Annu. Rev. Astron. Astrophys. 1991. 29: 325–62

INFLATION FOR ASTRONOMERS

J. V. Narlikar

Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411 007, India

T. Padmanabhan

Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400 005, India

KEY WORDS: early universe, cosmology

1. INTRODUCTION

The concept of inflation was introduced into cosmology by Guth (48) about a decade ago. It has generated a remarkable degree of response, both positive and negative, from physicists. By hindsight, the idea appears a natural consequence of the concept of the phase transition, which is believed to have occurred in the very early epochs of the big bang universe, when the breakdown of the so-called grand unification symmetry took place. When it was first proposed, the concept was somewhat difficult to understand, however, as it combined ideas from particle physics with those from the general theory of relativity. Even today, controversy remains about important questions, e.g.: Was there really an inflationary phase in the universe? If yes, what was the physical mechanism behind it? Given the mode of inflation, what tangible relics should that era have left for today?

Although excellent reviews are available on this subject (see e.g. 17), they have been written largely by and for the theoreticians working on the frontier between particle physics and cosmology. This review, as its title indicates, is written for astronomers. We present the basic idea in a form that is as free of the jargon of particle physics as possible, and focus attention on the last of the three questions posed above.

0066-4146/91/0915-0325$02.00

Table 1 Notation

FRW: Friedman-Robertson-Walker
MBR: Mirowave background radiation
LSS: the last scattering surface
GUTs: grand unified theories
HDM: hot dark matter
CDM: cold dark matter
$a(t)$ = the expansion factor of the universe at cosmic time t
t_0 = the age of the universe, taking $a(0) = 0$ (thus $t = t_0$ is the present epoch)
H_0 = Hubble's constant at the present epoch
h = Hubble's constant today (i.e. at $t = t_0$) in units of 100 $\text{kms}^{-1}\ \text{Mpc}^{-1}$
$\rho_c = (3H_0^2/8\pi G)$ = critical density separating the closed and open FRW models
Ω = mass density of *nonrelativistic* particles at $t = t_0$ in units of the present critical density
Ω_B = baryonic contribution to Ω at the present epoch
Ω_{grav} = energy density of gravitational waves at $t = t_0$ in units of the present critical density
Q_0 = any physical quantity Q evaluated at $t = t_0$
T_0 = temperature of MBR at $t = t_0$
$\theta = T_0/(2.75\ \text{K})$
t_{eq} = epoch when the matter and radiation energy densities were equal
t_{dec} = epoch when matter (baryons and leptons) decoupled from radiation
H = Hubble's constant during the inflationary phase
1 MeV = 10^6 electron volts (this energy corresponds to a temperature of 1.16×10^{10} K)
t_p = Planck epoch; m_p = Planck mass.

To help the reader (and ourselves!), an outline of the notation to be used in this article is presented in Table 1.

2. THE STANDARD MODEL AND ITS DIFFICULTIES

2.1 *The Friedman-Robertson-Walker Model with a Hot Big Bang*

The Friedman models (44, 45) were proposed as the simplest solutions of Einstein's equations without the Λ-term. Robertson (104) and Walker (121) showed that global symmetry arguments of homogeneity and isotropy lead to a spacetime geometry described by the line element

$$ds^2 = c^2dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)\right], \quad k = 0, \pm 1. \qquad 1.$$

For $k = 0$, the expression in parentheses describes the Euclidean metric for three dimensional space in spherical polar coordinates. For $k = +1$, the three-space has the closed topology of the surface of a hypersphere in four dimensions, whereas for $k = -1$, the three-space is open. The (r, θ, ϕ)

coordinates are the "comoving coordinates" of a typical "fundamental observer" who "sees" the universe as homogeneous and isotropic. The proper time of such an observer is measured by t, called the *cosmic time*, and the corresponding frame of reference is called the *cosmic rest frame*. The function $a(t)$ denotes the typical length scale of the universe and, for an expanding model, it is called the *expansion factor*.

This function $a(t)$ is determined by Einstein's equations if the energy momentum tensor of the physical contents of the universe are known. If ε is the energy density ($\equiv \rho c^2$) and p the pressure of these contents, these equations can be reduced to

$$\frac{\dot{a}^2+kc^2}{a^2} = \frac{8\pi G\varepsilon}{3c^2}, \qquad 2.$$

$$\frac{d}{da}(\varepsilon a^3)+3pa^2 = 0. \qquad 3.$$

The second equation is none other than that of energy conservation under adiabatic expansion.

The present state of the universe is *matter-dominated* in the sense that bulk of the contribution to ρ comes from matter that is either at rest in the cosmic rest frame or is moving slowly (compared to c) relative to it. For such matter $p_m \approx 0$, and from Eq. 3 we get $\rho \propto a^{-3}$. The simplest model of this kind is the *flat* model, for which $\rho_m \propto (\dot{a}/a)^2$. This leads therefore to a *critical density* $\rho_c \cong 2 \times 10^{-29}\,h^2\,gcm^{-3}$ of the matter. For $k = +1$ models, $\rho > \rho_c$, whereas for $k = -1$ models, $\rho < \rho_c$. In general we write the density ρ as $\Omega\rho_c$ with $\Omega > 1$ for $k = 1$, $\Omega < 1$ for $k = -1$.

The small component of radiation present today was, however, more dominant in the past. This is because for radiation with the equation of state $\varepsilon_r = 3p_r$, Eq. 3 gives $\varepsilon_r \propto a^{-4}$. We denote by t_{eq} and $a_{eq} \equiv a(t_{eq})$, the epoch and the expansion factor when $\varepsilon_m = \varepsilon_r$. Clearly,

$$\frac{a_{eq}}{a_0} = \left(\frac{\varepsilon_r}{\varepsilon_m}\right)_0. \qquad 4.$$

The microwave background radiation (MBR) energy density may be taken as a close approximation to $(\varepsilon_r)_0$. A simple calculation gives

$$\left(\frac{a_{eq}}{a_0}\right) = 4.31 \times 10^{-5}(\Omega h^2)^{-1}\theta^4. \qquad 5.$$

We may specify the redshift z of an epoch by the relation

$$1+z=\frac{a_0}{a}. \tag{6.}$$

Clearly $1+z_{\rm eq} \approx 2.3 \times 10^4(\Omega h^2)\theta^{-4}$; if we take $\rho_{\rm mo} \approx 3 \times 10^{-31}$ gcm^{-3}, then $z_{\rm eq} \approx 10^3$. This means that for $z > z_{\rm eq}$ the universe is *radiation-dominated*, whereas for $z < z_{\rm eq}$ it is *matter-dominated.*

At times $t \ll t_{\rm eq}$ (i.e. $z \gg z_{\rm eq}$), the radiation dominated over matter, and we find that the function $a(t)$ was approximately given by

$$a(t) \propto t^{1/2}. \tag{7.}$$

(This presupposes the neglect of the curvature term kc^2/a^2 in Eq. 2 in comparison with $\dot{a}^2/a^2$. As discussed below, this assumption is nontrivial.) Since $\varepsilon \propto T^4$ for radiation, it follows that $T \propto t^{-1/2}$. This time-temperature relationship can be written, more quantitatively, as

$$\left(\frac{t}{1\,{\rm s}}\right) = 2.4\, g^{-1/2}\left(\frac{T}{1\,{\rm MeV}}\right)^{-2}, \tag{8.}$$

where g denotes the effective degrees of freedom of relativistic particles present in thermodynamic equilibrium at that temperature. The number varies between a value of about 10^2 (at 10^{20} MeV) and 3 or so (at present).

