These lectures present a brief review of inflationary cosmology, provide an overview of the theory of cosmological perturbations, and then focus on the conceptual problems of the current paradigm of early universe cosmology, thus motivating an exploration of the potential of string theory to provide a new paradigm. Specifically, the string gas cosmology model is introduced, and a resulting mechanism for structure formation which does not require a period of cosmological inflation is discussed.
1. Introduction and Overview

These lectures focus on three topics. The first is an overview of the current paradigm of early universe cosmology, the inflationary universe scenario. The second topic is a pedagogical presentation of the theory of cosmological perturbations, the main tool of modern cosmology which allows us to connect theories of the very early universe with observational data. The third topic is string gas cosmology, an attempt to construct a new scenario of the very early universe based on established principles of string theory.

Over the past two and a half decades, cosmology has become a science dominated by data of rapidly increasing accuracy. Today, we have three-dimensional maps of the distribution of galaxies in space which contain more than one hundred thousand galaxies [1, 2]. They clearly indicate that luminous matter in the universe is neither uniformly nor randomly distributed. There are clear patterns to be seen: clusters of galaxies, superclusters, filaments and voids (regions of space empty of galaxies). The distribution can be quantified in terms of the luminosity power spectrum. A key challenge for cosmology is to understand the origin of these patterns in the distribution of matter.

Another observational window in cosmology is the cosmic microwave background (CMB) radiation. Overall, this radiation is characterized by a surprising isotropy. At a fractional level of a bit less than $10^{-4}$, however, there are anisotropies. These can be quantified in terms of their angular power spectrum. First, the sky map (two-dimensional) of anisotropies is expanded in spherical harmonics $Y_{lm}$:

$$\frac{\Delta T}{T}(\theta, \phi) = \sum_{l=1}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta, \phi) \quad (1.1)$$

where $\theta$ and $\phi$ are the usual angles on the surface of the sphere. If the fluctuations are due to a Gaussian random process with no distinguished direction in the sky, then the complete information about the fluctuations is given by the ensemble average (denoted by pointed parentheses) of the coefficients $a_{lm}$:

$$c_l = \langle |a_{lm}|^2 \rangle , \quad (1.2)$$

These $c_l$ coefficients define the angular power spectrum of CMB anisotropies. Figure 1 is the full sky map of CMB anisotropies from the WMAP experiment [3]. Figure 2 shows the resulting angular power spectrum (what is plotted on the vertical axis is $l(l+1)c_l/(2\pi)$). The key features are the flat region at small values of $l$ (large angular scales) and the characteristic oscillations of the spectrum at intermediate angular scales. Another crucial challenge for cosmology is to explain both the overall isotropy of the CMB, and the specific patterns of anisotropies.

According to our present understanding, we must look to the very early universe to find an explanation for the observed structures. The reason is that in Standard Big Bang cosmology (SBB), which well describes the cosmological evolution at late times (times later and including the period of nucleosynthesis) the physical wavelength of fixed comoving scales is increasing less fast than the Hubble radius (an important length scale which is defined and whose role is described at the end of this section). Hence, the scales which are currently observed were outside of the Hubble radius at early times, and no causal structure formation scenario is possible $^1$.

$^1$Topological defect models [4, 5, 6] provide a way to circumvent this reasoning. They, however, also involve new physics of the very early universe.
Figure 1: All sky map of the temperature anisotropies in the CMB from the WMAP satellite experiment. Credit: NASA/WMAP Science Team.

Figure 2: The power spectrum of CMB anisotropies as computed from the WMAP satellite experiment. The horizontal axis is the angular quantum number \( l \), the vertical axis gives the power of the CMB on the respective scales. The dots represent the data points (with their errors bars indicated), the solid curve is the prediction of the best-fit inflationary model. Credit: NASA/WMAP Science Team.

It is a remarkable success of inflationary cosmology, our current paradigm of early universe cosmology, that it, in addition to explaining why the universe is large, spatially flat, and containing a large amount of entropy, provides a causal mechanism for the origin of inhomogeneities in the universe. The solid curve in Figure 2 represents the predictions of inflationary cosmology. The number of free parameters, however, is much, much smaller than the number of data points.

\[\text{Several cosmological parameters, e.g. the current value of the cosmological constant and the fractional baryon density, have to be fixed in order to obtain this excellent agreement. The number of free parameters, however, is much, much smaller than the number of data points.}\]
predictions were made more than 15 years before the data shown in Figure 1 were collected.

In these lectures, however, I would like to focus less on the phenomenological successes of the inflationary paradigm, but more on the conceptual problems which current realizations of inflationary cosmology are confronted with. These problems call out for the development of a new paradigm, a paradigm which must include new fundamental physics, a theory which better describes space, time and matter and the highest densities. The best current candidate for such a theory is superstring theory. Hence, in the third main part of these lectures (Section 4) I will explore the possibility that string theory may lead to a new paradigm of early universe cosmology. My approach to string cosmology in this section is complementary to the one taken in the lectures of Cliff Burgess at this school (see [7]), in which avenues of obtaining inflation in the context of models coming from string theory are explored (see also [8, 9] for other reviews of such avenues).

The theory of cosmological perturbations plays a key role in modern cosmology, since it provides the techniques with which to calculate, in the context of any given scenario of the very early universe, the generation and evolution of the predicted inhomogeneities in the matter distribution and anisotropies in the CMB, and thus allows for a comparison between fundamental theory and observational data. The second part of these lectures (Section 3) provides a pedagogical overview of this theory.

I begin, however, with a discussion of inflationary cosmology, the current paradigm of cosmology.

To establish our notation and framework, we will be taking the background space-time to be homogeneous and isotropic, with a metric given by

$$ds^2 = dt^2 - a(t)^2 d\bar{x}^2,$$

where $t$ is physical time, $d\bar{x}^2$ is the Euclidean metric of the spatial hypersurfaces (here taken for simplicity to be spatially flat), and $a(t)$ is the scale factor. The scale factor determines the Hubble expansion rate via

$$H(t) = \frac{\dot{a}(t)}{a(t)}.$$

The coordinates $\bar{x}$ used above are comoving coordinates, coordinates painted onto the expanding spatial hypersurfaces.

We are interested in tracking the time evolution of the physical wavelength of the currently observed patterns in the distribution of matter and of the CMB anisotropies. Since the patterns are assumed to be frozen in in comoving coordinates, the physical wavelength scales as $a(t)$.

A key length scale in cosmology is the Hubble radius

$$l_H(t) = H^{-1}(t),$$

defined to be the inverse Hubble expansion rate. As will be explained later, the Hubble radius is the maximal distance that microphysics can act coherently over a Hubble expansion time - in particular it is the maximal distance on which any causal process could create fluctuations.

2. Inflationary Cosmology: The Current Paradigm

Standard Big Bang (SBB) cosmology, the precursor to inflationary cosmology as the paradigm for the evolution of the early universe, is based on a classical physics description of both space-time
Figure 3: Space-time diagram (sketch) showing the evolution of scales in inflationary cosmology. The vertical axis is time, and the period of inflation lasts between \( t_i \) and \( t_R \), and is followed by the radiation-dominated phase of standard big bang cosmology. During exponential inflation, the Hubble radius \( H^{-1} \) is constant in physical spatial coordinates (the horizontal axis), whereas it increases linearly in time after \( t_R \). The physical length corresponding to a fixed comoving length scale labelled by its wavenumber \( k \) increases exponentially during inflation but increases less fast than the Hubble radius (namely as \( t^{1/2} \)), after inflation.

(via Einstein’s theory of General Relativity) and matter (a superposition of two perfect fluids, the first describing pressureless matter - cold matter - the second describing radiation - the CMB). The key phenomenological success of the SBB model is the prediction of the existence and black body nature of the CMB, the black body nature of which was confirmed with spectacular accuracy by the COBE satellite [10] and UBC rocket experiments [11].

On the other hand, the SBB scenario leaves many crucial questions un-answered [12]. Why is the universe so close to spatially flat? Why is it so large and contains such a large entropy? Why is the CMB isotropic to an accuracy of better than \( 10^{-4} \) (after subtracting the dipole contributions due to our motion relative to the rest frame of the CMB and the effects due to the emission of our own galaxy)? Most importantly, what is the origin of the observed inhomogeneities in the distribution of matter and of the small CMB anisotropies? These are the “flatness”, “entropy”, “horizon” and “structure formation” problems of SBB cosmology. SBB also suffers from conceptual problems: the initial cosmological singularity tells us both that the theory must be incomplete, and that it is based on using the wrong fundamental physics input close to the singularity.

Inflationary cosmology [12] (see also [13, 14, 15]) provides an solution of the horizon, flatness and entropy problems. In addition, it provides a mechanism for the origin of structure in the universe based on causal physics [16] (see also [17, 13]).

The idea of inflationary cosmology is to assume that there was a period in the very early Universe during which the scale factor was accelerating, i.e. \( \ddot{a} > 0 \). This implies that the Hubble radius was shrinking in comoving coordinates, or, equivalently, that fixed comoving scales were “exiting” the Hubble radius. In the simplest models of inflation, the scale factor increases nearly exponentially.

\[^3\text{As we will see later on, these conceptual problems persist in inflationary cosmology.}\]
As illustrated in Figure 3, the basic geometry of inflationary cosmology provides a solution of the fluctuation problem. As long as the phase of inflation is sufficiently long, all length scales within our present Hubble radius today originate at the beginning of inflation with a wavelength smaller than the Hubble radius at that time. Thus, it is possible to create perturbations locally using physics obeying the laws of special relativity (in particular causality). As will be discussed later, it is quantum vacuum fluctuations of matter fields and their associated curvature perturbations which are predicted to be responsible for the structure we observe today.

Postulating a phase of inflation in the very early universe solves the horizon problem of the SBB, namely it explains how the causal horizon at the time \( t_{\text{rec}} \) when photons last scatter can be larger than the radius of the past light cone at \( t_{\text{rec}} \), the part of the last scattering surface which is visible today in CMB experiments. Inflation explains the near flatness of the universe: in a decelerating universe spatial flatness is an unstable fixed point of the dynamics, whereas in an accelerating universe it becomes an attractor. Another important feature of inflation is that the volume of space increases exponentially at constant energy density. If this energy density is successfully converted to ordinary matter at the end of the period of inflation, then the entropy of the universe is exponentially larger after compared to before inflation. In addition, with the exponential expansion of space it is easy to produce a universe of our size today from a Planck scale universe at the initial Planck time, something which is not possible in the SBB model. Thus inflation explains the large current size and entropy of the universe.

Let us now consider how it is possible to obtain a phase of cosmological inflation. We will assume that space-time is described using the equations of General Relativity. In this case, the dynamics of the scale factor \( a(t) \) is determined by the Friedmann-Robertson-Walker (FRW) equations

\[
\left( \frac{\dot{a}}{a} \right)^2 = 8\pi G \rho \tag{2.1}
\]

and

\[
\frac{\ddot{a}}{a} = -4\pi G(\rho + 3p) \tag{2.2}
\]

where for simplicity we have omitted the contributions of spatial curvature (since spatial curvature is diluted during inflation) and of the cosmological constant (since any small cosmological constant which might be present today has no effect in the early Universe since the associated energy density does not increase when going into the past). In the above, \( \rho \) and \( p \) denote the energy density and pressure, respectively. From (2.2) it is clear that in order to obtain an accelerating universe, matter with sufficiently negative pressure

\[
p < -\frac{1}{3}\rho \tag{2.3}
\]

is required. Exponential inflation is obtained for \( p = -\rho \).

Conventional perfect fluids have positive semi-definite pressure and thus cannot yield inflation. However, we know that a description of matter in terms of classical perfect fluids must break down at early times. An improved description of matter will be given in terms of quantum fields. Scalar matter fields are special in that they allow at the level of a renormalizable action the presence

---

\(^{\text{4}}\)Note, however, that the first model of exponential expansion of space [15] made use of a higher derivative gravitational action.
of a potential energy term. The energy density and pressure of a scalar field $\phi$ with canonically normalized action $^5$

$$S = \int d^4x \sqrt{-g} \left[ \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - V(\phi) \right]$$  \hspace{1cm} (2.4)

(where Greek indices label space-time coordinates, $V(\phi)$ is the potential energy density, and $g$ is the determinant of the metric) are given by

$$\rho = \frac{1}{2} (\phi)^2 + \frac{1}{2} a^{-2} (\nabla \phi)^2 + V(\phi)$$

$$p = \frac{1}{2} (\phi)^2 - \frac{1}{6} a^{-2} (\nabla \phi)^2 - V(\phi).$$  \hspace{1cm} (2.5)

Thus, it is possible to obtain an almost exponentially expanding universe provided the scalar field configuration $^6$ satisfies

$$\frac{1}{2} (\nabla_{\rho} \phi)^2 \ll V(\phi),$$ \hspace{1cm} (2.6)

$$\frac{1}{2} (\phi)^2 \ll V(\phi).$$ \hspace{1cm} (2.7)

In the above, $\nabla_{\rho} \equiv a^{-1} \nabla$ is the gradient with respect to physical as opposed to comoving coordinates. Since spatial gradients redshift as the universe expands, the first condition will (for single scalar field models) always be satisfied if it is satisfied at the initial time $^7$. It is the second condition which is harder to satisfy. In particular, this condition is in general not preserved in time even it is initially satisfied [20].

