin{aligned}
H_{eq} &= \sqrt{\frac{8\pi G}{3}\left(\rho_{m,eq} + \rho_{r,eq}\right)} \\
&= \sqrt{\frac{8\pi G}{3} 2\rho_{m,eq}} \\
&= \sqrt{\frac{8\pi G}{3} 2\rho_{m,0} a_{eq}^{-3}} \\
&= \sqrt{2\rho_{m,0}} \sqrt{\frac{8\pi G}{3}} a_{eq}^{-3/2} \\
&= \frac{\sqrt{2\rho_{m,0}}}{\sqrt{\rho_0}} \sqrt{\frac{8\pi G}{3}} \frac{\sqrt{\rho_0}}{1} a_{eq}^{-3/2} \\
&= \sqrt{\frac{2\Omega_{m,0}}{\Omega_0}} H_0 a_{eq}^{-3/2} \\
&= \sqrt{2\Omega_{m,0}} H_0 a_{eq}^{-3/2}
\end{aligned} \tag{B.8.4}$$

where ρ_0 represents the current total energy density and $\Omega_0 \approx 1$ is used in the final step.

Throughout the era of radiation-matter dominance, the constant value of ρ_Λ is small compared to the dynamic values of ρ_m and ρ_r. When calculating the Hubble parameter during this period, dark energy density may be neglected. Therefore, the Hubble parameter may be written

$$H = \sqrt{\frac{8\pi G}{3}} \sqrt{\rho_{m,0} \, a^{-3} + \rho_{r,0} \, a^{-4}} \, .$$

If $\sqrt{\rho_{m,0}}$ is factored out from the radical, $\sqrt{\Omega_0} = \frac{\sqrt{\rho_0}}{\sqrt{\rho_{critical,0}}} \approx 1$, and $a_{eq} = \rho_{r0}/\rho_{m0}$, the Hubble parameter may be written

$$\begin{aligned} H &= \sqrt{\frac{8\pi G}{3}}\sqrt{\rho_{m,0}}\,\sqrt{\Omega_0}\sqrt{a^{-3} + \frac{\rho_{r,0}}{\rho_{m,0}}a^{-4}} \\ &= \sqrt{\frac{8\pi G}{3}}\sqrt{\rho_{m,0}}\,\frac{\sqrt{\rho_0}}{\sqrt{\rho_{critical,0}}}\sqrt{a^{-3} + a_{eq}\,a^{-4}} \\ &= \sqrt{\frac{8\pi G\rho_0}{3}}\,\frac{\sqrt{\rho_{m,0}}}{\sqrt{\rho_{critical,0}}}\sqrt{a^{-3} + a_{eq}\,a^{-4}} \\ &= H_0\sqrt{\Omega_{m,0}}\sqrt{a^{-3} + a_{eq}\,a^{-4}}\,. \end{aligned}$$

By factoring out $a_{eq}^{-\frac{3}{2}}$, inserting the identity $\frac{\sqrt{2}}{\sqrt{2}}$, and substituting $H_{eq} = \sqrt{2\Omega_{m,0}}H_0 a_{eq}^{-3/2}$ from above, the Hubble parameter may be written

$$\begin{aligned} H &= \frac{\sqrt{2\Omega_{m,0}}H_0 a_{eq}^{-\frac{3}{2}}}{\sqrt{2}}\sqrt{\left(\frac{a_{eq}}{a}\right)^3 + \left(\frac{a_{eq}}{a}\right)^4} \\ &= \frac{H_{eq}}{\sqrt{2}}\sqrt{\left(\frac{a_{eq}}{a}\right)^3 + \left(\frac{a_{eq}}{a}\right)^4}\,. \end{aligned} \tag{B.8.5}$$

Since a_{eq} occurred at around $10^5\,years$ (may be calculated using Eq. (2.3.6) based on the value of z_{eq} published by the Planck Collaboration noted above), $a_{eq} \gg a_1$ where a_1 represents the beginning of the radiation era some brief time (preheating and heating period) after the end of inflation when the scale was a_e. Therefore, $\left(\frac{a_{eq}}{a_1}\right)^4 \gg \left(\frac{a_{eq}}{a_1}\right)^3$ and the Hubble parameter may be restated at the beginning of the radiation era as

$$H_1 = \frac{H_{eq}}{\sqrt{2}}\left(\frac{a_{eq}}{a_1}\right)^2$$

[106, Eq. (4.1.4)]. Solving this equation for a_1 yields $a_1 = 2^{-1/4}\sqrt{H_{eq}/H_1}a_{eq}$. By substituting this expression for a_1 into the previous equation, $a_1 H_1$ may be written

$$\begin{aligned} a_1 H_1 &= 2^{-1/4}\sqrt{H_{eq}/H_1}a_{eq} H_1 \\ &= 2^{-1/4}\sqrt{H_1}\sqrt{H_{eq}}a_{eq} \\ &= 2^{-1/4}\sqrt{H_1}\sqrt{\sqrt{2\Omega_{m,0}}H_0 a_{eq}^{-3/2}}a_{eq} \\ &= 2^{-1/4}\sqrt{H_1}2^{-1/4}\Omega_{m,0}^{1/4}\sqrt{H_0}a_{eq}^{-3/4}a_{eq} \\ &= \sqrt{H_1}\sqrt{H_0}\,\Omega_{m,0}^{1/4}\,a_{eq}^{1/4} \end{aligned} \tag{B.8.6}$$

where Eq. (B.8.4) was used in the third step.