Working backwards chronologically, there are three significant epochs in the early universe. In the first epoch, electrons combined with ions to form neutral atoms. The characteristic energy was the binding energy ($\sim$13.6 eV) of the H-atom, and the temperature was about 3000–4000 K. We denote this epoch by $t = t_{\rm dec}$ to indicate that the radiation *decoupled* from matter in the absence of free electrons as scatterers. During the second epoch, free neutrons and protons combined to form light nuclei at temperatures between 10^8–10^9 K. In the third epoch, that of the grand unified theories (GUTs), the breakdown of grand-unification symmetry at energies of about 10^{14} GeV led to the bifurcation of the electroweak interaction from the strong interaction. For this $t \sim 10^{-35}$ s. (There is also a fourth epoch preceding the GUTs epoch, prior to which the universe was governed by the laws of quantum gravity. Known as the Planck epoch, it was at $t_{\rm p} \sim 10^{-43}$ s. Classical general relativity could not be valid up to this epoch.)

The third epoch, at temperatures around 10^{14} GeV, is of interest to particle physicists. Some of the basic features of our universe—e.g. the photon-to-baryon-number ratio, which is presently observed to be $(N_\gamma/N_{\rm B}) = 3.52 \times 10^7(\Omega h^2)^{-1}\theta^3$—may have become frozen in at this epoch. Discrete structures (galaxies, clusters, superclusters, etc.) could conceivably grow from primordial seeds going back to this epoch. (For reviews and discussions of some of these problems, see 81, 84, 125).

The confidence with which physicists extrapolate their discussions to epochs as early as 10^{-35} s rests on the successes of the standard hot big-bang model at the first two stages, namely the prediction of relic abundances of light nuclei and the interpretation of MBR as the relic radiation. The standard theory encounters many problems of a fundamental nature, however, which require a radical rethinking of the very early scenario. This was the motivation for introducing the concept of inflation. Before considering the proposed remedy, it is appropriate to take a look at the problems.

2.2 *Some Puzzling Features of the FRW Models*

The problematic features of the standard hot big-bang model can be briefly summarized as follows:

2.2.1. THE HORIZON PROBLEM The typical fundamental observer O at a given epoch t has his past light cone truncated at $t = 0$, the epoch of the big-bang. A light signal in the radiation-dominated era could have only travelled a proper distance

$$R_{\mathrm{H}}(t) \equiv a(t)\int_0^t \frac{dx}{a(x)} = 2ct \qquad 9.$$

during the time interval $(0, t)$. Thus, any causal communication to O is limited by a sphere of radius R_{H} centered on O. This boundary is called the *particle horizon*. Two observers O and O', separated by a proper distance larger than $2R_{\mathrm{H}}(t)$ at epoch t, will therefore have totally disjoint spheres of communication (see Figure 1). Causal connection is a necessary requirement for establishing homogeneity across a large region. Therefore there is no *a priori* reason to expect O and O' to have a similar physical environment. In short, the existence of a particle horizon limits the physical processes that might have led to an attainment of homogeneity in the universe.

If the present features of the universe were essentially frozen at the GUTs epoch, we expect the sphere of radius $2R_{\mathrm{H}}$ at that epoch to have expanded sufficiently to encompass the present observable universe with a size of about $\sim 10^{28}$ cm. (This would provide a natural explanation for the observed homogeneity of our universe.) Since the expansion factor increases in inverse proportion to temperature from the GUTs epoch ($T \sim 10^{14}$ GeV) to the present MBR temperature of $T_0 \approx 2.4 \times 10^{-4}$ eV (=2.75 K), the overall expansion is a factor 4×10^{26}. At $t \sim 10^{-35}$ s, however, $2R_{\mathrm{H}} \sim 6 \times 10^{-25}$ cm. Thus the primordial sphere of homoge-

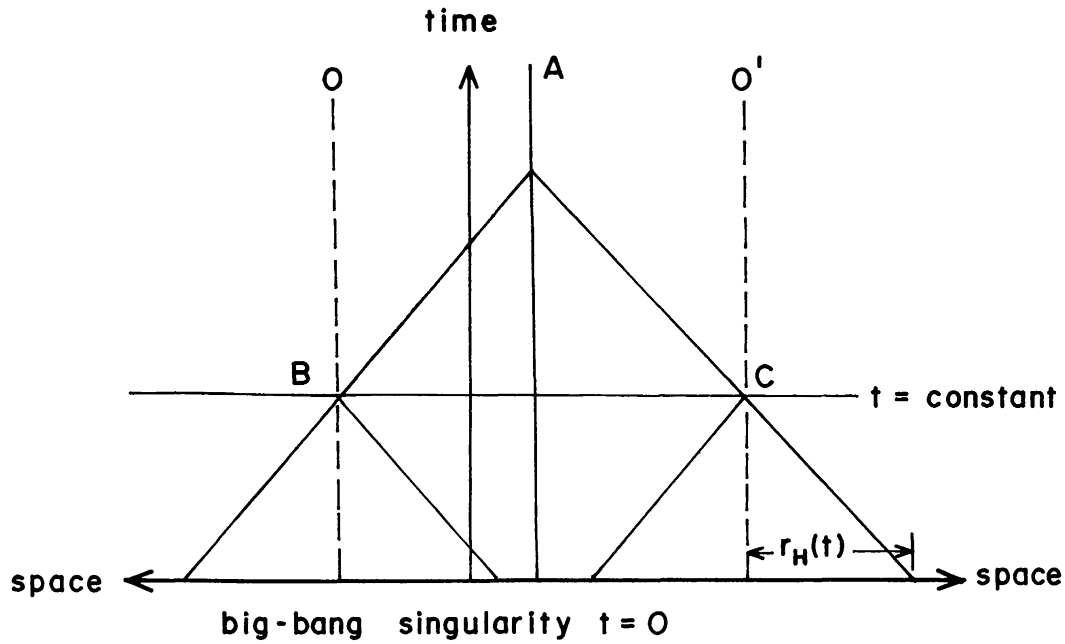

Figure 1 The particle horizons of O and O' at world points B and C are given by the segments of the $t = 0$ line where their past light cones interact it. In the situation illustrated above, O and O' are so far apart that their particle horizons are disjointed.

neity would have expanded only to a radius of about 2.4×10^2 cm, a value far short of the size of the present universe. So the presently observed large scale isotropy of the MBR and the very large scale homogeneity of discrete structures cannot be explained unless one postulates homogeneity at some initial epoch. The discrepancy is progressively reduced as one moves this epoch away from the big bang but is significant even when it is taken as late as when the radiation decoupled from matter at the last scattering surface.