It is sufficient to obtain a period of cosmological inflation that the slow-roll conditions for $\phi$ are satisfied. Recall that the equation of motion for a homogeneous scalar field in a cosmological space-time is (as follows from (2.4)) is

$$\ddot{\phi} + 3H \dot{\phi} = -V'(\phi),$$ \hspace{1cm} (2.8)

where a prime indicates the derivative with respect to $\phi$. In order that the scalar field roll slowly, it is necessary that

$$\ddot{\phi} \ll 3H \dot{\phi}$$ \hspace{1cm} (2.9)

such that the first term in the scalar field equation of motion (2.8) is negligible. In this case, the condition (2.7) becomes

$$\frac{(V')^2}{V} \ll 48\pi G$$ \hspace{1cm} (2.10)

and (2.9) becomes

$$\frac{V''}{V} \ll 24\pi G.$$ \hspace{1cm} (2.11)

In the initial model of inflation using scalar fields (“old inflation” [12]), it was assumed that $\phi$ was initially in a false vacuum with large potential energy. Hence, the conditions for inflation

$^5$See [18] for a discussion of fields with non-canonical kinetic terms.

$^6$The scalar field yielding inflation is called the inflaton.

$^7$In fact, careful studies [19] show that since the gradients decrease even in a non-inflationary backgrounds, they can become subdominant even if they initially dominate.
are trivially satisfied. To end inflation, a quantum tunneling event from the false vacuum to the true vacuum [21] was invoked (see e.g. [22] for a pedagogical review). This model, however, has a graceful exit problem since the tunneling leads to an initially microscopical bubble of the true vacuum which cannot grow to encompass our presently observed universe - the flatness problem of SBB cosmology in a new form. Hence, attention shifted to models in which the scalar field $\phi$ is slowly rolling during inflation.

There are many models of scalar field-driven inflation. Many of them can be divided into three groups [23]: small-field inflation, large-field inflation and hybrid inflation. Small-field inflationary models are based on ideas from spontaneous symmetry breaking in particle physics. We take the scalar field to have a potential of the form

$$ V(\phi) = \frac{1}{4} \lambda (\phi^2 - \sigma^2)^2, \quad (2.12) $$

where $\sigma$ can be interpreted as a symmetry breaking scale, and $\lambda$ is a dimensionless coupling constant. The hope of initial small-field models (“new inflation” [24]) was that the scalar field would begin rolling close to its symmetric point $\phi = 0$, where thermal equilibrium initial conditions would localize it in the early universe. At sufficiently high temperatures, $\phi = 0$ is a stable ground state of the one-loop finite temperature effective potential $V_T(\phi)$ (see e.g. [22] for a review). Once the temperature drops to a value smaller than the critical temperature $T_c$, $\phi = 0$ turns into an unstable local maximum of $V_T(\phi)$, and $\phi$ is free to roll towards a ground state of the zero temperature potential (2.12). The direction of the initial rolling is triggered by quantum fluctuations. The reader can easily check that for the potential (2.12) the slow-roll conditions cannot be satisfied if $\sigma \ll m_{pl}$, where $m_{pl}$ is the Planck mass which is related to $G$. If the potential is modified to a Coleman-Weinberg [25] form

$$ V(\phi) = \frac{\lambda}{4} \phi^4 \left[ \ln \left( \frac{\phi}{\sigma} \right) - \frac{1}{4} \right] + \frac{1}{16} \lambda \sigma^4 $$

(2.13)

(where $\sigma$ denotes the value of the minimum of the potential) then the slow-roll conditions can be satisfied. However, this corresponds to a severe fine-tuning of the shape of the potential. A further problem for most small-field models of inflation (see e.g. [20] for a review) is that the slow-roll trajectory is not an attractor in phase space. In order to end up close to the slow-roll trajectory, the initial field velocity must be constrained to be very small. This initial condition problem of small-field models of inflation affects a number of recently proposed brane inflation scenarios, see e.g. [26] for a discussion.

There is another reason for abandoning small-field inflation models: in order to obtain a sufficiently small amplitude of density fluctuations, the interaction coefficients of $\phi$ must be very small (this problem is discussed in detail at the beginning of Section 4). In particular, this makes it inconsistent to assume that $\phi$ started out in thermal equilibrium. In the absence of thermal equilibrium, the phase space of initial conditions is much larger for large values of $\phi$.

This brings us to the discussion of large-field inflation models, initially proposed in [27] under the name “chaotic inflation”. The simplest example is provided by a massive scalar field with potential

$$ V(\phi) = \frac{1}{2} m^2 \phi^2, \quad (2.14) $$
where $m$ is the mass. It is assumed that the scalar field rolls towards the origin from large values of $|\varphi|$. It is a simple exercise for the reader to verify that the slow-roll conditions (2.10) and (2.11) are satisfied provided

$$|\varphi| > \frac{1}{\sqrt{12\pi m_{pl}}}.$$  

(2.15)

Values of $|\varphi|$ comparable or greater than $m_{pl}$ are also required in other realizations of large-field inflation. Hence, one may worry whether such a toy model can consistently be embedded in a realistic particle physics model, e.g. supergravity. In many such models $V(\varphi)$ receives supergravity-induced correction terms which destroy the flatness of the potential for $|\varphi| > m_{pl}$. However, as discussed e.g. in [28], if the flatness of the potential is protected by some symmetry, then it can survive inclusion of the correction terms. As will be discussed later, a value of $m \sim 10^{13}\text{GeV}$ is required in order to obtain the observed amplitude of density fluctuations. Hence, the configuration space of field values with $|\varphi| > m_{pl}$ but $V(\varphi) < m_{pl}^4$ is huge. It can also be verified that the slow-roll trajectory is a local attractor in field initial condition space [19], even including metric fluctuations at the perturbative level [29].

With two scalar fields it is possible to construct a class of models which combine some of the nice features of large-field inflation (large phase space of initial conditions yielding inflation) and of small-field inflation (inflation taking place at sub-Planckian field values). These are models of hybrid inflation [30]. To give a prototypical example, consider two scalar fields $\varphi$ and $\chi$ with a potential

$$V(\varphi, \chi) = \frac{1}{4\lambda_\chi}(\chi^2 - \sigma^2)^2 + \frac{1}{2}m^2\varphi^2 - \frac{1}{2}g^2\varphi^2\chi^2.$$  

(2.16)

In the absence of thermal equilibrium, it is natural to assume that $|\varphi|$ begins at large values, values for which the effective mass of $\chi$ is positive and hence $\chi$ begins at $\chi = 0$. The parameters in the potential (2.16) are now chosen such that $\varphi$ is slowly rolling for values of $|\varphi|$ somewhat smaller than $m_{pl}$, but that the potential energy for these field values is dominated by the first term on the right-hand side of (2.16). The reader can easily verify that for this model it is no longer required to have values of $|\varphi|$ greater than $m_{pl}$ in order to obtain slow-rolling. The field $\varphi$ is slowly rolling whereas the potential energy is determined by the contribution from $\chi$. Once $|\varphi|$ drops to the value

$$|\varphi_c| = \frac{\sqrt{\lambda_\chi}}{g}\sigma$$  

(2.17)

the configuration $\chi = 0$ becomes unstable and decays to its ground state $|\chi| = \sigma$, yielding a graceful exit from inflation. Since in this example the ground state of $\chi$ is not unique, there is the possibility of the formation of topological defects at the end of inflation (see [4, 5, 6] for reviews of topological defects in cosmology, and the lectures by Polchinski [31] for a discussion of how this scenario arises in brane inflation models).

After the slow-roll conditions break down, the period of inflation ends, and the inflaton begins to oscillate around its ground state. Due to couplings of $\varphi$ to other matter fields, the energy of the universe, which at the end of the period of inflation is stored completely in $\varphi$, gets transferred to the matter fields of the particle physics Standard Model. Initially, the energy transfer was described

---

8Note that the slow-roll conditions (2.10) and (2.11) were derived assuming that $H$ is given by the contribution of $\varphi$ to $V$ which is not the case here.
perturbatively [32, 33]. Later, it was realized [34, 35, 36, 37] that through a parametric resonance instability, particles are very rapidly produced, leading to a fast energy transfer (“preheating”). The quanta later thermalize, and thereafter the universe evolves as described by SBB cosmology.

3. Theory of Cosmological Perturbations: A Short Review

After this review of inflationary cosmology (see e.g. [38] for a more complete recent review), we turn to the discussion of the main success of inflationary cosmology, namely the fact that it provides a causal mechanism for generating small amplitude inhomogeneities. The reader is referred to [39] for a comprehensive analysis of this theory of cosmological perturbations, and [40] for a pedagogical overview, from which most of this section is drawn.

First, we describe the Newtonian theory of cosmological perturbations, mainly to develop intuition for the main physical effects. The range of validity of the Newtonian analysis is restricted to sub-Hubble scales at late times. In the second subsection, we then summarize the general relativistic theory of fluctuations. Linear fluctuations can be trivially quantized. The quantization of cosmological perturbations is presented in Subsection 3.

3.1 Newtonian Theory of Cosmological Perturbations

The growth of density fluctuations is a consequence of the purely attractive nature of the gravitational force. Given an density excess \( \delta \rho \) localized about some point \( x \) in space. This inhomogeneity produces a force which attracts the surrounding matter towards \( x \). The magnitude of this force is proportional to \( \delta \rho \). Hence, in a non-expanding background, by Newton’s second law

\[
\dot{\delta \rho} \sim G \delta \rho, \tag{3.1}
\]

where \( G \) is Newton’s gravitational constant. Thus, an exponential growth of the fluctuations is induced.

If, as required by consistency in General Relativity, we consider density fluctuations in an expanding background, then the expansion of space leads to a friction term in (3.1). Hence, instead of an exponential instability to the development of fluctuations, the growth of fluctuations will be as a power of time. The main goal of the theory of cosmological perturbations is to determine how the power-law instability depends on the background cosmology and on the length scale of the fluctuations.

3.2 Perturbations about Minkowski Space-Time

To develop some physical intuition, we first consider the evolution of hydrodynamical matter fluctuations in a fixed non-expanding background.

In this context, matter is described by a perfect fluid, and gravity by the Newtonian gravitational potential \( \varphi \). The fluid variables are the energy density \( \rho \), the pressure \( p \), the fluid velocity \( \mathbf{v} \), and the entropy density \( S \). The basic hydrodynamical equations are

\[
\dot{\rho} + \nabla_p \cdot (\rho \mathbf{v}) = 0 \\
\dot{\mathbf{v}} + (\mathbf{v} \cdot \nabla_p) \mathbf{v} + \frac{1}{\rho} \nabla_p p + \nabla \varphi = 0
\]
\[ \nabla^2 \phi = 4\pi G \rho \]  
(3.2)

\[ \dot{S} + (v \cdot \nabla p) S = 0 \]

\[ p = p(\rho, S), \]

where the subscript \( p \) indicates that physical as opposed to comoving coordinates are used. The first equation is the continuity equation, the second is the Euler (force) equation, the third is the Poisson equation of Newtonian gravity, the fourth expresses entropy conservation, and the last describes the equation of state of matter. The derivative with respect to time is denoted by an over-dot.

The background is given by the background energy density \( \rho_0 \), the background pressure \( p_0 \), vanishing velocity, constant gravitational potential \( \phi_0 \) and constant entropy density \( S_0 \). Note that the background Poisson equation is not satisfied.

The equations for cosmological perturbations are obtained by perturbing the fluid variables about the background,

\[ \rho = \rho_0 + \delta \rho \]

\[ v = \delta v \]

\[ p = p_0 + \delta p \]

\[ \phi = \phi_0 + \delta \phi \]

\[ S = S_0 + \delta S, \]

where the fluctuating fields \( \delta \rho, \delta v, \delta p, \delta \phi \) and \( \delta S \) are functions of space and time, by inserting these expressions into the basic hydrodynamical equations (3.2), and by linearizing. After combining the resulting first order equations, we get the following differential equations for the energy density fluctuation \( \delta \rho \) and the entropy perturbation \( \delta S \)

\[ \delta \rho - c_s^2 \nabla^2 \delta \rho - 4\pi G \rho_0 \delta \rho = \sigma \nabla^2 \delta S \]

\[ \dot{\delta S} = 0, \]

(3.4)

where the variables \( c_s^2 \) and \( \sigma \) describe the equation of state

\[ \delta p = c_s^2 \delta \rho + \sigma \delta S \]

(3.5)

with

\[ c_s^2 = \left( \frac{\partial p}{\partial \rho} \right)_s \]

(3.6)

denoting the square of the speed of sound.

Since the equations are linear, we can work in Fourier space. Each Fourier component \( \delta \rho_k(t) \) of the fluctuation field \( \delta \rho(x,t) \)

\[ \delta \rho(x,t) = \int e^{i k \cdot x} \delta \rho_k(t) \]

(3.7)

evolves independently.

The fluctuations can be classified as follows: If \( \delta S \) vanishes, we have adiabatic fluctuations. If the \( \delta S \) is non-vanishing but \( \delta \rho = 0 \), we speak of an entropy fluctuation.