From the previous expression and $\Omega_{r0} = a_{eq}\,\Omega_{m0}$, $\frac{a_1 H_1}{a_0 H_0}$ may be written

$$\begin{aligned} \frac{a_1 H_1}{a_0 H_0} &= \frac{\sqrt{H_1}\sqrt{H_0}\,\Omega_{m,0}^{1/4}\,a_{eq}^{1/4}}{a_0 H_0} \\ &= \sqrt[4]{\Omega_{r0}}\sqrt{\frac{H_1}{H_0}}\,. \end{aligned} \tag{B.8.7}$$

The term $\sqrt{\frac{H_1}{H_0}}$ may be restated as

$$\begin{aligned} \sqrt{\frac{H_1}{H_0}} &= \sqrt{\sqrt{8\pi G\rho_1/3}/\sqrt{8\pi G\rho_0/3}} \\ &= \sqrt[4]{\rho_1/\rho_0} \\ &= \sqrt[4]{\frac{\rho_1}{\rho_{critical,0}}/\frac{\rho_0}{\rho_{critical,0}}} \\ &= \sqrt[4]{\frac{\rho_1}{\rho_{critical,0}}/\Omega_0} \\ &= \sqrt[4]{\rho_1/\rho_{critical,0}}\,. \end{aligned} \tag{B.8.8}$$

Substituting Eq. (B.8.8) into Eq. (B.8.7) yields

$$\frac{a_1 H_1}{a_0 H_0} = \sqrt[4]{\Omega_{r0} \frac{\rho_1}{\rho_{critical,0}}} . \tag{B.8.9}$$

Assuming that $a_1 H_1 \approx a_e H_e$ [106, Section 4.1] and substituting Eq. (B.8.9) into Eq. (B.8.2), the magnitude of current curvature density may be written

$$\left|\Omega_{k0}\right| = e^{-2N} \sqrt{\Omega_{r0} \frac{\rho_1}{\rho_{critical,0}}} \tag{B.8.10}$$

where ρ_1 is the energy density at the beginning of the radiation era. Therefore, the number of e-foldings during inflation required to produce the current magnitude of the curvature density parameter may be written

$$N = \frac{1}{4} \ln\left(\frac{\Omega_{r0}}{\Omega_{k0}^2} \frac{\rho_1}{\rho_{critical,0}} \right) . \tag{B.8.11}$$

B.9 Expansion: Horizon

In this section, an expression is derived for the quantity of expansion during the inflation epoch that explains the near perfect isotropy of the LSS. In this section, the subscripts i, e, 1, and $*$ represent values at the beginning of inflation, the end of inflation, the beginning of the radiation era, and the formation of the LSS, respectively. As determined on page 41, for these regions in the LSS to be thermally isotropic, inflation must increase the physical horizon, d_H, at the time the LSS was formed, t_*, so that the physical horizon exceeds the angular diameter distance, $d_A(t_*)$. Thus,

$$\frac{d_H(t_*)}{d_A(t_*)} > \mathcal{I} \geq 1 \tag{B.9.1}$$

leaving $\mathcal{I}$ as a parameter.

The following derivation is based on the following two assumptions:

- Space expanded approximately exponentially during the inflation epoch (hence $a \propto e^{\lambda t}$ for some constant λ and $H = \dot{a}/a \approx \lambda$) and
- The rate of expansion $\dot{a} = aH$ did not change much from the end of inflation at t_e to the beginning of the radiation era at t_1 (hence $a_e\, H_e \approx a_1\, H_1$).

Applying Eq. (2.3.10) and Eq. (2.3.17),

$$\begin{aligned}
\frac{d_H(t_*)}{d_A(t_*)} &= \frac{1}{d_A(t_*)}\frac{1}{z_*+1}\int_{t_i}^{t_*}\frac{dt'}{a(t')} \\
&= \frac{1}{d_A(t_*)}\frac{1}{z_*+1}\int_{t_i}^{t_e}\frac{dt'}{a(t')} + \underbrace{\frac{1}{d_A(t_*)}\frac{1}{z_*+1}\int_{t_e}^{t_*}\frac{dt'}{a(t')}}_{\approx 0.02 \text{ from p. 40}} \\
&\approx \frac{1}{d_A(t_*)}\frac{1}{z_*+1}\int_{t_i}^{t_e}\frac{dt'}{a(t')} + 0.02 \\
&\approx \frac{1}{d_A(t_*)}\frac{1}{z_*+1}\int_{t_i}^{t_e}\frac{dt'}{a(t')}\,.
\end{aligned}$$

Therefore, $d_H(t_*)$ may be written

$$d_H(t_*) \approx \frac{1}{z_*+1}\int_{t_i}^{t_e}\frac{dt'}{a(t')}\,.$$

Assuming that the scale factor was exponential during inflation with $a(t) = a_i\, e^{H_e\,(t-t_i)}$ [106] then, equivalently, $a(t) = a_e\, e^{-H_e\,(t_e-t)}$ and $d_H(t_*)$ may be written

$$\begin{aligned}
d_H(t_*) &\approx \frac{1}{z_*+1}\frac{1}{a_e}\int_{t_i}^{t_e} 1/e^{-H_e\,(t_e-t)}dt \\
&= \frac{1}{z_*+1}\frac{1}{a_e}\int_{t_i}^{t_e} e^{H_e\,(t_e-t)}dt \\
&= \frac{1}{z_*+1}\frac{-1}{a_e\,H_1}\left(e^{H_e\,0} - e^{H_e\,(t_e-t_i)}\right) \\
&= \frac{1}{z_*+1}\frac{1}{a_e\,H_1}\left(e^{H_e\,(t_e-t_i)} - 1\right)\,.
\end{aligned}$$