2.2.2. THE FLATNESS PROBLEM We have ignored the curvature term kc^2/a^2 in Eq. 2 in considering the very early universe. The use of the density parameter Ω in Eq. 2 leads to the relation

$$\Omega(t) - 1 = \frac{kc^2}{\dot{a}^2}. \qquad 10.$$

A comparison with the present epoch gives, assuming $k \neq 0$,

$$\Omega(t) - 1 = \frac{\dot{a}_0^2}{\dot{a}^2}(\Omega - 1). \qquad 11.$$

Since, as $t \to 0$, $|\dot{a}| \to \infty$, the convergence of $\Omega(t)$ onto the flat value 1 is very rapid as we approach $t = 0$. Expressed as a function of temperature, this relation becomes

$$\Omega(T) - 1 \cong 4 \times 10^{-15}(\Omega - 1)\left(\frac{T}{1\ \text{MeV}}\right)^{-2}. \qquad 12.$$

The present astronomical observations place Ω within a range of 0.1 to 2. Thus $|\Omega - 1|$ is of order unity. At the GUTs epoch, however, with $a \propto t^{1/2}$, we have $\dot{a}^2 \propto t^{-1}$. Putting in numbers, we find that $\dot{a}_0^2/\dot{a}^2 \leq 10^{-48}$. In other words, $|\Omega(t) - 1|$ is extremely fine-tuned to value zero.

Stated differently, had this fine-tuning not occurred, the universe would have contracted back to $a = 0$ (for $k = 1$) or diffused to $a = \infty$ (for $k = -1$) long before the present epoch. In the absence of any physical mechanism, this fine-tuning has to be imposed ad hoc at the GUTs epoch in the standard model.

The flatness problem can be posed alternatively as follows. The entropy density of photons at thermal equilibrium at temperature T is given by $(4\pi^2/45)T^3$. Since $a \propto T^{-1}$, the total entropy in a proper volume, $v \propto a^3$, in an expanding universe is conserved. Its present value in the observable region is about 10^{85}. Such a large value for a dimensionless conserved quantity is hard to explain.

2.2.3. MONOPOLES The spontaneous breakdown of symmetry in the GUT scenario causes the initial larger symmetry group of the unified interactions to decompose into smaller ones, which include the symmetry group of the electromagnetic theory. A general theorem tells us that whatever the initial symmetry group, the process of symmetry breaking inevitably generates physical solutions that describe magnetic monopoles. The energy of a monopole is about 10^{16} GeV, corresponding to a mass of about 1.8×10^{-8} g. Assuming that no more than one monopole is created in one horizon-size sphere, the mass density of monopoles at the GUTs epoch works out to about 1.7×10^{65} g cm^{-3}. This is diluted by expansion through a factor of about 4×10^{26} to a present mass density of 10^{-15} g cm^{-3}—far in excess of the critical density. With a density of this order, the universe would have contracted to $a = 0$ in time less than 10^5 years! Since monopoles are stable objects, they cannot be destroyed. Thus the problem of excess monopole density seems insurmountable within the standard model.

2.2.4. DOMAIN WALLS The GUTs phase transition produces certain characteristic discontinuities in the matter distribution. This happens during the spontaneous breakdown of symmetry, which leads the mediating scalar fields to have a discrete set of characteristic values. Thus one region or "domain" has one value of the scalar field, while its neighbor has another.

The boundary wall between the domains would therefore be a region of discontinuity. We expect to see this as a sheet of inhomogeneity in the large scale matter distribution in the universe. The present data do not reveal any such inhomogeneity in the observable universe.

2.2.5. SMOOTHNESS PROBLEM The horizon problem discussed above was in the context of the large-scale homogeneity in the universe. A related problem arises when we take into account the extraordinary smoothness of the MBR at small angles (for a review see 32, 96, 128). The difficulty for the standard model comes from the chronology of the proposed scenario in which discrete structures are supposed to evolve from primordial seeds. Most theories of structure formation predict that some imprint of this event remains on the MBR in the form of small fluctuations of the order $\Delta T/T \sim 10^{-4}$ in the temperature over angular sizes ranging from a few arc seconds to a few arc minutes. Contrary to these expectations no such fluctuations have been detected. The current upper bounds are $\Delta T/T \leq 10^{-5}$. We discuss this point at length in Section 2.3.

2.2.6. THE Λ-TERM The cosmological constant Λ was introduced by Einstein somewhat empirically to arrive at a static model of the universe (38). Although with the discovery of the cosmological redshift law by Hubble, the original need for Λ disappeared, it is still taken seriously by many cosmologists. Observationally, the value of Λ has to be of the order of the square of Hubble's constant at present. Thus $|\Lambda| \leq 10^{-35}\ \mathrm{s}^{-2}$.

The GUTs epoch generates a cosmological term purely from quantum field theoretical effects. The order of magnitude of this term is very large, however, about $\sim 10^{70}\ \mathrm{s}^{-2}$. Thus we have to find why the present Λ-term is smaller by an order $\sim 10^{-105}$ of the primordial Λ-term.

2.3 *The Problem of Formation of Structures*

The above discussion is based on an idealized model of the universe, populated by matter that is assumed to be homogeneous and isotropic; the density ρ and pressure p are taken to be spatially uniform: $\rho(t, \mathbf{x}) = \rho(t)$ and $p(t, \mathbf{x}) = p(t)$. The real universe, of course, is quite different and has a very inhomogeneous distribution of matter. One possible way of characterizing such inhomogeneities is to use the "two-point-correlation function." Suppose $\bar{n}$ is the constant mean number density of galaxies in space. If the galaxies are distributed at random, the chance of finding a galaxy in a given volume δV is simply $\bar{n}\delta V$. If, however, the galaxies are distributed inhomogeneously, the chance of finding a galaxy in a volume element δV at a distance r from another galaxy, picked at random, may be expressed as

$$\delta P = \bar{n}\{1+\xi(r)\}\delta V, \tag{13}$$

where $\xi(r)$ is called the *covariance function* or *two-point correlation function*. Detailed studies of the galaxy counts indicate that the two-point correlation function for galaxies is $\xi(r) \approx (r/5h^{-1}\ \text{Mpc})^{-1.8}$ showing significant clustering of matter at scales $r < 5h^{-1}$ Mpc (28, 98). On the other end, deep three-dimensional surveys indicate evidence for *filaments* ($100h^{-1}$ Mpc long and $5h^{-1}$ Mpc across) and *voids* (empty regions about $60h^{-1}$ Mpc in diameter) at very large scales (6, 34, 47, 68, 70, 114). The average density contrast in the universe also shows significant variations depending on the coarse-graining scale. It is about 10^5 at galactic scales and about 10^2 at the scale of clusters.