The first conclusions we can draw from the basic perturbation equations (3.4) are that

1) entropy fluctuations do not grow,
2) adiabatic fluctuations are time-dependent, and
3) entropy fluctuations seed an adiabatic mode.

Taking a closer look at the equation of motion for $\delta \rho$, we see that the third term on the left hand side represents the force due to gravity, a purely attractive force yielding an instability of flat space-time to the development of density fluctuations (as discussed earlier, see (3.1)). The second term on the left hand side of (3.4) represents a force due to the fluid pressure which tends to set up pressure waves. In the absence of entropy fluctuations, the evolution of $\delta \rho$ is governed by the combined action of both pressure and gravitational forces.

Restricting our attention to adiabatic fluctuations, we see from (3.4) that there is a critical wavelength, the Jeans length, whose wavenumber $k_J$ is (in physical coordinates) given by

$$k_J = \left( \frac{4\pi G \rho_0}{c_s^2} \right)^{1/2}. \tag{3.8}$$

Fluctuations with wavelength longer than the Jeans length ($k \ll k_J$) grow exponentially

$$\delta \rho_k(t) \sim e^{\omega_k t} \text{ with } \omega_k \sim 4(\pi G \rho_0)^{1/2} \tag{3.9}$$

whereas short wavelength modes ($k \gg k_J$) oscillate with frequency $\omega_k \sim c_s k$. Note that the value of the Jeans length depends on the equation of state of the background. For a background dominated by relativistic radiation, the Jeans length is large (of the order of the Hubble radius $H^{-1}(t)$), whereas for pressure-less matter it goes to zero.

Next, we study Newtonian cosmological fluctuations about an expanding background. In this case, the background equations are consistent (the non-vanishing average energy density leads to cosmological expansion). However, we are neglecting general relativistic effects (the fluctuations of the metric) which dominate on length scales larger than the Hubble radius $H^{-1}(t)$.

The background cosmological model is given by the energy density $\rho(t,x)$, the pressure $p(t,x)$, and the recessional velocity $v_0 = H(t) x_p$ where $x_p$ is the physical coordinate vector. The space- and time-dependent fluctuating fields are defined in analogy to the previous section:

$$\rho(t,x) = \rho_0(t) \left(1 + \delta \epsilon(t,x)\right)$$
$$v(t,x) = v_0(t,x) + \delta v(t,x)$$
$$p(t,x) = p_0(t) + \delta p(t,x), \tag{3.10}$$

where $\delta \epsilon$ is the fractional energy density perturbation (we are interested in the fractional rather than in the absolute energy density fluctuation!), and the pressure perturbation $\delta p$ is defined as in (3.5). In addition, there is the possibility of a non-vanishing entropy perturbation defined as in (3.3).

We now insert this ansatz into the basic hydrodynamical equations (3.2), linearize in the perturbation variables, and combine the first order differential equations for $\delta \epsilon$ and $\delta p$ into a single second order differential equation for $\delta \epsilon$. The result simplifies if we work in comoving coordinates $x$. After some algebra, we obtain the following equation which describes the time evolution of density fluctuations:

$$\ddot{\delta \epsilon} + 2H \dot{\delta \epsilon} - \frac{c_s^2}{a^2} \nabla^2 \delta \epsilon - 4\pi G \rho_0 \delta \epsilon = \frac{\sigma}{\rho_0 a^2} \delta S, \tag{3.11}$$
where $\nabla$ is the partial derivative vector with respect to comoving coordinates. In addition, we have the equation of entropy conservation

$$\delta S = 0.$$  \hspace{1cm} (3.12)

Comparing with the equations (3.4) obtained in the absence of an expanding background, we see that the only difference is the presence of a Hubble damping term in the equation for $\delta_\varepsilon$. This term will moderate the exponential instability of the background to long wavelength density fluctuations. In addition, it will lead to a damping of the oscillating solutions on short wavelengths.

More specifically, for physical wavenumbers $k_p \ll k_J$ (where $k_J$ is again given by (3.8)), and in a matter-dominated background cosmology, the general solution of (3.11) in the absence of any entropy fluctuations is given by

$$\delta_k(t) = c_1 t^{2/3} + c_2 t^{-1},$$  \hspace{1cm} (3.13)

where $c_1$ and $c_2$ are two constants determined by the initial conditions, and we have dropped the subscript $\varepsilon$ in expressions involving $\delta_\varepsilon$. There are two fundamental solutions, the first a growing mode with $\delta_k(t) \sim a(t)$, the second a decaying mode with $\delta_k(t) \sim t^{-1}$. On short wavelength, one obtains damped oscillatory motion:

$$\delta_k(t) \sim a^{-1/2}(t) \exp\left( \pm ik \int dt' a^{-1}(t') \right).$$  \hspace{1cm} (3.14)

Before going on to the relativistic theory of cosmological perturbations, we will pause to introduce terminology used in cosmology to describe the fluctuations.

Let us consider perturbations on a fixed comoving length scale given by a comoving wavenumber $k$. The corresponding physical length increases as $a(t)$. This is to be compared to the Hubble radius $H^{-1}(t)$ which scales as $t$ provided $a(t)$ grows as a power of $t$. In the late time Universe, $a(t) \sim t^{1/2}$ in the radiation-dominated phase (i.e. for $t < t_{eq}$), and $a(t) \sim t^{2/3}$ in the matter-dominated period ($t_{eq} < t < t_0$). Thus, at sufficiently early times, all comoving scales had a physical length larger than the Hubble radius. If we consider large cosmological scales (e.g. those corresponding to the observed CMB anisotropies or to galaxy clusters), the time $t_H(k)$ of “Hubble radius crossing” (when the physical length was equal to the Hubble radius) was in fact later than $t_{eq}$. The time of Hubble radius crossing plays an important role in the evolution of cosmological perturbations.

Cosmological fluctuations can be described either in position space or in momentum space. In position space, we compute the root mean square mass fluctuation $\delta M / M(k,t)$ in a sphere of radius $l = 2\pi/k$ at time $t$. A scale-invariant spectrum of fluctuations is defined by the relation

$$\frac{\delta M}{M}(k, t_H(k)) = \text{const}. \hspace{1cm} (3.15)$$

Such a spectrum was first suggested by Harrison [41] and Zeldovich [42] as a reasonable choice for the spectrum of cosmological fluctuations. The “spectral index” $n$ of cosmological fluctuations is defined by the relation

$$\left( \frac{\delta M}{M} \right)^2(k, t_H(k)) \sim k^{n-1}. \hspace{1cm} (3.16)$$

Thus, a scale-invariant spectrum corresponds to $n = 1$. 

13
To go to momentum space representation, the fractional spatial density contrast is expanded in a Fourier series:

\[ \delta(x,t) = \int d^3k \delta_\varepsilon(k,t)e^{ik\cdot x}. \]  

(3.17)

The power spectrum \( P_\delta \) of density fluctuations is defined by

\[ P_\delta(k) = \frac{1}{2\pi^2}k^3|\tilde{\delta}_\varepsilon(k)|^2, \]

(3.18)

where \( k \) is the magnitude of \( k \). For simplicity, the distribution of fluctuations is taken to be Gaussian so that the fluctuation amplitude only depends on \( k \).

The power spectrum of the gravitational potential \( \phi \) is defined by

\[ P_\phi(k) = \frac{1}{2\pi^2}k^3|\tilde{\delta}_\phi(k)|^2. \]

(3.19)

The two power spectra (3.18) and (3.19) are related by the Poisson equation (3.2)

\[ P_\phi(k) \sim k^{-4}P_\delta(k). \]

(3.20)

The condition of scale-invariance can be expressed in terms of the power spectrum evaluated at a fixed time. To obtain this condition, we first use the time dependence of the fractional density fluctuation from (3.13) to determine the mass fluctuations at a fixed time. We need to use the fact that the time of Hubble radius crossing is given by

\[ a_H(k)t_H(k)k^{-1} = \beta t_H(k), \]

(3.21)

where \( \beta = 2 \) or \( \beta = 3/2 \) in the radiation and matter dominated phases, respectively. Making use of (3.16) we find

\[ (\frac{\delta M}{M})^2(k,t) \sim k^{n+3}. \]

(3.22)

Since, for reasonable values of the index of the power spectrum, \( \delta M/M(k,t) \) is dominated by the Fourier modes with wavenumber \( k \), we find that (3.22) implies

\[ P_\delta(k) \sim k^{n+3}, \]

(3.23)

or, equivalently,

\[ P_\phi(k) \sim k^{n-1}. \]

(3.24)

### 3.3 Relativistic Theory of Cosmological Fluctuations

The Newtonian theory of cosmological fluctuations discussed in the previous section breaks down on scales larger than the Hubble radius because it neglects perturbations of the metric, and because on large scales the metric fluctuations dominate the dynamics.

To show why metric fluctuations are important on scales larger than the Hubble radius, we can use a “separate universe” argument. On such scales, one should be able to approximately describe the evolution of the space-time by applying the first FRW equation (2.1) of homogeneous and isotropic cosmology to the local Universe (this approximation is made more rigorous in [43]). Based on this equation, a large-scale fluctuation of the energy density will lead to a fluctuation
(“$\delta a$”) of the scale factor $a$ which grows in time, a manifestation of the self gravitational amplification of fluctuations on length scales $\lambda$ greater than the Hubble radius.

Let us now turn to the rigorous analysis of cosmological fluctuations in the context of general relativity, where both metric and matter inhomogeneities are taken into account. We will consider fluctuations about a homogeneous and isotropic background cosmology, given by the metric (1.3), which can be written in conformal time $\eta$ (defined by $dt = a(t)d\eta$) as

$$ds^2 = a(\eta)^2(d\eta^2 - dx^2).$$

(3.25)

The theory of cosmological perturbations is based on expanding the Einstein equations to linear order about the background metric. The theory was initially developed in pioneering works by Lifshitz [44]. Significant progress in the understanding of the physics of cosmological fluctuations was achieved by Bardeen [45] who realized the importance of subtracting gauge artifacts (see below) from the analysis (see also [46]). The following discussion is based on Part I of the comprehensive review article [39]. Other reviews - emphasizing different aspects or approaches - are [47, 48, 49, 50].

The first step in the analysis of metric fluctuations is to classify them according to their transformation properties under spatial rotations. There are scalar, vector and second rank tensor fluctuations. In linear theory, there is no coupling between the different fluctuation modes, and hence they evolve independently (for some subtleties in this classification, see [51]).

We begin by expanding the metric about the FRW background metric $g_{\mu\nu}^{(0)}$ given by (3.25):

$$g_{\mu\nu} = g_{\mu\nu}^{(0)} + \delta g_{\mu\nu}.$$  

(3.26)

The background metric depends only on time, whereas the metric fluctuations $\delta g_{\mu\nu}$ depend on both space and time. Since the metric is a symmetric tensor, there are at first sight 10 fluctuating degrees of freedom in $\delta g_{\mu\nu}$.

There are four degrees of freedom which correspond to scalar metric fluctuations (the only four ways of constructing a metric from scalar functions):

$$\delta g_{\mu\nu} = a^2 \begin{pmatrix} 2\phi & -B_j \\ -B_i & 2(\psi \delta_{ij} - E_{ij}) \end{pmatrix},$$  

(3.27)

where the four fluctuating degrees of freedom are denoted (following the notation of [39]) $\phi, B, E$, and $\psi$, a comma denotes the ordinary partial derivative (if we had included spatial curvature of the background metric, it would have been the covariant derivative with respect to the background spatial metric), and $\delta_{ij}$ is the Kronecker symbol.

There are four vector degrees of freedom of metric fluctuations, consisting of the four ways of constructing metric fluctuations from three vectors:

$$\delta g_{\mu\nu} = a^2 \begin{pmatrix} 0 & -S_i \\ -S_i & F_{ij} + F_{ji} \end{pmatrix},$$  

(3.28)

where $S_i$ and $F_i$ are two divergence-less vectors (for a vector with non-vanishing divergence, the divergence contributes to the scalar gravitational fluctuation modes).
Finally, there are two tensor modes which correspond to the two polarization states of gravitational waves:

\[ \delta g_{\mu \nu} = -a^2 \begin{pmatrix} 0 & 0 \\ 0 & h_{ij} \end{pmatrix}, \]  

where \( h_{ij} \) is trace-free and divergence-less

\[ h_i^i = h_j^j = 0. \]  

(3.29)

Gravitational waves do not couple at linear order to the matter fluctuations. Vector fluctuations decay in an expanding background cosmology and hence are not usually cosmologically important. Thus, the most important fluctuations, at least in inflationary cosmology, are the scalar metric fluctuations, the fluctuations which couple to matter inhomogeneities and which are the relativistic generalization of the Newtonian perturbations considered in the previous section.

The theory of cosmological perturbations is at first sight complicated by the issue of gauge invariance. The coordinates \( t, x \) of space-time carry no independent physical meaning. By performing a small-amplitude transformation of the space-time coordinates (called “gauge transformation” in the following), we can create “fictitious” fluctuations in a homogeneous and isotropic Universe. These modes are “gauge artifacts”.