Since, from above, $d_A = \frac{1}{z_*+1} H_0^{-1} \times 3.1$, $d_A(t_*)$ is of the order $\frac{1}{z_*+1} H_0^{-1}$. Defining $N = H_e\,(t_e - t_i)$, the condition in Eq. (B.9.1) may be written

$$\begin{aligned}
\mathcal{I} &< \frac{d_H(t_*)}{d_A(t_*)} \\
&\approx \frac{\frac{1}{z_*+1}\frac{1}{a_e\,H_1}\left(e^N - 1\right)}{\frac{1}{z_*+1} H_0^{-1}} \\
&= \frac{H_0}{a_e\,H_1}\left(e^N - 1\right) \\
&= \frac{a_0\,H_0}{a_e\,H_1}\left(e^N - 1\right)
\end{aligned}$$

or, since $e^N \gg 1$,

$$e^N > \mathcal{I} \cdot \frac{a_e\,H_1}{a_0\,H_0}$$

[106, Eq. (4.1.9)]. By Eq. (B.8.8) and the assumption that $a_e\,H_e = a_a\,H_1$, the previous expression may be written

$$e^N > \mathcal{I} \cdot \sqrt[4]{\Omega_{r0}\frac{\rho_1}{\rho_{critical,0}}}\,.$$

Therefore, the number of e-foldings during the inflation epoch that is required to explain the near perfect isotropy of the LSS in solid angles greater than about 7° is

$$N > \ln\left(\mathcal{I} \cdot \sqrt[4]{\frac{\Omega_{r0}}{\rho_{critical,0}}}\,\sqrt[4]{\rho_1}\right). \tag{B.9.2}$$

B.10 Conservation of the Curvature Perturbation

Consider a region of scale k^{-1} with causal exit at $t_1 \equiv k \sim aH$ and re-entry at $t_2 \equiv k \sim aH$. The region is causally disconnected from surrounding regions during the period $t_1 \rightarrow t_2$. It is an isolated region that evolves as a separate universe.

Establish a local scale factor for the isolated region produced by the curvature perturbation, $\zeta(\mathbf{x}, t)\Big|_k$, as

$$a(\mathbf{x}, t) = a(t)e^{\zeta(\mathbf{x},t)} \tag{B.10.1}$$

[77, Eq. (5.7)] where the scale indicator has been suppressed for clarity. It is evident that the scale factor loses its position dependence if ζ vanishes. All regions are smoothed to wavenumber k. The term $\zeta(\mathbf{x}, t)$ is understood to apply to all positions in the region. Therefore, the scale factor and the metic are smoothed to scale k^{-1}. Applying the continuity equation Eq. (B.3.22) to this region yields

$$\begin{aligned}
\dot{\rho}(t) &= -3\, H(\mathbf{x}, t)\, (\rho(t) + p(\mathbf{x}, t)) \\
&= -3\, \frac{\dot{a}(\mathbf{x}, t)}{a(\mathbf{x}, t)} (\rho(t) + p(\mathbf{x}, t) \\
&= -3\, \frac{\frac{d}{dt}\left(a(t)e^{\zeta}\right)}{a(t)e^{\zeta}}\, (\rho(t) + p(\mathbf{x}, t)) \\
&= -3\, \frac{\dot{a}e^{\zeta} + \dot{\zeta}\, a\, e^{\zeta}}{ae^{\zeta}}\, (\rho(t) + p(\mathbf{x}, t)) \\
&= -3\, \left(H + \dot{\zeta}\right) (\rho(t) + p(\mathbf{x}, t)) \ .
\end{aligned} \tag{B.10.2}$$

Re-arranging the previous expression and applying the equations of state and density formulas in Table 2.1 for $\alpha = \gamma, \nu, b, c, \Lambda$ ($\alpha = \phi$ with w_ϕ unknown but assumed constant as a

result of the homogeneous and isotopic background field), yields

$$\begin{aligned}
\frac{d\zeta}{dt} &= -\frac{1}{3}\frac{\dot{\rho}(t)}{(\rho(t)+p(\mathbf{x},t))} - H \\
&= -\frac{1}{3}\frac{\frac{d}{dt}\left(\sum_\alpha \left[\rho_{\alpha 0}\, a^{-3(1+w_\alpha)}\right]\right)}{\sum_\alpha \left[(1+w_\alpha)\rho_{\alpha 0}\, a^{-3(1+w_\alpha)}\right]} - H \\
&= \frac{\sum_\alpha \left[(1+w_\alpha)\rho_{\alpha 0}\, a^{-3(1+w_\alpha)-1}\,\dot{a}\right]}{\sum_\alpha \left[(1+w_\alpha)\rho_{\alpha 0}\, a^{-3(1+w_\alpha)}\right]} - H \\
&= \dot{a}a^{-1}\frac{\sum_\alpha \left[(1+w_\alpha)\rho_{\alpha 0}\, a^{-3(1+w_\alpha)}\right]}{\sum_\alpha \left[(1+w_\alpha)\rho_{\alpha 0}\, a^{-3(1+w_\alpha)}\right]} - H \\
&= \frac{\dot{a}}{a} - H \\
&= 0\,. \qquad \text{(B.10.3)}
\end{aligned}$$

Assuming that constant equations of state apply during this epoch is equivalent to assuming that pressure is adiabatic. Therefore, the adiabatic condition for pressure implies that $\zeta(\mathbf{x})\Big|_k$ is conserved while the region is isolated. The derivation may be adapted to demonstrate that conservation of the curvature perturbation requires adiabatic pressure.

B.11 Sky Map Expansion in Spherical Harmonics

In §2.7.3 *Statistical Analysis of Primordial Scalar Perturbations*, Fourier expansion was used to derive the statistical properties of Gaussian perturbations in three dimensional space. It was determined that the ensemble average of the perturbation is zero and that the variance of the perturbation is position independent and is the integral sum of the power spectrum over k-space.

For example, if temperature fluctuations in the CMB are a Gaussian random field, then the ensemble average of the temperature fluctuation at every point in space is zero. The variance of the fluctuation is position independent and is equal to the integral of the power spectrum taken over all k. The power spectrum at a given k is the two-point correlator of

the Fourier transformation of the fluctuation into k-space where k is the wavenumber. See §2.7.3 for an explanation of these terms.