This raises the question: How do we reconcile the existence of small-scale structures—galaxies, groups, and clusters—with the overall homogeneity of the universe? In particular, how should the theoretical framework be modified to take these inhomogeneities into account?

Two approaches are possible regarding this issue. We may try to obtain more realistic cosmological models in which the source energy momentum tensor of matter $T_{ik}(t, \mathbf{x})$ is inhomogeneous and anisotropic with an average value equal to that in a typical Friedman-Robertson-Walker (FRW) universe. The exact spacetime metric $g_{ik}(t, \mathbf{x})$ obtained as a solution to this problem will provide a more realistic description of our universe. Because Einstein's equations are nonlinear, no simple relation exists between the (large-scale) average value of the metric $\langle g_{ik}(t, \mathbf{x})\rangle$ and the average value of T_{ik}. Thus, this model for the universe may show significant departures from the FRW universe, even at large scales (39). This idea is extremely difficult to put into practice, however, because of our inability to find inhomogeneous solutions to Einstein's equations.

The alternative point of view, which enjoys considerable popularity, is the following: Let us assume that at some very early epoch $t = t_i$ (say), the physical variables describing the universe had small deviations around their mean values. The metric was $g_{ik}(t, \mathbf{x}) = \bar{g}_{ik}(t) + h_{ik}(t, \mathbf{x})$ and the source was $T_{ik}(t, \mathbf{x}) = \bar{T}_{ik}(t) + \tau_{ik}(t, \mathbf{x})$, where the quantities with an overhead bar denote the FRW values and the terms h_{ik} and τ_{ik} are small corrections to these mean values. Since these terms are small, one can linearize Einstein's equations in these variables and study the linear growth of metric and matter perturbations. These studies show that matter perturbations do grow under favorable circumstances, and will reach values comparable to their mean value in finite time (73, 98, 122). When this happens, the overdense region (effectively) decouples from the expansion of the universe and collapses further to form a condensation. It is generally believed that the structures in the universe arose by this process, that is, the gravitational growth of small, primordial seed perturbations.

The above picture runs into several difficulties when the details are worked out. Of these, the following three appear to be the most serious ones:

1. To have any predictive power in this model, we need to know the origin and magnitude of the initial density perturbations. In the conventional big-bang model, there is no natural seed for these perturbations. Truly primordial perturbations, e.g. those due to quantum gravitational effects at $t \approx t_P$, are likely to be of the order of one [$\mathcal{O}(1)$] and will be difficult to estimate. Thus, the theory lacks predictive power.

2. The second difficulty is of technical nature. We define the "density contrast" $\delta(t, \mathbf{x})$ and its Fourier transform $\delta_{\mathbf{k}}(t)$ by the relations

$$\delta(t,\mathbf{x}) = \frac{\rho(t,\mathbf{x})-\bar{\rho}(t)}{\bar{\rho}(t)} = \int \frac{d^3k}{(2\pi)^3}\,\delta_{\mathbf{k}}(t)e^{i\mathbf{k}\cdot\mathbf{x}}$$
$$= \int \frac{d^3k}{(2\pi)^3}\,\delta_{\mathbf{k}}(t)\exp\left\{i\frac{\mathbf{k}}{a(t)}\cdot \mathbf{l}\right\}, \qquad 14.$$

where (k/a) is the physical wave number and $|\mathbf{l}| \equiv |a(t)\mathbf{x}|$ is the proper distance. The Fourier transform separates a given inhomogeneity into components of different characteristic sizes, which may grow at different rates in the expanding universe. We can identify two effects in the time evolution of $\delta(t, \mathbf{x})$: (*a*) the growth of the amplitude $\delta_k(t)$ due to gravitational instability and (*b*) the kinematic "stretching" of the proper wave lengths $(2\pi/k)a(t)$ due to the overall expansion of the universe. Thus, a perturbation at a characteristic physical scale λ_0 today would correspond to a proper length of $\lambda_0[a(t)/a_0] \propto t^n$ in the past, if we take $a(t) \propto t^n$. The characteristic expansion scale of the universe, on the other hand, is given by the Hubble radius $cH^{-1}(t) = c(\dot{a}/a)^{-1} = cn^{-1}t$. In realistic cosmological models, $n < 1$ and hence the ratio $[\lambda(t)/cH^{-1}(t)]$ increases as we go to the earlier epochs. In other words, $\lambda(t)$ would have been larger than the Hubble radius at sufficiently high redshifts.

To every wavelength λ_0, we can associate the mass scale:

$$M(\lambda) = \frac{4\pi}{3}\bar{\rho}(t)\left(\frac{\lambda(t)}{2}\right)^3 = 1.5\times 10^{11} M_\odot \Omega h^2 \left(\frac{\lambda}{1\,\mathrm{Mpc}}\right)^3, \qquad 15.$$

which remains constant during expansion of the universe (since $\bar{\rho} \propto a^{-3}$ and $\lambda \propto a$). For example $\lambda \approx 1.88$ Mpc corresponds to the galactic mass of $10^{12}\ M_\odot$; of course, galaxies today are much smaller, because overdense regions have collapsed gravitationally and ceased to expand with the universe. Thus λ is the hypothetical size containing galactic mass, if it had continued with the cosmic expansion. This parametrization turns out to be extremely useful. The wavelength of this perturbation will be bigger than the Hubble radius at all redshifts $z > z_{\text{enter}}(M)$ where,

$$z_{\text{en}}(M) \approx \begin{cases} 1.41 \times 10^{5}(\Omega h^{2})^{1/3}(M/10^{12}\,M_{\odot})^{-1/3}; \\ \qquad M < M_{\text{eq}} \approx 3.2 \times 10^{14}\,M_{\odot}\theta^{6}(\Omega h^{2})^{-2} \\ 1.10 \times 10^{6}(\Omega h^{2})^{-1/3}(M/10^{12}\,M_{\odot})^{-2/3}; \\ \qquad M > M_{\text{eq}} \approx 3.2 \times 10^{14}\,M_{\odot}\theta^{6}(\Omega h^{2})^{-2} \end{cases}. \qquad 16.$$