In the following we take an “active” view of gauge transformation. Consider two space-time manifolds, one of them a homogeneous and isotropic Universe \( M_0 \), the other a physical Universe \( \mathcal{M} \) with inhomogeneities. A choice of coordinates can be considered as a mapping \( \mathcal{D} \) between the manifolds \( M_0 \) and \( \mathcal{M} \). A second mapping \( \mathcal{D}' \) will map the same point in \( M_0 \) into a different point in \( \mathcal{M} \). Using the inverse of these maps \( \mathcal{D} \) and \( \mathcal{D}' \), we can assign two different sets of coordinates to points in \( \mathcal{M} \).

Consider now a physical quantity \( Q \) (e.g. the Ricci scalar) on \( \mathcal{M} \), and the corresponding physical quantity \( Q^{(0)} \) on \( M_0 \). Then, in the first coordinate system given by the mapping \( \mathcal{D} \), the perturbation \( \delta Q \) of \( Q \) at the point \( p \in \mathcal{M} \) is defined by

\[ \delta Q(p) = Q(p) - Q^{(0)}(\mathcal{D}^{-1}(p)). \]  

(3.31)

In the second coordinate system given by \( \mathcal{D}' \), the perturbation is defined by

\[ \tilde{\delta} Q(p) = Q(p) - Q^{(0)}(\mathcal{D}'^{-1}(p)). \]  

(3.32)

The difference

\[ \Delta Q(p) = \tilde{\delta} Q(p) - \delta Q(p) \]  

(3.33)

is a gauge artifact and carries no physical significance.

Some of the metric perturbation degrees of freedom introduced in the first subsection are thus gauge artifacts. To isolate these, we must study how coordinate transformations act on the metric. There are four independent gauge degrees of freedom corresponding to the coordinate transformations

\[ x^\mu \rightarrow \tilde{x}^\mu = x^\mu + \xi^\mu. \]  

(3.34)

The time component \( \xi^0 \) of \( \xi^\mu \) leads to a scalar metric fluctuation. The spatial three vector \( \xi^i \) can be decomposed as

\[ \xi^i = \xi^i_{\text{tr}} + \gamma^{ij} \xi_j. \]  

(3.35)
(where $\gamma^i$ is the spatial background metric) into a transverse piece $\xi^i_{\tau}$ which has two degrees of freedom which yield vector perturbations, and the second term (given by the gradient of a scalar $\xi$) which gives a scalar fluctuation. Thus, there are two scalar gauge modes given by $\xi^0$ and $\xi$, and two vector modes given by the transverse three vector $\xi^i_{\tau}$. Thus, there remain two physical scalar and two vector fluctuation modes. The gravitational waves are gauge-invariant.

Let us now focus on how the scalar gauge transformations (i.e. the transformations given by $\xi^0$ and $\xi$) act on the scalar metric fluctuation variables $\phi, B, E$, and $\psi$. An immediate calculation yields:

$$
\tilde{\phi} = \phi - \frac{d'}{a} \xi^0 - (\xi^0)' \nonumber \\
\tilde{B} = B + \xi^0 - \xi' \nonumber \\
\tilde{E} = E - \xi \\
\tilde{\psi} = \psi + \frac{d'}{a} \xi^0 ,
$$

where a prime indicates the derivative with respect to conformal time $\eta$.

There are two approaches to deal with the gauge ambiguities. The first is to fix a gauge, i.e. to pick conditions on the coordinates which completely eliminate the gauge freedom, the second is to work with a basis of gauge-invariant variables.

If one wants to adopt the gauge-fixed approach, there are many different gauge choices. Note that the often used synchronous gauge determined by $\delta g^0_{\mu} = 0$ does not totally fix the gauge. A convenient system which completely fixes the coordinates is the so-called longitudinal or conformal Newtonian gauge defined by $B = E = 0$.

If one prefers a gauge-invariant approach, there are many choices of gauge-invariant variables. A convenient basis first introduced by [45] is the basis $\Phi, \Psi$ given by

$$
\Phi = \phi + \frac{1}{a} [(B - E')a]' \nonumber \\
\Psi = \psi - \frac{d'}{a} (B - E') . \nonumber
$$

The gauge-invariant variables $\Phi$ and $\Psi$ coincide with the corresponding diagonal metric perturbations $\phi$ and $\psi$ in longitudinal gauge.

Note that the variables defined in (3.37) are gauge-invariant only under linear space-time coordinate transformations. Beyond linear order, the structure of perturbation theory becomes much more involved. In fact, one can show [52] that the only fluctuation variables which are invariant under all coordinate transformations are perturbations of variables which are constant in the background space-time. Beyond linear order there is also mixing of scalar, vector and tensor modes.

To derive the equations of motion for the fluctuations, the starting point is the set of Einstein equations

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

where $G_{\mu\nu}$ is the Einstein tensor associated with the space-time metric $g_{\mu\nu}$, and $T_{\mu\nu}$ is the energy-momentum tensor of matter. We insert the ansatz for metric and matter perturbed about a FRW
background \((g^{(0)}_{\mu \nu}(\eta), \varphi_0(\eta))\):

\[
g_{\mu \nu}(x, \eta) = g^{(0)}_{\mu \nu}(\eta) + \delta g_{\mu \nu}(x, \eta) \tag{3.40}
\]

\[
\varphi(x, \eta) = \varphi_0(\eta) + \delta \varphi(x, \eta), \tag{3.41}
\]

(where we have for simplicity replaced general matter by a scalar matter field \(\varphi\)) and expand to linear order in the fluctuating fields, obtaining the following equations:

\[
\delta G_{\mu \nu} = 8\pi G \delta T_{\mu \nu}. \tag{3.42}
\]

In the above, \(\delta g_{\mu \nu}\) is the perturbation in the metric and \(\delta \varphi\) is the fluctuation of the matter field \(\varphi\).

In a gauge-fixed approach, one can start with the metric in longitudinal gauge

\[
d\tilde{s}^2 = a^2[(1 + 2\phi)d\eta^2 - (1 - 2\psi)|\eta_d x_i dx^i] \tag{3.43}
\]

and insert this ansatz into the general perturbation equations (3.42). This yields the following set of equations of motion:

\[
-3\mathpzc{H}(\mathpzc{H} \phi + \psi') + \nabla^2 \psi = 4\pi G a^2 \delta T^0_0 \quad (\mathpzc{H}\phi + \psi')_i = 4\pi G a^2 \delta T^0_i \tag{3.44}
\]

\[
[2\mathpzc{H} + \mathpzc{H}^2] \phi + \mathpzc{H} \phi' + \psi'' + 2\mathpzc{H} \psi' \delta_{ij} + \frac{1}{2}\nabla^2 D \delta_j - \frac{1}{2} \eta^{ik} D_{kj} = -4\pi G a^2 \delta T^j_i,
\]

where \(D \equiv \phi - \psi\) and \(\mathpzc{H} = \dot{a}/a\).

If no anisotropic stress is present in the matter at linear order in fluctuating fields, i.e. if \(\delta T^i_j = 0\) for \(i \neq j\), then the two metric fluctuation variables coincide, i.e. \(\phi = \psi\). This will be the case in most simple cosmological models, e.g. in theories with matter described by a set of scalar fields with canonical form of the action, and in the case of a perfect fluid with no anisotropic stress.

In the simple case in which matter described in terms of a single scalar field \(\varphi\), then in longitudinal gauge (3.44) reduce to the following set of equations of motion

\[
\nabla^2 \phi - 3\mathpzc{H} \phi' - (\mathpzc{H}' + 2\mathpzc{H}^2) \phi = 4\pi G (\phi_0' \delta \varphi + V' a^2 \delta \varphi) \quad \phi' + \mathpzc{H} \phi = 4\pi G \phi_0' \delta \varphi \quad \phi'' + 3\mathpzc{H} \phi' + (\mathpzc{H}' + 2\mathpzc{H}^2) \phi = 4\pi G (\phi_0' \delta \varphi - V' a^2 \delta \varphi), \tag{3.45}
\]

where \(V'\) denotes the derivative of \(V\) with respect to \(\varphi\). These equations can be combined to give the following second order differential equation for the relativistic potential \(\phi\):

\[
\phi'' + 2\left(\mathpzc{H} - \frac{\phi''_0}{\phi_0}\right) \phi' - \nabla^2 \phi + 2\left(\mathpzc{H}' - \mathpzc{H} \frac{\phi''_0}{\phi_0}\right) \phi = 0. \tag{3.46}
\]

This is the final result for the classical evolution of cosmological fluctuations.\(^9\)

\(^9\)Note that we have implicitly assumed that the background matter field is slowly rolling, as it does in slow-roll inflation. In the case that it is time-independent, then the leading metric fluctuations are quadratic in the matter inhomogeneities, as discussed in [53].
There are similarities between the above equation of motion for the relativistic perturbations and the equation (3.11) obtained in the Newtonian theory. The final term in (3.46) is the force due to gravity leading to the instability, the second to last term is the pressure force leading to oscillations (relativistic since we are considering matter to be a relativistic field), and the second term is the Hubble friction term. For each wavenumber there are two fundamental solutions. On small scales \( k > H \), the solutions correspond to damped oscillations, on large scales \( k < H \) the oscillations freeze out and the dynamics is governed by the gravitational force competing with the Hubble friction term. Note, in particular, how the Hubble radius naturally emerges as the scale where the nature of the fluctuating modes changes from oscillatory to frozen.

Considering the equation in a bit more detail, observe that if the equation of state of the background is independent of time (which will be the case if \( \mathcal{H}' = \phi'' = 0 \)), then in an expanding background, the dominant mode of (3.46) is constant, and the sub-dominant mode decays. If the equation of state is not constant, then the dominant mode is not constant in time. Specifically, at the end of inflation \( \mathcal{H}' < 0 \), and this leads to a growth of \( \phi \) (see below).

To study the quantitative implications of the equation of motion (3.46), it is convenient to introduce \([54, 55]\) the variable \( \zeta \) (which, up to correction term of the order \( \nabla^2 \phi \) which is unimportant for large-scale fluctuations, is equal to the curvature perturbation \( \mathcal{R} \) in comoving gauge \([56]\)) by

\[
\zeta \equiv \phi + \frac{2}{3} \frac{(H^{-1} \dot{\phi} + \phi)}{1 + w},
\]  

(3.47)

where

\[
w = \frac{p}{\rho},
\]  

(3.48)

characterizes the equation of state of matter. In terms of \( \zeta \), the equation of motion (3.46) takes on the form

\[
\frac{3}{2} \xi H (1 + w) = \mathcal{O}(\nabla^2 \phi).
\]  

(3.49)

On large scales, the right hand side of the equation is negligible, which leads to the conclusion that large-scale cosmological fluctuations satisfy

\[
\dot{\zeta} (1 + w) = 0.
\]  

(3.50)

This implies that \( \zeta \) is constant except possibly if \( 1 + w = 0 \) at some point in time during the cosmological evolution (which occurs during reheating in inflationary cosmology if the inflaton field undergoes oscillations - see \([57]\) and \([58, 59]\) for discussions of the consequences in single and double field inflationary models, respectively). In single matter field models it is indeed possible to show that \( \dot{\zeta} = 0 \) on super-Hubble scales independent of assumptions on the equation of state \([60, 61]\). This "conservation law" makes it easy to relate initial fluctuations to final fluctuations in inflationary cosmology, as will be illustrated in the following.

Consider an application to inflationary cosmology. We return to the space-time sketch of the evolution of fluctuations - see Figure (1) - and use the conservation law (3.50) - in the form \( \zeta = \text{const} \) on large scales - to relate the amplitude of \( \phi \) at initial Hubble radius crossing during the inflationary phase (at \( t = t_i(k) \)) with the amplitude at final Hubble radius crossing at late times (at
topics in cosmology

Robert Brandenberger

$t = t_f(k)$. Since both at early times and at late times $\dot{\phi} = 0$ on super-Hubble scales, as the equation of state is not changing, (3.50) and (3.47) lead to

$$\phi(t_f(k)) \simeq \frac{(1 + w)(t_f(k))}{(1 + w)(t_i(k))} \phi(t_i(k)).$$

(3.51)

If the initial values of the perturbations are known, then the above equation allows us to evaluate the fluctuation amplitude at the time $t_f(k)$ of re-entry into the Hubble radius.

The time-time perturbed Einstein equation (the first equation of (3.44)) relates the value of $\phi$ at the initial Hubble radius crossing to the amplitude of the fractional energy density fluctuations at that time. This, together with the fact that the amplitude of the scalar matter field quantum vacuum fluctuations is of the order $H$, yields

$$\phi(t_i(k)) \sim H \frac{V'(t_i(k))}{V}.$$  

(3.52)

In the late time radiation dominated phase, $w = 1/3$, whereas during slow-roll inflation

$$1 + w(t_i(k)) \simeq \frac{\dot{\phi}_0^2}{V}(t_i(k)).$$

(3.53)

Making, in addition, use of the slow roll conditions satisfied during the inflationary period

$$3H \dot{\phi}_0 \simeq -V',$$

$$H^2 \simeq \frac{8\pi G}{3} V,$$

(3.54)

we arrive at the final result

$$\phi(t_f(k)) \sim \frac{V^{3/2}}{\sqrt{m^3_{pl}}}(t_i(k)),$$

(3.55)

which gives the position space amplitude of cosmological fluctuations on a scale labelled by the comoving wavenumber $k$ at the time when the scale re-enters the Hubble radius at late times, a result first obtained in the case of the Starobinsky model [15] of inflation in [16], and later in the context of scalar field-driven inflation in [62, 63, 64, 54].