The method of spherical harmonics is a variant of the Fourier method adapted to functions over a two-dimensional sphere. Since the CMB is observed as a spherical shell of light, spherical harmonics is used instead of Fourier analysis to find the statistical properties of temperature fluctuations in the CMB.

If $f(\mathbf{n})$ is a real function over a sphere where $\mathbf{n}$ is a unit vector, then $f(\mathbf{n})$ may be expanded in spherical harmonics by the following transformation and its inverse:

$$a_{\ell m} = \int_{\Omega} f(\mathbf{n})\, Y^*_{\ell m}(\mathbf{n})\, d\Omega \tag{B.11.1}$$

and

$$f(\mathbf{n}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}\, Y_{\ell m}(\mathbf{n})\,. \tag{B.11.2}$$

$Y_{\ell m}(\mathbf{n})$ is the spherical harmonic function of multipole ℓ and order m [83, Eq. (21.78)]. The orthogonality of the spherical harmonic functions is given by

$$\int_{\Omega} Y^*_{\ell m}(\mathbf{n})\, Y_{\ell' m'}(\mathbf{n}')\, d\Omega = \delta_{\ell\,\ell'}\, \delta_{m\,m'} \tag{B.11.3}$$

[83, Eq. (21.79)]. If $f(\mathbf{n})$ represents random temperature fluctuations with Gaussian distribution on the CMB "shell" at position $\mathbf{n}$:

$$f(\mathbf{n}) = \frac{T(\mathbf{n}) - T_0}{T_0} = \frac{\delta T(\mathbf{n})}{T_0}\,, \tag{B.11.4}$$

then $f(\mathbf{n})$ and its transform, $a_{\ell m}$, are random fields. The former is a random field in real space, namely the CMB shell. The latter is a random field in the spherical harmonic space defined by the transformation that will be called the "co-space". As was the case in the

Fourier space, the coefficients are uncorrelated except as required by the reality condition. Therefore the two-point correlator in co-space is the ensemble average:

$$\langle a_{\ell m}\, a_{\ell' m'} \rangle = \delta_{\ell\,\ell'}\, \delta_{m\,m'} \left\langle |a_{\ell m}|^2 \right\rangle . \tag{B.11.5}$$

The angular power spectrum for multipole ℓ is the predicted ensemble average:

$$C_\ell = \left\langle |a_{\ell m}|^2 \right\rangle . \tag{B.11.6}$$

This is the uncountably infinite sum of each realization of $a_{\ell m}$ for a specific amplitude weighted by the probability of the realization. Since there is only one realization, our universe, the ergodic property (as shown for the Fourier method) is invoked to determine the average values over a sample of observed amplitudes as a substitute for the ensemble average. The ergodic estimates will be called ergodic spectra. The ergodic spectra and the observed amplitudes are decorated with tildes to distinguish them from the values predicted by theory:

$$\begin{aligned} \tilde{C}_\ell &= \frac{1}{2\,\ell + 1} \sum_{m=-\ell}^{\ell} \tilde{a}_{\ell m}\, \tilde{a}^*_{\ell m} \\ &= \frac{1}{2\,\ell + 1} \sum_{m=-\ell}^{\ell} \tilde{a}_{\ell m}\, \tilde{a}_{\ell m} \end{aligned} \tag{B.11.7}$$

[41, Eq. (A3)] where the complex conjugate notion is dropped since the amplitudes are real. For low-ℓ, the number of terms used in the average is small and, therefore, does not provide a good representation of the ensemble average. For larger values of ℓ, the average is a good proxy for the ensemble average.

The process used to collect data representing maps of the CMB fluctuations is described in §2.6. Sets of data are available from the Planck Explorer, the Wilkinson Microwave Anisotropy Probe (WMAP), and the Cosmic Background Explorer (COBE) as well as data

from various balloon and land based instruments including the BICEP2 instrument operating at the South Pole. A set of temperature fluctuations over a range of pixel locations $\mathbf{p} = p(\theta, \phi)$ is collected from each receiver over a long period of time. Data are usually collected in pairs separated by a fixed angle, facilitating the application of two-point correlator functions to calculate power spectra [83]. The pixel size varies by receiver. The pixel size corresponds to a co-moving scale k_p^{-1}. The pixel size of a receiver is determined by its angular resolution. The angular resolution of the Planck Explorer receivers range from about $5' \rightarrow 33'$ ($1' \equiv 1^\circ/60$). The resolution of the WMAP receivers range from $14' \rightarrow 56'$. The COBE resolution is about 7°. Fluctuations at larger scales are developed by combining pixels.

A temperature fluctuation map collected by a receiver is a data set,

$$\{\, \delta T(\mathbf{p}); \mathbf{p} \subset LSS \,\}_{g(\nu)} \,, \tag{B.11.8}$$

where $\delta T(\mathbf{p}) = T(\mathbf{p}) - T_0$; $g \in \{T, E, B\}$ referring to the temperature fluctuations of all filtered photons striking the bolometer surface (T), the electric component of filtered polarized photons striking the bolometer surface (E-mode), and the magnetic component of filtered polarized photons striking the bolometer surface (B-mode); and, ν refers to the frequency of the receiver. The maps are indexed below as δT_i. The amplitudes for each map are estimated as:

$$\begin{aligned} \tilde{a}^i_{\ell m} &= \int_\Omega f(\mathbf{n})\, Y^*_{\ell m}(\mathbf{n})\, d\Omega \\ &\simeq \Omega_p \sum_p \delta T_i(\mathbf{p}) W(\mathbf{p})\, Y^*_{\ell m}(\mathbf{p}) \end{aligned} \tag{B.11.9}$$

where Ω_p is the pixel size of the receiver and $W(\mathbf{p})$ is the weight given to the pixel based on the estimated contamination by foreground (dust, galaxies, hot gas, etc.) [41, Eq. (A1)].