It is usual to say that the perturbation carrying mass M *enters* the Hubble radius at $z = z_{\text{enter}}$ or that the perturbation with mass M was *outside* the Hubble radius at $z > z_{\text{enter}}$. The discontinuity in the two forms for Eq. 16 arises because the universe changes from the radiation-dominated phase to the matter-dominated phase at some $t_{\text{eq}} = 4.36 \times 10^{10}(\Omega h^{2})^{-2}\theta^{6}$ s with $z_{\text{eq}} = 2.32 \times 10^{4}(\Omega h^{2})\theta^{-4}$. A scale of $\lambda_{\text{eq}} \simeq 13$ Mpc $(\Omega h^{2})^{-1}\theta^{2}$ enters the Hubble radius at t_{eq}, carrying the mass $M_{\text{eq}} \cong 3.19 \times 10^{14}\ M_{\odot}$. Scales with $\lambda < \lambda_{\text{eq}}$ $(M < M_{\text{eq}})$ enter the Hubble radius before $(z > z_{\text{eq}})$, when $a(t) \propto t^{1/2}$, whereas the scales $\lambda > \lambda_{\text{eq}}$ enter later $(z < z_{\text{eq}})$, when $a(t) \propto t^{2/3}$.

Notice that according to the above calculation, a galactic mass perturbation was bigger than the Hubble radius for redshifts larger than a moderate value of about 10^{6}. This result leads to a major difficulty in conventional cosmology. It is usually assumed that physical processes can act coherently only over sizes smaller than the Hubble radius (see e.g. 8, 17). Thus any physical processes leading to small density perturbations at some early epoch $t = t_i$ could have only operated at scales smaller than $cH^{-1}(t_i)$, but most of the relevant astrophysical scales (corresponding to clusters, groups, galaxies, etc.) were much bigger than $cH^{-1}(t)$ for reasonably early epochs! Therefore, if we want the seed perturbations to have originated in the very early universe, it is difficult to understand how any physical process could have contributed to it.[1]

3. A further major difficulty in the study of structure formation is the following: It can be shown that the linear baryonic density perturbations

[1] The conclusion above based on $cH^{-1}(t)$ as a distance scale limiting causal interaction is accepted by most physicists, though it is very difficult to prove in general terms and cannot be proved by causality arguments because $cH^{-1}(t)$ and the horizon size can be quite different in a general FRW model (see 40, 93). Normally, the particle horizon at an epoch limits the range of causal communication from the past, while the event horizon limits the range of such communication to the future. The radiation-dominated era has a particle horizon of the order of Hubble radius, whereas the de Sitter universe has an event horizon of the same order. In the above argument, both the horizons are used to limit causal communication in the same sense. Moreover, for its limited duration, the inflationary phase cannot claim even to have an event horizon! Thus one must consider the Hubble radius in the above argument only as a hand-waving device.

(at scales smaller than the Hubble radius) grew as $\lambda(t) \propto a(t)$ in a matter-dominated universe from $z = z_{dec}$ up to a redshift of about $(\Omega^{-1} - 1)$ (see e.g. 98). Since $z_{dec} \simeq 1.1 \times 10^3 \theta^{-1}$, these perturbations could have grown only by a factor of about Ωz_{dec} since then. Today $\delta = (\delta\rho/\rho)$ is of order unity at scales larger than $\lambda \approx 8h^{-1}$ Mpc. Therefore, in a baryonic universe, these perturbations must have been at least $\delta \approx 0.8 \times 10^{-3}\, \Omega_B^{-1} \gtrsim 5 \times 10^{-3}$ at decoupling if we take into account the constraint $\Omega_B < 0.16$ coming from the primordial nucleosynthesis calculations (130). In the simplest models, this will lead to a temperature anisotropy of MBR of $(\Delta T/T) \approx 0.3\delta \approx 1.7 \times 10^{-3}$ at angular scales of about 4.4′. [The angular scale of anisotropy θ and the linear scale λ of the perturbation producing the anisotropy are related by $\theta \approx 0.55' \Omega h(\lambda/1 \text{ Mpc})$.] No such anisotropy has been discovered. The bounds on $(\Delta T/T)$ [less than 3×10^{-5} at 4.5′ (119)] suggest that δ is less than at least 10^{-4} at the time of decoupling. This density contrast is insufficient to produce the observed structures in a purely baryonic universe.

The perturbations described here, and elsewhere in this review, are called *adiabatic perturbations* or *curvature perturbations*. They actually change the curvature of space-time [locally, the equivalent Newtonian potential changes by $\delta\phi_k \simeq (\lambda_{phy} c^{-1} H)\delta_k$]. It is possible to think of another kind of perturbation (*isothermal*) in which $\delta\rho = 0$ but the $p = p(\rho)$ equation changes from place to place. We do not discuss them here because it seems unlikely that they are able to account for structures in the universe. [For a more detailed discussion of this and related issues, see (7, 101, 118).]

3. THE CONCEPT OF INFLATION

3.1 *The Basic Mechanism*

Some of these difficulties of the standard FRW models may be approached by modifying the dynamics of the very early universe. The trick lies in introducing a temporary phase during which the universe expanded *exponentially* as in the classical de Sitter (35) model. Such an exponentially expanding phase is called *inflation*. The de Sitter model described an empty universe with expansion caused by negative stresses due to the Λ-term. The inflationary universe also requires a Λ-term; but here it arises and is supposed to last only during the transient stage in which the GUTs phase transition is taking place. We consider the actual mechanisms proposed for this purpose in Section 3.2. Here we outline the actual model that emerges and the way it can handle some of the awkward features of the standard model.

Consider a model for the universe in which the universe was radiation-dominated up to, say, $t = t_i$, but expanded exponentially in the interval $t_i < t < t_f$:

$$a(t) = a_i \exp H(t - t_i) \qquad t_i \leq t \leq t_f. \tag{17.}$$

For $t > t_f$, evolution is again radiation-dominated $[a(t) \propto t^{1/2}]$ until $t = t_{eq} \cong 4.36 \times 10^{10} (\Omega h^2)^{-2}$ s. Evolution becomes matter-dominated for $t_{eq} < t < t_{now} = t_0$. Typical values for t_i, t_f, and H suggested in the literature are

$$t_0 \approx 10^{-35}\,\text{s}; \qquad H \approx 10^{10}\,\text{GeV}; \qquad t_f \approx 70\,H^{-1}, \tag{18.}$$

which give an overall inflation of about $A \equiv \exp N \cong \exp(70) \approx 2.5 \times 10^{30}$ to the scale factor in the period $t_i < t < t_f$. At $t = t_i$, the temperature of the universe is about 10^{14} GeV. During this exponential inflation, the temperature drops drastically but the matter is expected to be reheated to the initial temperature $\sim 10^{14}$ GeV at $t \approx t_f$. The reheating takes place when the phase transition is over, and the energy released in the process is passed on to the radiation content of the universe. The situation is analogous to the reheating that takes place when supercooled steam condenses and releases its latent heat. Thus, inflation effectively changes the value of $S \equiv T(t)a(t)$ by a factor $A = \exp(70) \approx 10^{30}$. Note that this quantity S is conserved during the noninflationary phases of the expansion.