In the case of slow roll inflation, the right hand side of (3.55) is, to a first approximation, independent of $k$, and hence the resulting spectrum of fluctuations is nearly scale-invariant.

3.4 Quantum Theory of Cosmological Fluctuations

In many models of the very early Universe, in particular in inflationary cosmology, primordial inhomogeneities emerge from quantum vacuum fluctuations on microscopic scales (wavelengths smaller than the Hubble radius). The wavelength is then stretched relative to the Hubble radius, becomes larger than the Hubble radius at some time and the perturbation then propagates on super-Hubble scales until re-entering at late cosmological times. In the context of a Universe with a de Sitter phase, the quantum origin of cosmological fluctuations was first discussed in [16]. In an inflationary universe, it is easy to justify focusing attention on the quantum fluctuations: any classical fluctuations present at the beginning of inflation are red-shifted during the period of inflation, and will thus be irrelevant to scales probed in observations today. On small scales, a quantum vacuum remains.
To understand the role of the Hubble radius, consider the equation of a free scalar matter field \( \varphi \) on an unperturbed expanding background:

\[
\dddot{\varphi} + 3H \dot{\varphi} - \frac{\nabla^2}{a^2} \varphi = 0.
\] (3.56)

The second term on the left hand side of this equation leads to damping of \( \varphi \) with a characteristic decay rate given by \( H \). As a consequence, in the absence of the spatial gradient term, \( \dot{\varphi} \) would be of the order of magnitude \( H \varphi \). Thus, comparing the second and the third terms on the left hand side, we immediately see that the microscopic (spatial gradient) term dominates on length scales smaller than the Hubble radius, leading to oscillatory motion, whereas this term is negligible on scales larger than the Hubble radius, and the evolution of \( \varphi \) is determined primarily by gravity.

Note that in general cosmological models the Hubble radius is much smaller than the horizon (the forward light cone calculated from the initial time). In an inflationary universe, the horizon is larger by a factor of \( \exp(N) \), where \( N \) is the number of e-foldings of inflation. It is very important to realize this difference, a difference which is obscured in most articles on cosmology in which the term “horizon” is used when “Hubble radius” is meant. Note, in particular, that the homogeneous inflaton field contains causal information on super-Hubble but sub-horizon scales. Hence, it is completely consistent with causality [57] to have a microphysical process related to the background scalar matter field lead to exponential amplification of the amplitude of fluctuations during reheating on such scales, as it does in models in which entropy perturbations are present and not suppressed during inflation [58, 59].

To understand the generation and evolution of fluctuations in current models of the very early Universe, we need both Quantum Mechanics and General Relativity, i.e. quantum gravity. At first sight, we are thus faced with an intractable problem, since the theory of quantum gravity is not yet established. We are saved by the fact that today on large cosmological scales the fractional amplitude of the fluctuations is smaller than 1. Since gravity is a purely attractive force, the fluctuations had to have been - at least in the context of an eternally expanding background cosmology - very small in the early Universe. Thus, a linearized analysis of the fluctuations (about a classical cosmological background) is self-consistent.

From the classical theory of cosmological perturbations discussed in the previous subsection, it follows that the analysis of scalar metric inhomogeneities can be reduced - after extracting gauge artifacts - to the study of the evolution of a single fluctuating variable. Thus, the quantum theory of cosmological perturbations must be reducible to the quantum theory of a single free scalar field which we will denote by \( v \). Since the background in which this scalar field evolves is time-dependent, the mass of \( v \) will be time-dependent. The time-dependence of the mass will lead to quantum particle production over time if we start the evolution in the vacuum state for \( v \). As we will see, this quantum particle production corresponds to the development and growth of the cosmological fluctuations. The quantum theory of cosmological fluctuations provides a consistent framework to study both the generation and the evolution of metric perturbations.

In order to obtain the action for linearized cosmological perturbations, we expand the action to quadratic order in the fluctuating degrees of freedom. The linear terms cancel because the background is taken to satisfy the background equations of motion.
We begin with the Einstein-Hilbert action for gravity and the action of a scalar matter field

\[ S = \int d^4x\sqrt{-g}\left[ -\frac{1}{16\pi G}R + \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi) \right], \]  

(3.57)

where \( R \) is the Ricci curvature scalar.

The simplest way to proceed is to work in longitudinal gauge. The next step is to reduce the number of degrees of freedom. Since for scalar field matter there are no anisotropic stresses to linear order, the off-diagonal spatial Einstein equations force \( \psi = \phi \). The two remaining fluctuating variables \( \phi \) and \( \varphi \) are linked by the Einstein constraint equations since there cannot be matter fluctuations without induced metric fluctuations.

The two nontrivial tasks of the lengthy [39] computation of the quadratic piece of the action is to find out what combination of \( \varphi \) and \( \phi \) yields the variable \( v \) in terms of which the action has canonical kinetic term, and what the form of the time-dependent mass is. In the context of scalar field matter, the quantum theory of cosmological fluctuations was developed by Mukhanov [65, 66] and Sasaki [67]. The result is the following form of the action quadratic in the perturbations:

\[ S^{(2)} = \frac{1}{2} \int d^4x [v'^2 - v_jv_j + \frac{z''}{z}v^2], \]  

(3.58)

where the canonical variable \( v \) (the “Sasaki-Mukhanov variable” introduced in [66] - see also [68]) is given by

\[ v = a \left[ \delta \phi + \frac{\dot{\phi}}{H} \phi \right], \]  

(3.59)

and

\[ z = \frac{a\dot{\phi}}{H}. \]  

(3.60)

As long as the equation of state does not change over time \( \mathcal{H} \) and \( \phi_0' \) are proportional and hence

\[ z(\eta) \sim a(\eta). \]  

(3.61)

Note that the variable \( v \) is related to the curvature perturbation \( \mathcal{R} \) in comoving coordinates introduced in [56] and closely related to the variable \( \zeta \) used in [54, 55]:

\[ v = z\mathcal{R}. \]  

(3.62)

The equation of motion which follows from the action (3.58) is (in momentum space)

\[ v_k'' + k^2v_k - \frac{z''}{z}v_k = 0, \]  

(3.63)

where \( v_k \) is the \( k \)'th Fourier mode of \( v \). As a consequence of (3.61), the tachyonic mass term in the above equation is given by the Hubble scale

\[ k_H^2 \equiv \frac{z''}{z} \simeq H^2. \]  

(3.64)

Thus, it immediately follows from (3.63) that on small length scales, i.e. for \( k > k_H \), the solutions for \( v_k \) are constant amplitude oscillations. These oscillations freeze out at Hubble radius crossing, i.e. when \( k = k_H \). On longer scales (\( k \ll k_H \)), the solutions for \( v_k \) increase as \( z \):

\[ v_k \sim z, \quad k \ll k_H. \]  

(3.65)
The state of the fluctuations becomes a squeezed quantum state.

Given the action (3.58), the quantization of the cosmological perturbations can be performed by canonical quantization (in the same way that a scalar matter field on a fixed cosmological background is quantized [69]).

The final step in the quantum theory of cosmological perturbations is to specify an initial state. Since in inflationary cosmology all pre-existing classical fluctuations are red-shifted by the accelerated expansion of space, one usually assumes (we will return to a criticism of this point when discussing the trans-Planckian problem of inflationary cosmology) that the field $v$ starts out at the initial time $t_i$ mode by mode in its vacuum state. Two questions immediately emerge: what is the initial time $t_i$, and which of the many possible vacuum states should be chosen. It is usually assumed that since the fluctuations only oscillate on sub-Hubble scales, the choice of the initial time is not important, as long as it is earlier than the time when scales of cosmological interest today cross the Hubble radius during the inflationary phase. The state is usually taken to be the Bunch-Davies vacuum (see e.g. [69]), since this state is empty of particles at $t_i$ in the coordinate frame determined by the FRW coordinates (see e.g. [70] for a discussion of this point), and since the Bunch-Davies state is a local attractor in the space of initial states in an expanding background (see e.g. [71]). Thus, we choose the initial conditions

$$v_k(\eta_i) = \frac{1}{\sqrt{2\omega_k}}$$
$$v'_k(\eta_i) = \frac{\sqrt{\omega_k}}{\sqrt{2}}$$

where here $\omega_k = k$, and $\eta_i$ is the conformal time corresponding to the physical time $t_i$.

Let us briefly summarize the quantum theory of cosmological perturbations. In the linearized theory, fluctuations are set up at some initial time $t_i$ mode by mode in their vacuum state. While the wavelength is smaller than the Hubble radius, the state undergoes quantum vacuum fluctuations. The accelerated expansion of the background redshifts the length scale beyond the Hubble radius. The fluctuations freeze out when the length scale is equal to the Hubble radius. On larger scales, the amplitude of $v_k$ increases as the scale factor. This corresponds to the squeezing of the quantum state present at Hubble radius crossing (in terms of classical general relativity, it is self-gravity which leads to this growth of fluctuations). As discussed e.g. in [72], the squeezing of the quantum vacuum state sets up the classical correlations in the wave function of the fluctuations which are an essential ingredient in the classicalization of the perturbations. Squeezing also leads to the phase coherence of fluctuations on all scales, which in turn is responsible for producing the oscillations in the CMB angular power spectrum, as realized long before the advent of inflationary cosmology in [73, 74].

We end this subsection with a calculation of the spectrum of curvature fluctuations in inflationary cosmology.

We need to compute the power spectrum $P_R(k)$ of the curvature fluctuation $R$ defined in (3.62). The idea in calculating the power spectrum at a late time $t$ is to first relate the power spectrum via the growth rate (3.65) of $v$ on super-Hubble scales to the power spectrum at the time $t_H(k)$ of Hubble radius crossing, and to then use the constancy of the amplitude of $v$ on sub-Hubble
scales to relate it to the initial conditions (3.66). Thus
\[
2\pi^2 \mathcal{P}_R(k,t) \equiv k^3 |\mathcal{R}_k|^2(t) = k^3 z^{-2}(t) |v_k(t)|^2
\]
(3.67)
\[
= k^3 z^{-2}(t) \left( \frac{z(t)}{z(t_H(k))} \right)^2 |v_k(t_H(k))|^2
\]
\[
= k^3 z^{-2}(t_H(k)) |v_k(t_H(k))|^2
\]
\[
\sim k^3 a^{-2} |v_k(t_H(k))|^2,
\]
where in the final step we have used (3.61) and the constancy of the amplitude of $v$ on sub-Hubble scales. Making use of the condition $a^{-1}(t_H(k))k = H$ for Hubble radius crossing, and of the initial conditions (3.66), we immediately see that
\[
\mathcal{P}_R(k,t) \sim k^3 k^{-2} k^{-1} H^2,
\]
and that thus a scale invariant power spectrum with amplitude proportional to $H^2$ results.

The quantization of gravitational waves parallels the quantization of scalar metric fluctuations, but is more simple because there are no gauge ambiguities. The starting point is the action (3.57), into which we insert the metric which corresponds to a classical cosmological background plus tensor metric fluctuations:
\[
ds^2 = a^2(\eta) \left[ d\eta^2 - (\delta_{ij} + h_{ij})dx^i dx^j \right],
\]
(3.69)
where the second rank tensor $h_{ij}(\eta, x)$ represents the gravitational waves, and in turn can be decomposed as
\[
h_{ij}(\eta, x) = h_+ (\eta, x) e_{ij}^+ + h_\times (\eta, x) e_{ij}^\times
\]
(3.70)
into the two polarization states. Here, $e_{ij}^+$ and $e_{ij}^\times$ are two fixed polarization tensors, and $h_+$ and $h_\times$ are the two coefficient functions.

To quadratic order in the fluctuating fields, the action consists of separate terms involving $h_+$ and $h_\times$. Each term is of the form
\[
S^{(2)} = \int d^4x a^2 \frac{d^2}{2} [h'^2 - (\nabla h)^2],
\]
(3.71)
leading to the equation of motion
\[
h_+'' + 2 \frac{d'}{d} h_+' + k^2 h_+ = 0.
\]
(3.72)
The variable in terms of which the action (3.71) has canonical kinetic term is
\[
\mu_k \equiv ah_k,
\]
(3.73)
and its equation of motion is
\[
\mu_k'' + \left( k^2 - \frac{d''}{d} \right) \mu_k = 0.
\]
(3.74)
This equation is very similar to the corresponding equation (3.63) for scalar gravitational inhomogeneities, except that in the mass term the scale factor $a(\eta)$ replaces $z(\eta)$, which leads to a very
different evolution of scalar and tensor modes during the reheating phase in inflationary cosmology during which the equation of state of the background matter changes dramatically.