Applying Eq. (B.11.7), ergodic auto-spectra are

$$\tilde{C}_\ell^{ii} = \frac{1}{2\,\ell + 1} \sum_{m=-\ell}^{\ell} \tilde{a}_{\ell m}^{i}\, \tilde{a}_{\ell m}^{i} \tag{B.11.10}$$

where the amplitudes are determined from a single map. Ergodic cross-spectra are defined as

$$\tilde{C}_\ell^{ij} = \frac{1}{2\,\ell + 1} \sum_{m=-\ell}^{\ell} \tilde{a}_{\ell m}^{i}\, \tilde{a}_{\ell m}^{j} \tag{B.11.11}$$

where the amplitudes are determined from two different maps that may be of different types (e.g., T and E) [41, Eq. (A9)].

The ergodic spectra, are combined to provide increased sky coverage and improved accuracy using an inverse coupling matrix, $G^{-1}_{\ell\,\ell'}$, to produce ${}^{obs}C_\ell$ that represents an estimate of the full sky power spectrum of the CMB.

For example, to produce the TT power spectrum, some set of power spectra, $\{\,\tilde{C}_\ell^{ij}\,\}$, are combined. The i and j maps used to determine each $\tilde{C}_\ell^{ij}$ are produced by receivers where each receiver is set to measure fluctuations of all filtered photons striking the bolometer surface. The estimate of the full sky power spectrum of the CMB based on the combined observations is

$${}^{obs}C_\ell^{TT} = \sum_{\ell'} G^{-1}_{\ell\,\ell'} \tilde{C}_{\ell'}^{ij} \tag{B.11.12}$$

[41, Eq. (A8)]. The power spectra ${}^{obs}C_\ell^{TE}$, ${}^{obs}\,C_\ell^{EE}$, and ${}^{obs}C_\ell^{BB}$ are determined similarly. This method allows maps produced by receivers on multiple instruments such as the Planck, WMAP, and, BICEP2 to be combined to form power spectra.

The coupling matrix, $G_{\ell\,\ell'}$, is developed based on the weight functions, $W(\mathbf{n})$. See Ref. [41] for additional information on this method.

Methods used to estimate power spectra, including the method of two-point correlator functions, are described in References [83, 75, 11]. The methods used by the Planck Collaboration to develop the CMB power spectra are described in *Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters* [5].

B.12 Critical Energy Density in cgs Units

In cgs units, $[G] = \frac{cm^3}{g\ sec^2}$. Therefore, the value of the critical density in cgs units at time t is

$$\begin{aligned}\rho_{critical}(t) &= \frac{3H^2(t)}{8\pi G}\frac{g\ sec^2}{cm^3}\left(\frac{km}{Mpc\cdot\ sec}\right)^2 \\ &\approx \frac{3}{8\pi G}\left(\frac{h(t)\times 10^2\times 10^3\times 10^2}{3.086\times 10^{24}}\right)^2\frac{g\ sec^2}{cm^3}\left(\frac{cm}{cm\cdot\ sec}\right)^2 \\ &\approx 1.8772\,h^2(t)\times 10^{-29}\,g\,cm^{-3} \end{aligned} \tag{B.12.1}$$

where $h(t) = H(t)/100$ (the reduced Hubble parameter), $1Mpc \approx 3.086\times 10^{24}\,cm$ and $G \approx 6.67684\times 10^{-8}\frac{g\ sec^2}{cm^3}$.

By Eq. (2.3.20), the current radius of the observable universe is $4.45\times 10^{28}\,cm$. Flat space means that energy density is equal to critical energy density. Therefore, the total mass of the universe may be determined based on Eq. (B.12.1) and $H_0 \simeq 67.04$:

$$\begin{aligned} M &= \frac{4}{3}\left(4.45\times 10^{28}\right)^3 1.8772\left(\frac{67.04}{100}\right)^2\times 10^{-29} \\ &= 3\times 10^{57}\,g\,. \end{aligned} \tag{B.12.2}$$

B.13 Densities and Time In Thermal Equilibrium

The expressions in this section are presented without derivation. Derivations are given in §3.1 of [106].

Notation. The energy of a particle is

$$\begin{aligned} E &= \sqrt{q^2c^2 + (mc^2)^2} \\ &= c\sqrt{q^2 + m^2c^2}\,. \end{aligned} \tag{B.13.1}$$

where the rest mass of the particles is m and the momentum is q (where q is being used to avoid conflict with p used for pressure).

The radiation energy constant is

$$a_B = \frac{8\pi^5 k_B^4}{15c^3h^3} \equiv \frac{\pi^2 k_B^4}{15c^3\hbar^3}\,. \tag{B.13.2}$$

The effective particle number in a plasma of highly relativistic particles in thermal equilibrium is

$$\mathcal{N} = \sum g_b + \frac{7}{8}\sum g_f\,. \tag{B.13.3}$$

The symbol, g, in this context represents the number of spin states of the particle and its antiparticle as listed in §B.7. The first term on the right side represents bosons and the second term represents fermions.