Such an evolution, if it can be implemented dynamically, has several attractive features. Let us first explore how this evolution helps us to overcome some of the difficulties of the FRW models (48, 67, 107).

Consider first the "flatness problem," which concerns the unusually small value of the curvature term (kc^2/a^2) in the early epochs. Inflation of $a(t)$ by a factor A decreases the value of this term by a factor $A^{-2} \cong 10^{-60}$. Thus one can start with moderate values of (kc^2/a^2) before inflation and bring it down to a very small value for $t \gtrsim 10^{-33}$ s. This solves the flatness problem, interpreted as the smallness of (kc^2/a^2). Notice that no classical process can change a $k = \pm 1$ universe to a $k = 0$ universe, since doing so involves a change in topological properties. What inflation does is to decrease the importance of (kc^2/a^2) so much that, for all practical purposes, we can ignore the dynamical effect of the curvature term, thus having the same effect as setting $k = 0$.

Inflation also solves the horizon problem by bringing the entire observed region of the last scattering surface into a causally connected patch. The coordinate size of the region in the LSS from which we receive signals today is

$$l(t_0, t_{\rm dec}) = \int_{t_{\rm dec}}^{t_0} \frac{cdt}{a(t)} \cong \frac{3c}{a_{\rm dec}} (t_{\rm dec}^{2/3} t_0^{1/3}), \quad 19.$$

while the coordinate size of the horizon at $t = t_{\rm dec}$ will be

$$l(t_{\rm dec}, 0) = \int_0^{t_{\rm dec}} \frac{dct}{a(t)} \cong \frac{4ct_i}{a_{\rm dec}} \left(\frac{t_{\rm dec}}{t_f}\right)^{1/2} A. \quad 20.$$

[We have used the relations $t_0 \gg t_{\rm dec}$, $A \gg 1$, $t_i \simeq H^{-1}$ and $a_{\rm dec} = a_i A (t_{\rm dec}/t_{\rm f})^{1/2}$]. The ratio

$$R = \frac{l(t_{\rm dec}, 0)}{l(t_0, t_{\rm dec})} \cong 2A \cdot \frac{t_i}{(t_f t_{\rm dec})^{1/2}} \left[\frac{2}{3}\left(\frac{t_{\rm dec}}{t_0}\right)^{1/3}\right] \approx 4 \times 10^4 \left(\frac{A}{10^{30}}\right) \quad 21.$$

is far larger than unity for $A \simeq 10^{30}$. Thus, all the signals we receive are from a causally connected domain in the last scattering surface (LSS). Note that, in the absence of inflation, $l(t_{\rm dec}, 0) = (2ct_{\rm dec}/a_{\rm dec})$, so that $R = 2/3(t_{\rm dec}/t_0)^{1/3} \ll 1$. This value is amplified by the factor $2At_i(t_f t_{\rm dec})^{-1/2} \approx 10^7$ in the course of inflation.

Inflation can also reduce the density of any stable, relic particle, e.g. the magnetic monopoles, by a dilution factor $A^{-3} \approx 10^{-90}$, provided these relics were produced before the onset of inflation. Likewise, any discontinuities, e.g. the domain walls etc., are expanded away from one another so that the chance of observing one by a typical observer is made negligibly small. Further, because of the reheating at the end of inflation, one arrives at the large entropy presently observed in the universe. By the same token, any relic of the early universe that survives to the present (like baryon number) must be generated after the end of inflation.

The most attractive feature of the inflationary model is, however, the possibility of generating the seed perturbations that can grow to form the large-scale structures (8, 49, 52, 111). This can be achieved in the following manner:

In the FRW models with $a(t) \propto t^n$ $(n < 1)$, the physical wavelengths (which grow as $\lambda \propto a \propto t^n$) will be far larger than the Hubble radius [which grows as $H(t)^{-1} \propto t$] in the early phases. This situation is drastically altered in an inflationary model. During inflation, physical wavelengths grow exponentially [$\lambda \propto a \propto \exp Ht$] while the Hubble radius remains constant. Therefore, a given length scale has the possibility of crossing the Hubble radius twice in the inflationary models. Consider, for example, a wavelength $\lambda_0 \sim 2$ Mpc today [which, according to (15) contains a mass of a typical galaxy $1.2 \times 10^{12}(\Omega h^2)M_\odot$]. This scale would have been

$$\lambda(t_f) = \lambda_0 \frac{a(t_f)}{a(t_0)} = 2\,\mathrm{Mpc}\left(\frac{T_0}{T(t_f)}\right) \simeq 1.8 \times 10^{-2}\,\mathrm{cm} \qquad 22.$$

at the end of inflation. (This is, of course, much larger than the typical Hubble radius at that epoch, $cH^{-1} \approx 1.4 \times 10^{-24}$ cms.) But at the beginning of inflation, its proper length would have been

$$\lambda(t_i) = \lambda(t_f) \cdot \frac{a(t_i)}{a(t_f)} = A^{-1}\lambda(t_i) = 1.8 \times 10^{-32}\,\mathrm{cms}. \qquad 23.$$

This is much smaller than the Hubble radius. This Hubble radius remains constant throughout the inflation while λ increases exponentially. In an interval of $\Delta t = t - t_i \simeq 18H^{-1}$, λ will grow as big as the Hubble radius.

The situation is summarized in Figure 2. We see that the scales that are astrophysically relevant today were much smaller than the Hubble radius at the onset of inflation. (Therefore, causal processes could have operated at these scales.) During inflation, the proper wavelength grows, and becomes equal to the Hubble radius cH^{-1} at some time $t = t_{\rm exit}$. For a mode labeled by a wave vector $\mathbf{k}$, this happens at $t_{\rm exit}(\mathbf{k})$ where,

$$\frac{2\pi}{k}\, a(t_{\rm exit}) = cH^{-1}. \qquad 24.$$

That is, when $(kc/aH) = 2\pi$. In the radiation-dominated era after the inflation, the proper length grows only as $t^{1/2}$ whereas the Hubble radius grows as t; thus the Hubble radius "catches up" with the proper wavelength at some $t = t_{\rm enter}(k)$. For $t > t_{\rm enter}$, this wavelength will be completely within the Hubble radius.

Inflationary models thus allow λ to be less than cH^{-1} at two different epochs: an early phase, $t < t_{\rm exit}(k)$ and a late phase, $t > t_{\rm enter}(k)$. Any perturbations generated by physical processes at $t < t_{\rm exit}$ can be preserved intact during $t_{\rm exit} < t < t_{\rm enter}$ and can lead to formation of structures at $t > t_{\rm enter}$.