Based on the above discussion we have the following theory for the generation and evolution of gravitational waves in an accelerating Universe (first developed by Grishchuk [75]): waves exist as quantum vacuum fluctuations at the initial time on all scales. They oscillate until the length scale crosses the Hubble radius. At that point, the oscillations freeze out and the quantum state of gravitational waves begins to be squeezed in the sense that

\[ \mu_k(\eta) \sim a(\eta), \]  
(3.75)

which, from (3.73) corresponds to constant amplitude of \( h_k \). The squeezing of the vacuum state leads to the emergence of classical properties of this state, as in the case of scalar metric fluctuations.

4. Towards a New Paradigm: String Gas Cosmology

In spite of the phenomenological successes of inflationary cosmology in addressing some key problems of standard cosmology and in providing a predictive theory of structure formation, current models of inflation are faced with some key conceptual problems which motivate the search for a new early universe paradigm based on new fundamental physics, string theory being the most promising candidate. Below, we first list some of the key problems for inflation, and then discuss a toy model designed to explore possible cosmological consequences of some of the key new features of string theory.

4.1 Conceptual Problems of Scalar Field-Driven Inflation

Nature of the Inflaton

In the context of General Relativity as the theory of space-time, matter with an equation of state \( p \simeq -\rho \) is required in order to obtain almost exponential expansion of space. If we describe matter in terms of fields with canonical kinetic terms, a scalar field is required since in the context of usual renormalizable field theories it is only for scalar fields that a potential energy function in the Lagrangian is allowed, and of all energy terms only the potential energy can yield the required equation of state.

In order for scalar fields to generate a period of cosmological inflation, the potential energy needs to dominate over the kinetic and spatial gradient energies. It is generally assumed that spatial gradient terms can be neglected. This is, however, not true in general. Next, assuming a homogeneous field configuration, we must ensure that the potential energy dominates over the kinetic energy. This leads to the first “slow-roll” condition. Requiring the period of inflation to last sufficiently long leads to a second slow-roll condition, namely that the \( \ddot{\phi} \) term in the Klein-Gordon equation for the inflaton \( \phi \) be negligible. Scalar fields charged with respect to the Standard Model symmetry groups do not satisfy the slow-roll conditions.

Assuming that both slow-roll conditions hold, one obtains a “slow-roll trajectory” in the phase space of homogeneous \( \phi \) configurations. In large-field inflation models such as “chaotic inflation” [27] and “hybrid inflation” [30], the slow-roll trajectory is a local attractor in initial condition space.
[19] (even when linearized metric perturbations are taken into account [29]), whereas this is not the case [20] in small-field models such as “new inflation” [24]. As shown in [26], this leads to problems for some models of inflation which have recently been proposed in the context of string theory. To address this problem, it has been proposed that inflation may be future- eternal [76] and that it is hence sufficient that there be some configurations in initial condition space which give rise to inflation within one Hubble patch, inflation being then self-sustaining into the future. However, one must still ensure that slow-roll inflation can locally be satisfied.

Many models of particle physics beyond the Standard Model contain a plethora of new scalar fields. One of the most conservative extensions of the Standard Model is the MSSM, the “Minimal Supersymmetric Standard Model”. According to a recent study, among the many scalar fields in this model, only a hand-full can be candidates for a slow-roll inflaton, and even then very special initial conditions are required [77]. The situation in supergravity and superstring-inspired field theories may be more optimistic, but the issues are not settled (see e.g. [7, 8, 9] for recent reviews).

**Hierarchy Problem**

Assuming for the sake of argument that a successful model of slow-roll inflation has been found, one must still build in a hierarchy into the field theory model in order to obtain an acceptable amplitude of the density fluctuations (this is sometimes also called the “amplitude problem”). Unless this hierarchy is observed, the density fluctuations will be too large and the model is observationally ruled out.

In a wide class of inflationary models, obtaining the correct amplitude requires the introduction of a hierarchy in scales, namely [78]

\[ \frac{V(\phi)}{\Delta \phi^4} \leq 10^{-12}, \]  

where \( \Delta \phi \) is the change in the inflaton field during one Hubble expansion time (during inflation), and \( V(\phi) \) is the potential energy during inflation.

This problem should be contrasted with the success of topological defect models (see e.g. [4, 5, 6] for reviews) in predicting the right order of magnitude of density fluctuations without introducing a new scale of physics. The GUT scale as the scale of the symmetry breaking phase transition (which produces the defects) yields the correct magnitude of the spectrum of density fluctuations [79]. Topological defects, however, cannot be the prime mechanism for the origin of fluctuations since they do not give rise to coherent adiabatic fluctuations and hence fail to yield acoustic oscillations in the angular power spectrum of the CMB anisotropies [80].

**Trans-Planckian Problem**

A more serious problem is the “trans-Planckian problem” [81]. Returning to the space-time diagram of Figure 1, we can immediately deduce that, provided that the period of inflation lasted sufficiently long (for GUT scale inflation the number is about 70 e-foldings), then all scales inside of the Hubble radius today started out with a physical wavelength smaller than the Planck scale at the beginning of inflation. Now, the theory of cosmological perturbations is based on Einstein’s theory of General Relativity coupled to a simple semi-classical description of matter. It is clear that these building blocks of the theory are inapplicable on scales comparable and smaller than the Planck scale. Thus, the key successful prediction of inflation (the theory of the origin of fluctuations) is based on suspect calculations since new physics must enter into a correct computation.
of the spectrum of cosmological perturbations. The key question is as to whether the predictions obtained using the current theory are sensitive to the specifics of the unknown theory which takes over on small scales.

One approach to study the sensitivity of the usual predictions of inflationary cosmology to the unknown physics on trans-Planckian scales is to study toy models of ultraviolet physics which allow explicit calculations. The first approach which was used [82, 83] is to replace the usual linear dispersion relation for the Fourier modes of the fluctuations by a modified dispersion relation, a dispersion relation which is linear for physical wavenumbers smaller than the scale of new physics, but deviates on larger scales. Such dispersion relations were used previously to test the sensitivity of black hole radiation on the unknown physics of the ultra-violet [84, 85]. It was found [82] that if the evolution of modes on the trans-Planckian scales is non-adiabatic, then substantial deviations of the spectrum of fluctuations from the usual results are possible. Non-adiabatic evolution turns an initial state minimizing the energy density into a state which is excited once the wavelength becomes larger than the cutoff scale. Back-reaction effects of these excitations may limit the magnitude of the trans-Planckian effects, but - based on our recent study [86] - not to the extent initially expected [87, 88].

From the point of view of fundamental physics, the trans-Planckian problem is not a problem. Rather, it yields a window of opportunity to probe new fundamental physics in current and future observations, even if the scale of the new fundamental physics is close to the Planck scale. The point is that if the universe in fact underwent a period of inflation, then trans-Planckian physics leaves an imprint on the spectrum of fluctuations. The exponential expansion of space amplifies the wavelength of the perturbations to observable scales. At the present time, it is our ignorance about quantum gravity which prevents us from making any specific predictions. For example, we do not understand string theory in time-dependent backgrounds sufficiently well to be able to at this time make any predictions for observations.

Singularity Problem

The next problem is the “singularity problem”. This problem, one of the key problems of Standard Cosmology, has not been resolved in models of scalar field-driven inflation.

As follows from the Penrose-Hawking singularity theorems of General Relativity (see e.g. [89] for a textbook discussion), an initial cosmological singularity is unavoidable if space-time is described in terms of General Relativity, and if the matter sources obey the weak energy conditions. Recently, the singularity theorems have been generalized to apply to Einstein gravity coupled to scalar field matter, i.e. to scalar field-driven inflationary cosmology [90]. It is shown that in this context, a past singularity at some point in space is unavoidable.

In the same way that the appearance of an initial singularity in Standard Cosmology told us that Standard Cosmology cannot be the correct description of the very early universe, the appearance of an initial singularity in current models of inflation tell us that inflationary cosmology cannot yield the correct description of the very, very early universe. At sufficiently high densities, a new description will take over. In the same way that inflationary cosmology contains late-time standard cosmology, it is possible that the new cosmology will contain, at later times, inflationary cosmology. However, one should keep an open mind to the possibility that the new cosmology will connect to present observations via a route which does not contain inflation, a possibility explored later in this section.
Breakdown of Validity of Einstein Gravity

The Achilles heel of scalar field-driven inflationary cosmology is, however, the use of intuition from Einstein gravity at energy scales not far removed from the Planck and string scales, scales where correction terms to the Einstein-Hilbert term in the gravitational action dominate and where intuition based on applying the Einstein equations break down (see also [91] for arguments along these lines).

All approaches to quantum gravity predict correction terms in the action which dominate at energies close to the Planck scale - in some cases in fact even much lower. Semiclassical gravity leads to higher curvature terms, and may (see e.g. [92, 93]) lead to bouncing cosmologies without a singularity). Loop quantum cosmology leads to similar modifications of early universe cosmology (see e.g. [94] for a recent review). String theory, the theory we will focus on in the following sections, has a maximal temperature for a string gas in thermal equilibrium [95], which may lead to an almost static phase in the early universe - the Hagedorn phase [96].

Common to all of these approaches to quantum gravity corrections to early universe cosmology is the fact that a transition from a contracting (or quasi-static) early universe phase to the rapidly expanding radiation phase of standard cosmology can occur without violating the usual energy conditions for matter. In particular, it is possible (as is predicted by the string gas cosmology model discussed below) that the universe in an early high temperature phase is almost static. This may be a common feature to a large class of models which resolve the cosmological singularity.

Closely related to the above is the “cosmological constant problem” for inflationary cosmology. We know from observations that the large quantum vacuum energy of field theories does not gravitate today. However, to obtain a period of inflation one is using precisely the part of the energy-momentum tensor of the inflaton field which looks like the vacuum energy. In the absence of a convincing solution of the cosmological constant problem it is unclear whether scalar field-driven inflation is robust, i.e. whether the mechanism which renders the quantum vacuum energy gravitationally inert today will not also prevent the vacuum energy from gravitating during the period of slow-rolling of the inflaton field\(^\text{10}\).

4.2 String Gas Cosmology

An immediate problem which arises when trying to connect string theory with cosmology is the dimensionality problem. Superstring theory is perturbatively consistent only in ten space-time dimensions, but we only see three large spatial dimensions. The original approach to addressing this problem was to assume that the six extra dimensions are compactified on a very small space which cannot be probed with our available energies. However, from the point of view of cosmology, it is quite unsatisfactory not to be able to understand why it is precisely three dimensions which are not compactified and why the compact dimensions are stable. Brane world cosmology [98] provides another approach to this problem: it assumes that we live on a three-dimensional brane embedded in a large nine-dimensional space. Once again, a cosmologically satisfactory theory should explain why it is likely that we will end up exactly on a three-dimensional brane (for some interesting work addressing this issue see [99, 100, 101]).

\(^\text{10}\)Note that the approach to addressing the cosmological constant problem making use of the gravitational back-reaction of long range fluctuations (see [97] for a summary of this approach) does not prevent a long period of inflation in the early universe.
Finding a natural solution to the dimensionality problem is thus one of the key challenges for superstring cosmology. This challenge has various aspects. First, there must be a mechanism which singles out three dimensions as the number of spatial dimensions we live in. Second, the moduli fields which describe the volume and the shape of the unobserved dimensions must be stabilized (any strong time-dependence of these fields would lead to serious phenomenological constraints). This is the moduli problem for superstring cosmology. As mentioned above, solving the singularity problem is another of the main challenges. These are the three problems which string gas cosmology [96, 102, 103] explicitly addresses at the present level of development.

In the absence of a non-perturbative formulation of string theory, the approach to string cosmology which we have suggested, string gas cosmology [96, 102, 103] (see also [104] for early work, and [105, 106, 107, 108] for reviews), is to focus on symmetries and degrees of freedom which are new to string theory (compared to point particle theories) and which will be part of a non-perturbative string theory, and to use them to develop a new cosmology. The symmetry we make use of is T-duality, and the new degrees of freedom are string winding modes and string oscillatory modes.

We take all spatial directions to be toroidal, with $R$ denoting the radius of the torus. Strings have three types of states: momentum modes which represent the center of mass motion of the string, oscillatory modes which represent the fluctuations of the strings, and winding modes counting the number of times a string wraps the torus. Both oscillatory and winding states are special to strings. Point particle theories do not contain these modes.

The energy of an oscillatory mode is independent of $R$, momentum mode energies are quantized in units of $1/R$, i.e.

$$E_n = n \frac{1}{R}, \quad (4.2)$$

whereas the winding mode energies are quantized in units of $R$, i.e.

$$E_m = mR, \quad (4.3)$$

where both $n$ and $m$ are integers. The energy of oscillatory modes does not depend on $R$.

The T-duality symmetry is the invariance of the spectrum of string states under the change

$$R \rightarrow \frac{1}{R}, \quad (4.4)$$

in the radius of the torus (in units of the string length $l_s$). Under such a change, the energy spectrum of string states is not modified if winding and momentum quantum numbers are interchanged

$$(n,m) \rightarrow (m,n). \quad (4.5)$$

The string vertex operators are consistent with this symmetry, and thus T-duality is a symmetry of perturbative string theory. Postulating that T-duality extends to non-perturbative string theory leads [109] to the need of adding D-branes to the list of fundamental objects in string theory. With this addition, T-duality is expected to be a symmetry of non-perturbative string theory. Specifically, T-duality will take a spectrum of stable Type IIA branes and map it into a corresponding spectrum of stable Type IIB branes with identical masses [110].