Densities. The number density of a particle at a given temperature, T, is the integral of the product of the number of particles at momentum, q, for each value of q. The equation for the number density of a particle of mass m is

$$n(T) = \frac{4\pi\, g}{(2\pi\,\hbar)^3}\int_0^\infty \frac{q^2}{e^{\frac{c\sqrt{q^2+m^2c^2}}{k_B T}} \pm 1}\,dq\,. \tag{B.13.4}$$

The energy density at a given temperature T is the integral of the product of the number of particles at momentum, q, multiplied by the energy for each value of q. The equation for

the energy density of a particle of mass m is

$$\rho(T) = \int_0^\infty dq\, n(q,T) c \sqrt{q^2 + m^2c^2} \tag{B.13.5}$$

and the equation for the pressure of a particle of mass m is

$$p(T) = \int_0^\infty dq\, n(q,T) \frac{q^2c^2}{3c\sqrt{q^2 + m^2c^2}}\,. \tag{B.13.6}$$

Entropy density is the sum of the energy density and the pressure divided by the temperature:

$$s(T) = \frac{1}{T}\int_0^\infty dq\, n(q,T)\left[c\sqrt{q^2 + m^2c^2} + \frac{q^2c^2}{3c\sqrt{q^2 + m^2c^2}}\right]. \tag{B.13.7}$$

For massless particles and highly relativistic massive particles where $q >> mc^2$ and $c\sqrt{q^2 + m^2c^2} \approx c\,q$:

$$n(T) = \begin{cases} \frac{15\,g\zeta(3)}{\pi^4}\frac{a_B T^3}{k_B} & bosons \\ \frac{3}{4} \times \frac{15\,g\zeta(3)}{\pi^4}\frac{a_B T^3}{k_B} & fermions \end{cases}, \tag{B.13.8}$$

$$\rho(T) = \begin{cases} \frac{g\,a_B T^4}{2} & bosons \\ \frac{7}{8}\frac{g\,a_B T^4}{2} & fermions \end{cases}, \text{ and} \tag{B.13.9}$$

$$s(T) = \begin{cases} \frac{2\,g\,a_B T^3}{3} & bosons \\ \frac{7}{8}\frac{2\,g\,a_B T^3}{3} & fermions \end{cases}. \tag{B.13.10}$$

For example, radiation energy density is given by

$$\rho_\gamma(T) = \frac{g\, a_B\, T^4}{2} = \frac{2\, a_B\, T^4}{2} = a_B\, T^4\,. \tag{B.13.11}$$

Time. Time determined during an epoch during which the universe consists of a highly relativistic particle plasma in thermal equilibrium is given by

$$t = \sqrt{\frac{3c^2}{16\pi G \mathcal{N} a_B}}\frac{1}{T^2}\,. \tag{B.13.12}$$

At temperatures below $10^{11}K$, neutrinos have begun to decouple. Eq. (B.13.12) may be used if the neutrino temperature is adjusted. For example, at $T = 5 \times 10^8 K$, electron-neutrinos are the only fermions in free formation. Therefore,

$$\mathcal{N} = 2 + \frac{7}{8} \times 2 \times \left(\left(\frac{4}{11}\right)^{\frac{1}{3}} S^{\frac{1}{3}}\left(\frac{m_e c^2}{k_B T}\right)\right)^4$$

would be used in Eq. (B.13.12).

B.14 Gaussian Distribution

The probability density of the Gaussian distribution is given by

$$\frac{1}{\sigma\sqrt{2\pi}}\, exp\left(\frac{(g-\mu)^2}{2\sigma^2}\right) \tag{B.14.1}$$

where μ and σ are the mean and the standard deviation respectively [11, Eq. (23.56)]. Thus, the probability that the value of random variable X with Gaussian distribution falls between values a and b is given by

$$Pr[a \le X \le b] = \int_a^b \frac{dg}{\sigma\sqrt{2\pi}}\, exp\left(\frac{(g-\mu)^2}{2\sigma^2}\right). \tag{B.14.2}$$

B.15 Precision Cosmology

Precision cosmology compares the theoretical evolution of the CMB anisotropies to detailed observations of the anisotropies. The theoretical evolution is derived from computer models of nucleosynthesis, recombination, and the power spectra. Parameter fitting programs are used to determine the values of cosmological parameters that best fit observed CMB anisotropies. The programs used by the Planck Collaboration to produce their 2015 reports are described in Section 2 of *Planck 2015 results. XIII. Cosmological parameters* [88].

The nucleosynthesis computer models establish the neutron decay constant, the helium abundance by weight, the relative abundances of light element nuclei, and the effective number of neutrinos. The principal parameter that is used to derive these values is the observed ratio of the number density of baryons to the number density of photons, $\eta \approx 4 \times 10^{-9}$ [106].

The recombination codes and the development of the simulated anisotropy maps with power spectra use the best fit parameters determined from nucleosynthesis simulations as input together with an array of other estimated cosmological parameters. The principal inputs used under the assumption of Gaussian adiabatic perturbations are the average temperature of the CMB, the radiation density parameter, the baryonic matter density parameter, the cold dark matter density parameter, the dark energy density parameter, the Hubble parameter, the redshift at re-ionization, and the spectral index. The parameter fitting process is described in *Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters* [5], *Planck 2015 results. XIII. Cosmological parameters* [88], and*Planck 2015 results. XX. Constraints on inflation* [89].

A cartoon version of the parameter fitting process may be described as follows. Initial values of cosmological parameters are developed from preceding experiments. Prior values could come from previous measurements of CMB anisotropies or from standard astrophysical observations. For example, the curvature of space may be measured based

on the measurements of the total angles of large triangles formed between observers and pairs of distant objects. The total density parameter of ordinary matter can be estimated by the number density of galaxies and the average number of solar masses in a galaxy. The Hubble constant can be approximated by measuring the redshifts of light from Type 1a supernovae. Dark matter in a galaxy can be inferred from the orbital speed of stars in the galaxy versus the estimated mass of ordinary matter in the galaxy based on the stars observed in the galaxy. The redshift at re-ionization may be estimated from the largest redshift of observed galaxies. Total scale invariance may be used as the initial estimate of the spectral index. Prior estimates are used as initial values in the theoretical model of the power spectrum (for example the CAMB codes [88]) to produce a simulated power spectrum curve. The simulated power spectrum is compared to the observed power spectrum. The initial cosmological parameters are incrementally adjusted in tandem to change the shape of the theoretical power spectrum curve to best fit the observed power spectrum. The parameter fitting process used in precision cosmology is, of course, far more complicated (for example, the CosmoMC codes [88]).