What can lead to perturbations at $t < t_{\rm exit}$? Since the physical processes taking place in this epoch are quantum mechanical by nature, quantum fluctuations in matter fields are obvious candidates as seed perturbations. Therefore, in principle, we can now generate (and compute) the density inhomogeneities in the universe—indeed a major achievement of inflation.

Notice that all the above conclusions only depend on the scale factor growing rapidly (by a factor 10^{30} or so) in a short time. For example, if the energy density $\varepsilon(t)$ varies slowly during $t_i < t < t_f$, then one has near exponential expansion with

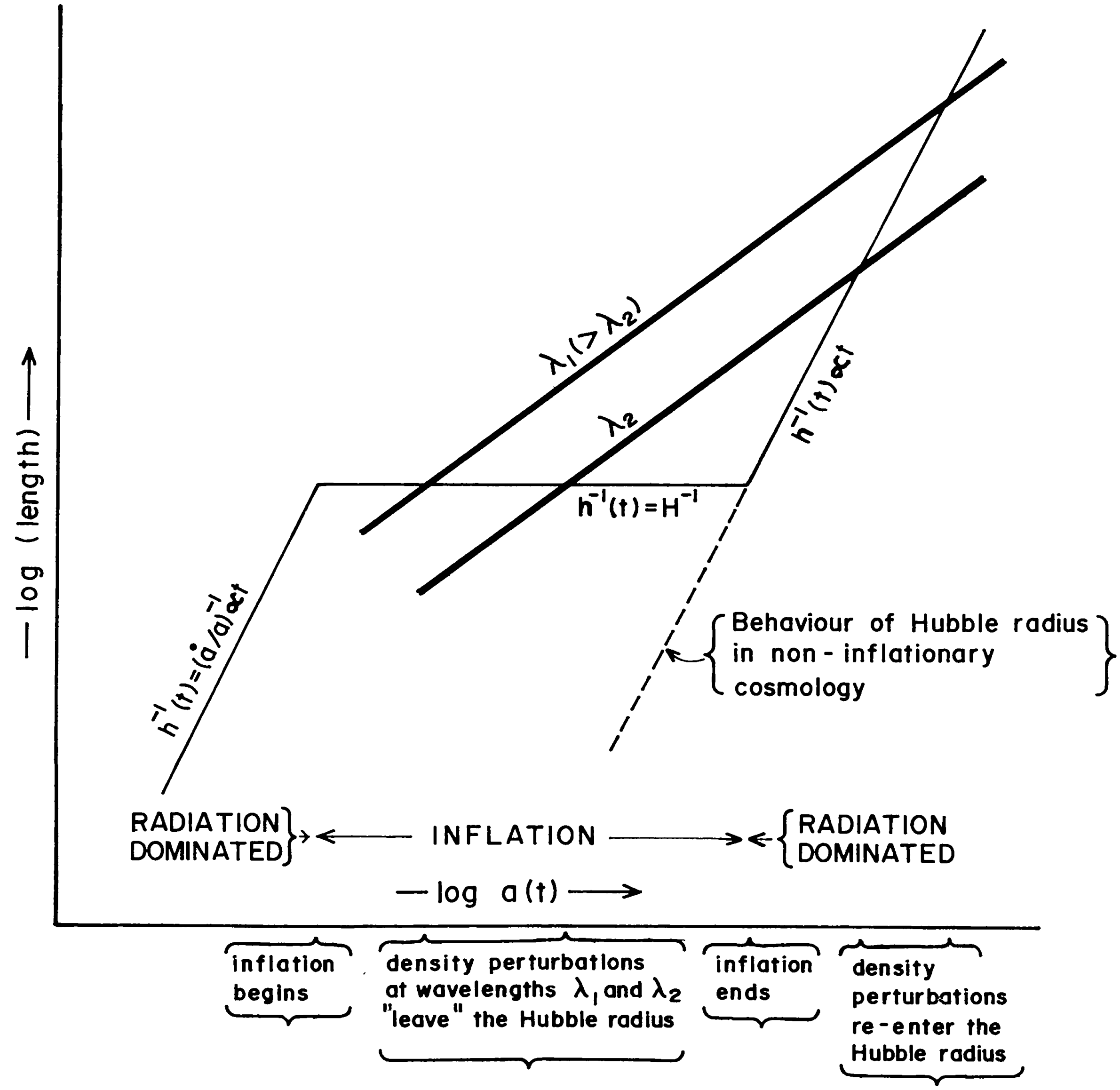

Figure 2 The figure illustrates the growths of the length scales of primordial fluctuations in relation to the Hubble radius during the inflationary and Friedman phases of expansion.

$$a(t_f) = a(t_i) \exp \int_{t_i}^{t_f} H(t)dt \equiv a(t_i) \exp N \qquad 25.$$

where $H^2(t) = [8\pi G\varepsilon(t)/3c^2]$. This provides a general definition of N.

3.2 *The Epicycles of Inflation*

Since the inflationary idea seemed to be quite attractive, several mechanisms were devised by which this idea can be implemented. Each of these models has some advantages and disadvantages and none of them is completely satisfactory. We briefly highlight three models.

For the universe to expand exponentially, the energy density should remain (at least approximately) constant. Various models of inflation differ in the process by which this is achieved. In most of them the quasi-constant energy density $\varepsilon(t)$ is derived from phase transition at the GUTs epoch. Although in specific details one grand unified theory may differ from another, almost all of them involve gauge theories with a mediating role played by the Higgs scalar field ϕ. We need not go into the intricacies of how ϕ is related to the other matter fields. The feature of interest to us is that the potential energy density V of the scalar field ϕ depends on the ambient temperature T.

At any given temperature T that is higher than a critical temperature T_c, the minimum value of V is found to be at the expected zero of ϕ. We may term this minimum at $\phi = 0$ as the *vacuum state* of ϕ. As the temperature is lowered, however, the minimum of V may no longer remain at $\phi = 0$ but may shift to a finite value $\phi = \sigma$. This phase transition occurs at $T = T_c$ and may be likened to the condensation of steam. Thus ϕ would tend to transit from $\phi = 0$ to $\phi = \sigma$.

If ϕ were to condense immediately at T_c, all the excess energy could be released at once. In the more likely case of supercooling, however, ϕ may continue at $\phi = 0$ and move to the true minimum $\phi = \sigma$ later. During this transitional stage, the state $\phi = 0$ is called the *false vacuum state*, since the true vacuum is now at $\phi = \sigma$. The original model for the inflation, due to Guth (48), invoked this temperature dependence of the potential energy of the Higgs field $V(\phi, T)$. [The potential energy has the form shown in Figure 3. Here $T_c \approx 3 \times 10^{14}$ GeV.]