Since the number of string oscillatory modes increases exponentially as the string mode energy increases, there is a maximal temperature of a gas of strings in thermal equilibrium, the Hagedorn
Figure 4: Sketch (based on the analysis of [96]) of the evolution of temperature $T$ as a function of the radius $R$ of space of a gas of strings in thermal equilibrium. The top curve is characterized by an entropy higher than the bottom curve, and leads to a longer region of Hagedorn behaviour.

temperature $T_H$ [95]. If we imagine taking a box of strings and compressing it, the temperature will never exceed $T_H$. In fact, as the radius $R$ decreases below the string radius, the temperature will start to decrease, obeying the duality relation [96]

$$T(R) = T(1/R). \quad (4.6)$$

This argument shows that string theory has the potential of taming singularities in physical observables. Figure 4 provides a sketch of how the temperature $T$ changes as a function of $R$.

If we imagine that there is a dynamical principle that tells us how $R$ evolves in time, then Figure 2 can be interpreted as depicting how the temperature changes as a function of time. If $R$ is a monotonic function of time, then two interesting possibilities for cosmology emerge. If $\ln R$ decreases to zero at some fixed time (which without loss of generality we can call $t = 0$), and continues to decrease, we obtain a temperature profile which is symmetric with respect to $t = 0$ and which (since small $R$ is physically equivalent to large $R$) represents a bouncing cosmology (see [111] for a concrete recent realization of this scenario). If, on the other hand, it takes an infinite amount of time to reach $R = 0$, an emergent universe scenario [112] is realized.

It is important to realize that in both of the cosmological scenarios which, as argued above, seem to follow from string theory symmetry considerations alone, a large energy density does not lead to rapid expansion in the Hagedorn phase, in spite of the fact that the matter sources we are considering (namely a gas of strings) obey all of the usual energy conditions discussed e.g. in [89]). These considerations are telling us that intuition drawn from Einstein gravity will give us a completely incorrect picture of the early universe.

Any physical theory requires both a specification of the equations of motion and of the initial conditions. We assume that the universe starts out small and hot. For simplicity, we take space to be toroidal, with radii in all spatial directions given by the string scale. We assume that the initial energy density is very high, with an effective temperature which is close to the Hagedorn temperature, the maximal temperature of perturbative string theory.
In this context, it was argued [96] that in order for spatial sections to become large, the winding modes need to decay. This decay, at least on a background with stable one cycles such as a torus, is only possible if two winding modes meet and annihilate. Since string world sheets have measure zero probability for intersecting in more than four space-time dimensions, winding modes can annihilate only in three spatial dimensions (see, however, the recent caveats to this conclusion based on the work of [113, 114]). Thus, only three spatial dimensions can become large, hence explaining the observed dimensionality of space-time. As was shown later [103], adding branes to the system does not change these conclusions since at later times the strings dominate the cosmological dynamics. Note that in the three dimensions which are becoming large there is a natural mechanism of isotropization as long as some winding modes persist [115].

Some of the above heuristic arguments can be put on a more firm mathematical basis, albeit in the context of a toy model, a model consisting of a classical background coupled to a gas of strings. From the point of view of rigorous string theory, this separation between classical background and stringy matter is not satisfactory when dealing with very early times when the typical length scale might be the string scale. However, in the absence of a non-perturbative formulation of string theory, at the present time we are forced to make this separation. Note that this separation between classical background geometry and string matter is common to all current approaches to string cosmology.

The background is described by dilaton gravity. The dilaton must be included since it arises in string theory at the same level as the graviton, and also (in the context of string gas cosmology) because it and not Einstein gravity is consistent with the T-duality symmetry. Note, however, that the background dynamics inevitably drives the system into a parameter region where the dilaton is strongly coupled and hence beyond the region of validity of the approximations made.

The action for dilaton gravity coupled to a matter action $S_m$ is

$$S = \frac{1}{2\kappa^2} \int d^{10}x \sqrt{-g} e^{-2\phi} \left[ R + 4 \partial^\mu \phi \partial_\mu \phi \right] + S_m,$$

(4.7)

where $g$ is the determinant of the metric, $R$ is the Ricci scalar, $\phi$ is the dilaton, and $\kappa$ is the reduced gravitational constant in ten dimensions. The metric appearing in the above action is the metric in the string frame.

For a homogeneous and isotropic metric

$$ds^2 = dt^2 - a(t)^2 dx^2,$$

(4.8)

the resulting equations of motion in the string frame are [102] (see also [116])

$$-d\dot{\lambda}^2 + \phi^2 = e^\phi E$$

(4.9)

$$\dot{\lambda} - \phi \dot{\lambda} = \frac{1}{2} e^\phi P$$

(4.10)

$$\phi - d\dot{\lambda}^2 = \frac{1}{2} e^\phi E,$$

(4.11)

where $E$ and $P$ denote the total energy and pressure, respectively, $d$ is the number of spatial dimensions, and we have introduced the logarithm of the scale factor

$$\lambda(t) = \log(a(t)).$$

(4.12)
and the rescaled dilaton

\[ \varphi = 2\phi - d\lambda. \]  (4.13)

The second of these equations indicates that a gas of strings containing both stable winding and momentum modes will lead to the stabilization of the radius of the torus: windings prevent expansion, momenta prevent the contraction. The right hand side of the equation can be interpreted as resulting from a confining potential for the scale factor. One of the key issues when dealing with theories with extra dimensions is the question of how the size and shape moduli of the extra-dimensional spaces are stabilized. String gas cosmology provides a simple and string-specific mechanism to stabilize most of these moduli. This topic will not be reviewed here (see [107, 108] for recent reviews). The outstanding issue is how to stabilize the dilaton.

Note that the dilaton is evolving at the time when the radius of the torus is at the minimum of its potential. For the branch of solutions we are considering, the dilaton is increasing as we go into the past. At some point, therefore, it becomes greater than zero. At this point, we enter the region of strong coupling. As already discussed in [117], a different dynamical framework is required to analyze this phase. In particular, the fundamental strings are no longer the lightest degrees of freedom. We will call this phase the "strongly coupled Hagedorn phase" [118] for which we lack an analytical description. Since the energy density in this phase is of the string scale, the background equations should also be very different from the dilaton gravity equations used above. In the following, we assume that the dilaton is frozen in the strongly coupled Hagedorn phase. This could be a consequence of S-duality (see e.g. [119]).

4.3 String Gas Cosmology and Structure Formation

The following are key aspects of the string gas cosmology background which emerge from the previous discussion. First, in thermal equilibrium at the string scale (\( R \approx l_s \)), the self-dual radius, the number of winding and momentum modes are equal. Since winding and momentum modes give an opposite contribution to the pressure, the pressure of the string gas in thermal equilibrium at the self-dual radius will vanish. From the dilaton gravity equations of motion (4.9 - 4.11) it then follows that a static phase \( \lambda = 0 \) will be a fixed point of the dynamical system. This phase is the Hagedorn phase.

On the other hand, for large values of \( R \) in thermal equilibrium the energy will be exclusively in momentum modes. These act as usual radiation. Inserting the radiative equation of state into the above equations (4.9 - 4.11) it follows that the source in the dilaton equation of motion vanishes and the dilaton approaches a constant as a consequence of the Hubble damping term in its equation of motion. Consequently, the scale factor expands as in the usual radiation-dominated universe. The transition between the Hagedorn phase and the radiation-dominated phase with fixed dilaton is achieved via the annihilation of winding modes, as studied in detail in [120]. The main point is that, starting in a Hagedorn phase, there will be a smooth transition to the radiation-dominated phase of standard cosmology with fixed dilaton.

Our new cosmological background is obtained by following our currently observed universe into the past according to the string gas cosmology equations. The radiation phase of standard
cosmology is unchanged. In particular, the dilaton is fixed in this phase \(^{11}\). However, as the temperature of the radiation bath approaches the Hagedorn temperature, the equation of state of string gas matter changes. The equation of state parameter \(w = P/E\) decreases towards a pressureless state and the string frame metric becomes static. Note that, in order for the present size of the universe to be larger than our current Hubble radius, the size of the spatial sections in the Hagedorn phase must be at least 1 mm \(^{12}\). We have entered the Hagedorn phase.

As we go back in time in the Hagedorn phase, the dilaton increases. At the time \(t_c\) when the dilaton equals zero, a second transition occurs, the transition to a “strongly coupled Hagedorn phase” (using the terminology introduced in [118]). We take the dilaton to be fixed in this phase. In this case, the strongly coupled Hagedorn phase may have a duration which is very long compared to the Einstein frame Hubble time immediately following \(t_R\). It is in this cosmological background that we will study the generation of fluctuations. We will denote the time when the Einstein frame Hubble radius reaches a minimum by \(t_R\), to evoke the analogy with the time of reheating in inflationary cosmology. The end of the Hagedorn phase (the time when the radiation phase of standard cosmology begins) is slightly later.

It is instructive to compare the background evolution of string gas cosmology with the background of inflationary cosmology. Figure 5 is a sketch of the space-time evolution in string gas cosmology. For times \(t < t_R\), we are in the static Hagedorn phase and the Hubble radius is infinite. For \(t > t_R\), the Einstein frame Hubble radius is expanding as in standard cosmology. To understand why string gas cosmology can lead to a causal mechanism of structure formation, we must compare the evolution of the physical wavelength corresponding to a fixed comoving scale with that of the Einstein frame Hubble radius \(H^{-1}(t)\). Recall that the Einstein frame Hubble radius separates scales on which fluctuations oscillate (wavelengths smaller than the Hubble radius) from wavelengths on which the fluctuations are frozen in and cannot be effected by microphysics. Causal microphysical processes can generate fluctuations only on sub-Hubble scales \(^{13}\). The key point is that for \(t < t_i(k)\), the fluctuation mode \(k\) is inside the Hubble radius, and thus a causal generation mechanism for fluctuations is possible.

In contrast, in inflationary cosmology (Figure 1) the Hubble radius is constant during inflation \((t < t_R\), where here \(t_R\) is the time of inflationary reheating\), whereas the physical wavelength corresponding to a fixed comoving scale expands exponentially. Thus, as long as the period of inflation is sufficiently long, all scales of interest for current cosmological observations are sub-Hubble at the beginning of inflation.

There are both important similarities and key differences between the structure formation mechanisms in inflationary cosmology and string gas cosmology. In both cases, scales are sub-Hubble during the early stages, thus allowing for a causal generation mechanism. Also, in both cases the fluctuations evolve on super-Hubble scales for a long time in the radiation phase of standard cosmology, thus leading to their squeezing, which in turn leads to the phase coherence of the

\(^{11}\)The dilaton comes to rest, but it is not pinned to a particular value by a potential. Thus, in order to obtain consistency with late time cosmology, an additional mechanism operative at late times which fixes the dilaton is required.

\(^{12}\)How to obtain this initial size starting from string-scale initial conditions constitutes the *entropy problem* of our scenario. A possible solution making use of an initial phase of bulk dynamics is given in [121].

\(^{13}\)Matter which couples minimally to gravity in the string frame is prevented from oscillating on scales larger than the Einstein frame Hubble radius by the dilaton friction term in its equation of motion.
Figure 5: Space-time diagram (sketch) showing the evolution of fixed comoving scales in string gas cosmology. The vertical axis is time, the horizontal axis is physical distance. The solid curve represents the Einstein frame Hubble radius $H^{-1}$ which shrinks abruptly to a microphysical scale at $t_R$ and then increases linearly in time for $t > t_R$. Fixed comoving scales (the dotted lines labeled by $k_1$ and $k_2$) which are currently probed in cosmological observations have wavelengths which are smaller than the Hubble radius before $t_R$. They exit the Hubble radius at times $t_i(k)$ just prior to $t_R$, and propagate with a wavelength larger than the Hubble radius until they reenter the Hubble radius at times $t_f(k)$.

fluctuations which generate the acoustic oscillations in the angular power spectrum of the CMB. However, the actual generation mechanism for fluctuations is completely different. In inflationary cosmology, any thermal fluctuations present before the onset of inflation are red-shifted away, leaving us with a quantum vacuum state, whereas in the quasi-static Hagedorn phase of string gas cosmology matter is in a thermal state. Hence, whereas in inflationary cosmology the fluctuations originate as quantum vacuum perturbations, in string gas cosmology the inhomogeneities are created by the thermal fluctuations of the string gas.

As we have shown in [122, 123, 124], string thermodynamical fluctuations in the Hagedorn phase of string gas cosmology yield an almost scale-invariant spectrum of both scalar and tensor modes. This result stems from the holographic scaling of the specific heat $C_V(R)$ (evaluated for fixed volume) as a function of the radius $R$ of the box

$$C_V(R) \sim R^2.$$  \hspace{1cm} (4.14)

As derived in [125], this result holds true for a gas of closed strings in a space-time in which the three large spatial dimensions are compact (see [128, 129] for recent papers emphasizing the role
of holography). The scaling (4.14) is an intrinsically stringy result: thermal fluctuations of a gas of particles would lead to a very different scaling.