Examples of the impact on the theoretical power spectrum curve produced by varying cosmological parameters from their best fit values are provided below. All images were developed based on WMAP observations using the "CMB Analyzer" on the NASA website at http://lambda.gsfc.nasa.gov/education/cmb_plotter/cmb_plot.swf. The red curves with dots (data points) represent the observed power spectrum (which is the same in each graph) and blue curves represent the theoretical power spectrum.

Varying Curvature from Best Fit Value. The best fit value of the curvature density parameter is $\Omega_{k0} = 0$. The pair of graphs in Figure B.1 illustrates the impact on the theoretical power spectrum curve produced by varying the curvature of space from flat space by varying Ω_{k0} from its best fit value. The left side of Figure B.1 shows that the theoretical power spectrum curve is shifted to the left when $\Omega_{k0} = 1 - \Omega_0 < 0$, corresponding to positively

curved space. The right side of Figure B.1 shows that the theoretical power spectrum curve is shifted to the right when $\Omega_{k0} > 0$, resulting in negatively curved space.

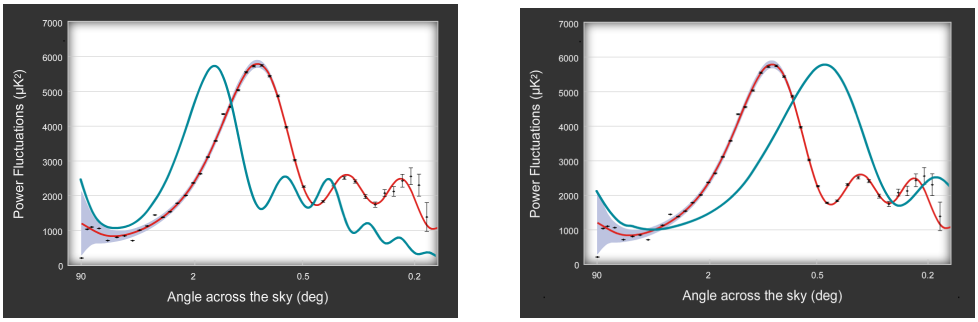


Figure B.1: The curvature density parameter in the theoretical curve is set below the best fit value, $\Omega_{k0} = 0$, on the left side and above the best fit value on the right side.

Varying the Hubble Constant from Best Fit Value. The pair of graphs in Figure B.2 illustrates the impact on the theoretical power spectrum curve produced by varying the Hubble constant from its best fit value, $H_0 \approx 70$. The left side shows that the theoretical curve is shifted upward beginning at about the first peak when the Hubble constant is set below its best fit value and is shifted downward beginning at about the first peak when the Hubble constant is set above its best fit value.

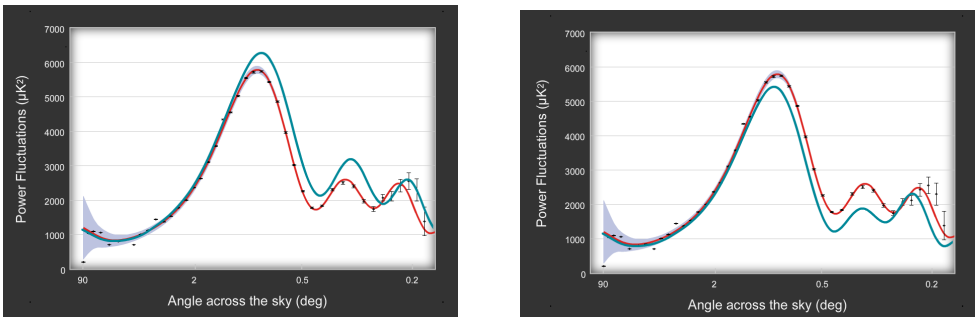


Figure B.2: Hubble parameter in the theoretical curve is set below the best fit value of the Hubble constant on the left side and and above the best fit value on the right side.

Other Variations. The second and third peaks are sensitive to the relative contributions of baryonic and dark matter to the total density parameter. Figure 2.5 in Chapter 2 is based on the best fit values of all parameters. As can be seen in that figure (and in the curves representing the observed power spectrum in the figures above), the second and third peaks have approximately the same height using the best fit values.

If the density parameter of baryonic matter, Ω_{b0}, is increased, then the odd peaks are enhanced over the even peaks. This is because the odd peaks represent the maximum compression phases of baryon acoustic oscillations. Increasing the density parameter of baryonic matter causes greater compression of the plasma by gravity before the radiation pressure causes the plasma to rebound to its maximum rarefaction phase. The even peaks represent the maximum rarefaction phases of baryon acoustic oscillations. Compression heats the plasma; rarefaction cools the plasma. Therefore, increasing the baryon density parameter increases the change in temperature between the extrema.

Baryon acoustic oscillations form a harmonic series. The first mode of the oscillations in the series corresponds to the fundamental scale. The wavelength of the first mode is two times the fundamental scale. The fundamental scale is the sound horizon (the cosmological parameter, $100\theta_{MC}$, described in §2.2). The sound horizon is the co-moving distance

"sound" could travel through the plasma from the beginning of the radiation era to recombination. Therefore, the first mode, k_1, is π divided by the sound horizon. The first mode is at maximum compression (maximum temperature) at recombination. Since oscillations cease at recombination the first mode oscillation is caught at its maximum temperature. The maximum temperature at this angular scale or multipole is imprinted on the CMB. The first peak represents the temperature extrema of mode k_1.

The first harmonic of the fundamental scale is the second mode. The second mode, k_2, has half the wavelength of the first mode ($k_2 = 2\,k_1$) and oscillates twice as fast as the first mode. Therefore, it has had time to reach maximum compression and rebound to reach maximum rarefaction phase at recombination. The second mode is at maximum rarefaction (minimum temperature) at recombination. Since oscillations cease at recombination, the second mode oscillation is caught at its minimum temperature. The minimum temperature at this angular scale or multipole is imprinted on the CMB. The second peak represents the temperature extrema of mode k_2.

The second harmonic of the fundamental scale is the third mode. The third mode, k_3, has one-third the wavelength of the first mode ($k_3 = 3\,k_1$) and oscillates three times as fast as the first mode. Therefore, it has had time to reach maximum compression, rebound to reach maximum rarefaction, and re-compress to reach maximum compression again at recombination. The third mode is at maximum compression (maximum temperature) at recombination. Since oscillations cease at recombination, the third mode oscillation is caught at its maximum temperature. The maximum temperature at this angular scale or multipole is imprinted on the CMB. The third peak represents the temperature extrema of mode k_3. Oscillations occur at all harmonics of the fundamental scale and follow the same pattern.

If the redshift at re-ionization is increased from its best fit value (i.e., star formation begins earlier), then the theoretical power spectrum shifts downward; and if its value is set

lower than its best fit value (i.e., star formation begins later), then the theoretical power spectrum shifts upward.

If the spectral index (also called the tilt) is increased from its best fit value, then the theoretical power spectrum is tilted upward to the right as shown in Figure B.3. Alternatively, if its value is set lower than its best fit value, then the theoretical power spectrum tilts to the left.

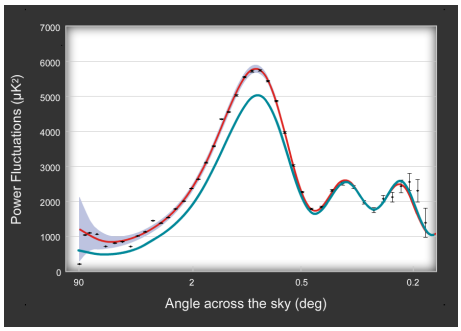
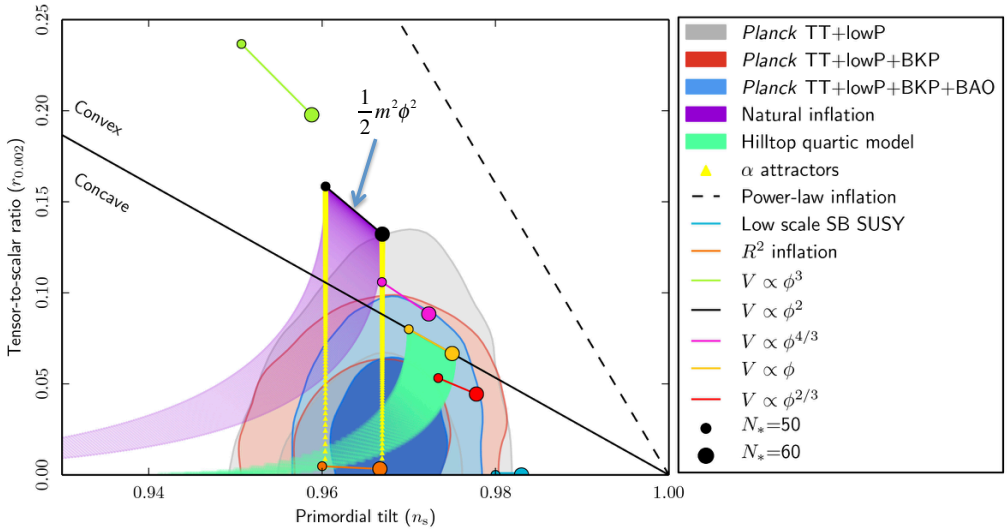


Figure B.3: The theoretical curve tilts upward to the right when the spectral index is set higher than its best fit value.

Using Likelihood Contours to Fit Parameter Pairs. Figure 4.5, copied below, is an illustration of determining the best fit values of two inflation observables, n_s and r, based on likelihood contours developed from the analysis of Planck temperature and polarization data, BICEP2/Keck Array data, and baryon acoustic oscillation analysis.


Tensor-to-scalar ratio ($r_{0.002}$)
Primordial tilt (n_s)
0.00
0.05
0.10
0.15
0.20
0.25
0.94
0.96
0.98
1.00
Convex
Concave
$\frac{1}{2}m^2\phi^2$
Planck TT+lowP
Planck TT+lowP+BKP
Planck TT+lowP+BKP+BAO
Natural inflation
Hilltop quartic model
α attractors
Power-law inflation
Low scale SB SUSY
R^2 inflation
$V \propto \phi^3$
$V \propto \phi^2$
$V \propto \phi^{4/3}$
$V \propto \phi$
$V \propto \phi^{2/3}$
N_*=50
N_*=60

Vita

Peter Alexander Farago was born on September 29, 1951 in Washington, District of Columbia and is a U.S. citizen. He graduated from Thomas Jefferson High School, Richmond, Virginia in 1969. He received his Bachelor of Arts in Mathematics from The George Washington University, Washington, District of Columbia in 1975. He was inducted into the Pi Mu Epsilon Honor Society in Mathematics, GWU Chapter in 1975. He is a retired Enrolled Actuary, Member of the American Academy of Actuaries, Associate of Society of Actuaries, Member of the American Society of Pension Actuaries, and a Fellow in the Conference of Consulting Actuaries. Prior to starting an actuarial and software consulting practice in 1984 he taught mathematics and worked for the Life Insurance Company of Virginia and Peat, Marwick, Mitchell & Co. He sold the majority of his consulting practice in 2000 and the remainder in 2007. He began studying physics at the University of Richmond in 2011 before entering the Master of Science program in physics at Virginia Commonwealth University in 2012. He was inducted into the The Honor Society of Phi Kappa Phi, VCU Chapter in 2014.