Since the primordial perturbations in our scenario are of thermal origin (and there are no non-vanishing chemical potentials), they will be adiabatic. The spectrum of scalar metric fluctuations has a slight red tilt. As a distinctive feature [126], our scenario predicts a slight blue tilt for the spectrum of gravitational waves. The red tilt for the scalar modes is due to the fact that the temperature when short wavelength modes exit the Hubble radius is slightly lower than the temperature when longer wavelength modes exit. The gravitational wave amplitude, in contrast, is determined by the pressure. Since the pressure is closer to zero the deeper in the Hagedorn phase we are, a slight blue tilt for the tensor fluctuations results. These results are explained in more detail in a recent review [108] and in the original references. Here, we will only focus on some of the key steps.

Our approximation scheme for computing the cosmological perturbations and gravitational wave spectra from string gas cosmology is as follows (the analysis is similar to how the calculations were performed in [54, 55] in the case of inflationary cosmology). For a fixed comoving scale $k$ we follow the matter fluctuations until the time $t_i(k)$ shortly before the end of the Hagedorn phase when the scale exits the Hubble radius. At that time, we use the Einstein constraint equations (discussed below) to compute the values of $\Phi(k)$ and $h(k)$ ($h$ is the amplitude of the gravitational wave tensor), and then we propagate the metric fluctuations according to the standard gravitational perturbation equations until scales re-enter the Hubble radius at late times.

The first key point is to show how the scalar and tensor metric fluctuations can be extracted from knowledge of the energy-momentum tensor of the string gas. Working in conformal time $\eta$ and in the longitudinal gauge for the scalar metric fluctuations, the metric of a homogeneous and isotropic background space-time perturbed by linear cosmological perturbations and gravitational waves can be written in the form

$$ds^2 = a^2(\eta) \left\{ (1 + 2\Phi) d\eta^2 - [(1 - 2\Phi) \delta_{ij} + h_{ij}] dx^i dx^j \right\}.$$  \hfill (4.15)

Here, $\Phi$ (which is a function of space and time) describes the scalar metric fluctuations. The tensor $h_{ij}$ is transverse and traceless and contains the two polarization states of the gravitational waves. We have assumed that there is no anisotropic stress.

Inserting the metric (4.15) into the Einstein equations, subtracting the background terms and truncating the perturbative expansion at linear order leads to the following system of equations

$$-3\dot{\mathcal{H}} (\mathcal{H} \Phi + \dot{\Phi}) + \nabla^2 \Phi = 4\pi G a^2  \delta T_{00}^i,$$

$$\left( \mathcal{H} \Phi + \dot{\Phi} \right)_i = 4\pi G a^2 \delta T_i^0,$$

$$\left[ 2\dot{\mathcal{H}} + \mathcal{H}^2 \right] \Phi + 3\mathcal{H} \Phi' + \Phi'' = -4\pi G a^2 \delta T^i_i.$$
In the Hagedorn phase, these equations simplify substantially and allow us to extract the scalar and tensor metric fluctuations individually. Replacing comoving by physical coordinates, we obtain from the 00 equation

\[ \nabla^2 \Phi = 4 \pi G \delta T^{0}_0 \]  

and from the \( i \neq j \) equation

\[ \nabla^2 h_{ij} = -4 \pi G \delta T^{i}_j . \]  

The above equations (4.17) and (4.18) allow us to compute the power spectra of scalar and tensor metric fluctuations in terms of correlation functions of the string energy-momentum tensor. Since the metric perturbations are small in amplitude we can consistently work in Fourier space. Specifically,

\[ \langle |\Phi(k)|^2 \rangle = 16 \pi^2 G^2 k^{-4} \langle \delta T^{0}_0(k) \delta T^{0}_0(k) \rangle , \]  

where the pointed brackets indicate expectation values, and

\[ \langle |h(k)|^2 \rangle = 16 \pi^2 G^2 k^{-4} \langle \delta T^{i}_j(k) \delta T^{i}_j(k) \rangle , \]  

where on the right hand side of (4.20) we mean the average over the correlation functions with \( i \neq j \).

The second key step is to compute the matter fluctuations in the Hagedorn phase. Since this phase is dominated by the gas of strings, fluctuations in our scenario are the thermal fluctuations of a string gas. We will consider a gas of closed strings in a compact space, i.e. our three-dimensional space is considered to be large but compact. Specifically, it is important to have winding modes in the spectrum of string states.

General thermodynamical relations allow us to compute the matter correlation functions in terms of the the specific heat \( C_V \) and the pressure \( p \). The result for the energy density fluctuation is

\[ \langle \delta \rho^2 \rangle = -\frac{1}{R^3} \frac{\partial}{\partial \beta} \left( F + \beta \frac{\partial F}{\partial \beta} \right) = \frac{T^2}{R^3} C_V , \]  

where \( F \) is the free energy of the string gas and \( \beta \) is the inverse temperature. The off-diagonal pressure fluctuations, in turn, are given by

\[ \langle \delta T^{i}_j \rangle = \langle T^{i}_j \rangle - \langle T^i \rangle^2 \]  

\[ = \frac{1}{\beta R^3} \frac{1}{\partial \ln R} \left( -\frac{1}{R^3} \frac{\partial F}{\partial \ln R} \right) = \frac{1}{\beta R^2} \frac{\partial p}{\partial R} . \]  

Now, we apply these relations to the thermodynamics of strings. In [125], the thermodynamical properties of a gas of closed strings in a toroidal space of radius \( R \) were computed. To compute the fluctuations in a region of radius \( R \) which forms part of our three-dimensional compact space, we will apply the results of [125] to a box of strings in a volume \( V = R^3 \).
The starting point of the computation is the formula for the density of states $\Omega(E,R)$ which then determines the specific heat and the pressure. The result for the specific heat is

$$C_V \approx 2 \frac{R^2/\ell^3}{T (1 - T/T_H)}.$$ (4.23)

The ‘holographic’ scaling $C_V(R) \sim R^2$ is responsible for the overall scale-invariance of the spectrum of cosmological perturbations. The factor $(1 - T/T_H)$ in the denominator is responsible for giving the spectrum a slight red tilt.

For the pressure, we obtain

$$p(E,R) \approx n_H T_H - \frac{2}{3} \frac{(1 - T/T_H)}{\ell^3 R} \ln \left[ \frac{\ell^3 T}{R^2 (1 - T/T_H)} \right],$$ (4.24)

which immediately yields

$$\langle \delta T^2 \rangle \approx \frac{T}{\ell^2 R} \frac{(1 - T/T_H)}{\ell^3 R} \ln^2 \left[ \frac{R^2}{\ell^2} (1 - T/T_H) \right].$$ (4.25)

Note that the factor $(1 - T/T_H)$ (which appears in the numerator of the key term in the expression for the microcanonical partition function) is in the numerator, whereas it was in the denominator in the expression for the specific heat. The reason is that in deriving the specific heat from the microcanonical partition function, a temperature derivative was taken, but not so in deriving the pressure. This leads to the slight blue tilt of the spectrum of gravitational waves, characteristic of our proposed structure formation scenario. As mentioned earlier, the physical reason for this blue tilt is that larger wavelength modes exit the Hubble radius deeper in the Hagedorn phase where the pressure is smaller and thus the strength of the tensor modes is less.

The third key step in the analysis is to compute the power spectra for the scalar and tensor modes based on the earlier results. The power spectrum of scalar metric fluctuations is given by

$$P_\Phi(k) = \frac{1}{2\pi^2} \frac{k^3 |\Phi(k)|^2}{|\Phi(k)|^2} = 8G^2 k^{-1} < |\delta \rho(k)|^2 > .$$ (4.26)

where in the first step we have used (4.19) to replace the expectation value of $|\Phi(k)|^2$ in terms of the correlation function of the energy density, and in the second step we have made the transition to position space (note that $k = R^{-1}$).

According to (4.21), the density correlation function is given by the specific heat via $T^2 R^{-6} C_V$. Inserting the expression from (4.23) for the specific heat of a string gas on a scale $R$ yields to the final result

$$P_\Phi(k) = 8G^2 \frac{T}{\ell^3 R} \frac{1}{1 - T/T_H}$$ (4.27)

for the power spectrum of cosmological fluctuations. In the above equation, the temperature $T$ is to be evaluated at the time $t_i(k)$ when the mode $k$ exits the Hubble radius. Since modes with larger
values of $k$ exit the Hubble radius slightly later when the temperature is slightly lower, a small red tilt of the spectrum is induced. The amplitude $\mathcal{A}_S$ of the power spectrum is given by

$$\mathcal{A}_S \sim \left( \frac{l_{pl}}{l_s} \right)^4 \frac{1}{1 - T/T_H}.$$  

(4.28)

Taking the last factor to be of order unity, we find that a string length three orders of magnitude larger than the Planck length, a string length which was assumed in early studies of string theory, gives the correct amplitude of the spectrum. Thus, it appears that the string gas cosmology structure formation mechanism does not have a serious amplitude problem.

Similarly, we can compute the power spectrum of the gravitational waves and obtain

$$P_h(k) \sim 8G^2 \frac{T}{\ell_s^2} \left( 1 - T/T_H \right) \ln^2 \left[ \frac{1}{\ell_s^2 k^2} \left( 1 - T/T_H \right)^{-1} \right].$$  

(4.29)

This shows that the spectrum of tensor modes is - to a first approximation, namely neglecting the logarithmic factor and neglecting the $k$-dependence of $T(t(k))$ - scale-invariant. The $k$-dependence of the temperature at Hubble radius crossing induces a small blue tilt for the spectrum of gravitational waves.

Comparing (4.27) and (4.29) we see that the tensor to scalar ratio is suppressed by the factor $(1 - T/T_H)^2$. Given a good understanding of the exit from the Hagedorn phase we would be able to compute this ratio as well as the magnitude of the spectral tilts for both scalar and tensor modes.

4.4 Discussion

In order to put our structure formation scenario on a firm basis, we need a consistent description of the Hagedorn phase. The dilaton gravity background allows us to understand the onset of the Hagedorn phase (going backwards in time), but since the dilaton blows up, we rapidly leave the domain of applicability of the model.

A background in which our string gas structure formation scenario can be implemented [111] is the ghost-free and asymptotically free higher derivative gravity model proposed in [93] given by the gravitational action

$$S = \int d^4x \sqrt{-g} F(R)$$  

(4.30)

with

$$F(R) = R + \sum_{n=0}^{\infty} \frac{c_n}{M_s^{2n}} R \left( \frac{\partial^2}{\partial^2 t} - \nabla^2 \right)^n R,$$  

(4.31)

where $M_s$ is the string mass scale (more generally, it is the scale where non-perturbative effects start to dominate), and the $c_n$ are coefficients of order unity.

As shown in [93] and [111], this action has bouncing cosmological solutions. If the temperature during the bounce phase is sufficiently high, then a gas of strings will be excited in this phase. In the absence of initial cosmological perturbations in the contracting phase, our string gas structure formation scenario is realized. The string network will contain winding modes in the same way that a string network formed during a cosmological phase transition will contain infinite strings. The dilaton is fixed in this scenario, thus putting the calculation of the cosmological perturbations on a firm basis. There are no additional dynamical degrees of freedom compared to
those in Einstein gravity. The higher derivative corrections to the equations of motion (in particular to the Poisson equation) are suppressed by factors of \((k/M_s)^2\). Thus, all of the conditions on a cosmological background to successfully realize the string gas cosmology structure formation scenario are realized.

5. Conclusions

These lectures have focused on three topics in theoretical cosmology. The first is the inflationary universe scenario, the current paradigm of early universe cosmology. Inflation has been an extremely successful scenario. It explains why the universe is spatially flat and isotropic to the extent it is observed to be, and why it is so large and contains such a large amount of entropy, thus resolving several mysteries which the previous paradigm of cosmology, the SBB model, was not able to explain. Possibly more importantly, inflation provides a causal mechanism for the origin of the fluctuations which are now mapped out to great accuracy by recent cosmological observations. Inflation correctly (and fifteen years before the precision observations) predicted the power spectrum of CMB anisotropies.

The theory of cosmological perturbations is the key tool which allows us to take theories of the very early universe and calculate predictions for late time cosmology. This theory is applicable whatever the paradigm of early universe cosmology might be, and it is the second topic discussed in these lectures.

In spite of the successes of the inflationary universe scenario, key conceptual questions remain, making it clear that new input from fundamental physics is required in order to develop a better theory of the very early universe. The third topic discussed in these lectures, string gas cosmology, is an attempt to explore possible consequences for cosmology of superstring theory, the likely candidate for the new physics required for early universe cosmology. In particular, an alternative structure formation scenario not requiring inflationary dynamics emerges.

Acknowledgements

I wish to thank the organizers of the summer school for inviting me to lecture and for their hospitality in Dubrovnik. My research is supported by funds from NSERC and from the Canada Research Chair program.

References


H. Sato, “Galaxy Formation by Cosmic Strings,” Prog. Theor. Phys. 75, 1342 (1986);


S. Shankaranarayanan, “Is there an imprint of Planck scale physics on inflationary cosmology?,” Class. Quant. Grav. 20, 75 (2003), [arXiv:gr-qc/0203060];