At temperatures $T \gg T_c$, the potential V has only one minimum (at $\phi = 0$) with $V(0) \approx (10^{14}\,\mathrm{GeV})^4$. As the temperature is lowered to $T \sim T_c$, a second minimum appears at $\phi = \sigma$. For $T \ll T_c$, the $\phi = \sigma$ minimum is the "true" minimum [i.e. $V(\sigma) \approx 0 \ll V(0)$]. Now consider what happens in the early universe as matter cools through $T \approx T_c$. At $T \gg T_c$, the minimum configuration corresponds to $\phi = 0$ whereas for $T \sim T_c$ it is $\phi = \sigma$. Matter in the universe does not instantaneously switch over from $\phi = 0$ to $\phi = \sigma$, however. The universe can get "stuck" at $\phi = 0$ (the false vacuum), with $V = V(0)$, even at $T < T_c$, and will expand exponentially because the dominant energy density driving the expansion is the constant $V(0) - V(\sigma) \approx V(0)$. Over the course of time, thermal fluctuations and quantum tunneling will induce a transition from the false vacuum $\phi = 0$ to the true vacuum $\phi = \sigma$, thereby ending the inflation in localized regions ("bubbles"). The phase transition is expected to be completed by the expanding bubbles colliding, coalescing, and reheating the matter.

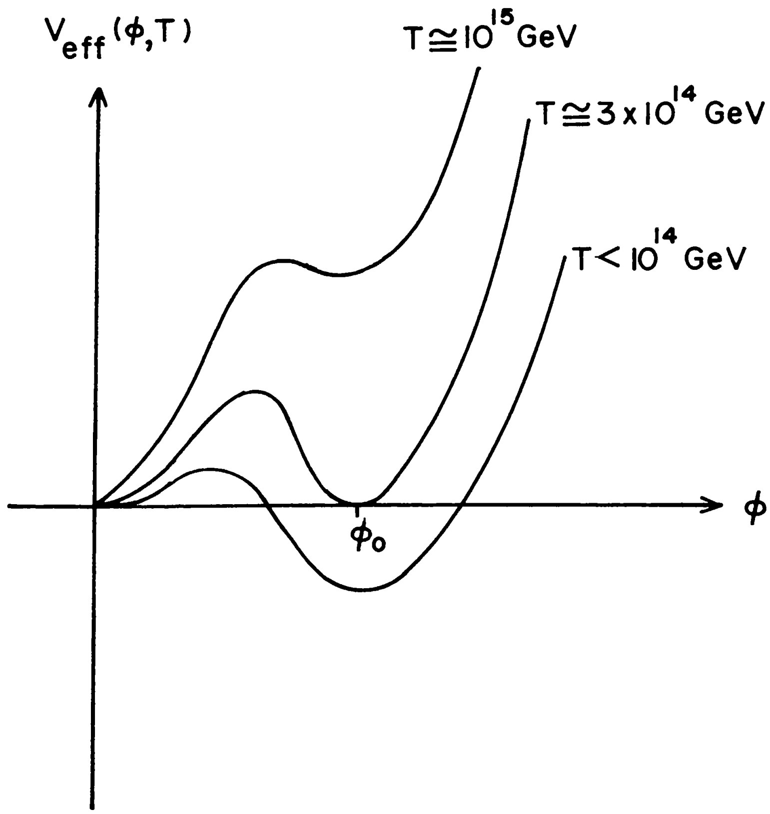

Figure 3 The potential energy of the Higgs field ϕ at various temperatures in the original model proposed by Guth.

Detailed analysis, however, shows that this model does not work (50). In order to have sufficient amount of inflation, it is necessary to keep the "false" vacuum fairly stable. In such a case, the bubble nucleation rate is small and even the resulting bubbles do not coalesce together efficiently. The final configuration is very inhomogeneous and quite different from the universe we need.

The original model was soon replaced by a version based on a very special form for $V(\phi)$ called the Coleman-Weinberg potential (3, 74, 75). At zero temperature, this potential is given by (25)

$$V(\phi) = \frac{1}{2}B\sigma^4 + B\phi^4\left[\ln\frac{\phi^2}{\sigma^2} - \frac{1}{2}\right]; \quad B \approx 10^{-3}; \quad \sigma \approx 2 \times 10^{15}\,\text{GeV}. \qquad 26.$$

This potential is extremely flat for $\phi \lesssim \sigma$ and drops rapidly near $\phi \approx \sigma$. At finite temperatures, the potential picks up a small barrier near the origin [at $\phi \simeq \mathcal{O}(T)$] with height $\mathcal{O}(T^4)$, creating a local minimum at $\phi = 0$

(see Figure 4). This false vacuum, however, is quite unstable when the temperature becomes $\mathcal{O}(10^9\ \text{GeV})$ (74). The scalar field rapidly tunnels to $\phi \approx \phi_0 \approx \mathcal{O}(H)$, and starts "rolling down" the gently sloped potential toward $\phi = \sigma$. Since the potential is nearly flat in this region, the energy density driving the universe is approximately constant and about $V(0) \approx (3 \times 10^{14}\ \text{GeV})^4$. The evolution of the scalar field in this *slow roll-over phase* can be approximated as

$$\Box\phi + V'(\phi) = \ddot{\phi} + 3H\dot{\phi} + V'(\phi) \approx 3H\dot{\phi} + V'(\phi) = 0, \tag{27.}$$

where we have ignored the $\ddot{\phi}$ term and $H = (4\pi BG\sigma^4/3c^2) \approx 2 \times 10^{10}$ GeV (in energy units). If the slow roll over lasts when ϕ varies from $\phi_{\text{start}} \simeq \mathcal{O}(H)$ to some $\phi_{\text{end}} \lesssim \mathcal{O}(\sigma)$, then

$$N \equiv \int_{t_i}^{t_f} H\,\mathrm{d}t = H\int_{\phi_s}^{\phi_e} \frac{d\phi}{|\dot{\phi}|} \approx 3\int_{\phi_s}^{\phi_e} \frac{H^2}{|V'(\phi)|}\,\mathrm{d}\phi. \tag{28.}$$

For the typical values of the Coleman-Weinberg potential, this number can easily be about 10^2, thus ensuring sufficient inflation.

As ϕ_0 approaches σ, the field "falls down" the potential and oscillates around the minimum at $\phi = \sigma$ with the frequency $\omega^2 = V''(\sigma) \approx (2 \times 10^{14}\ \text{GeV})^2 \gg H^2$. These oscillations are damped by the decay of ϕ into other particles (with some decay time Γ^{-1}, say), and by the expansion of the universe. If $\Gamma^{-1} \ll H^{-1}$, the coherent field energy $(\frac{1}{2}\dot{\phi}^2 + V)$ will be con-

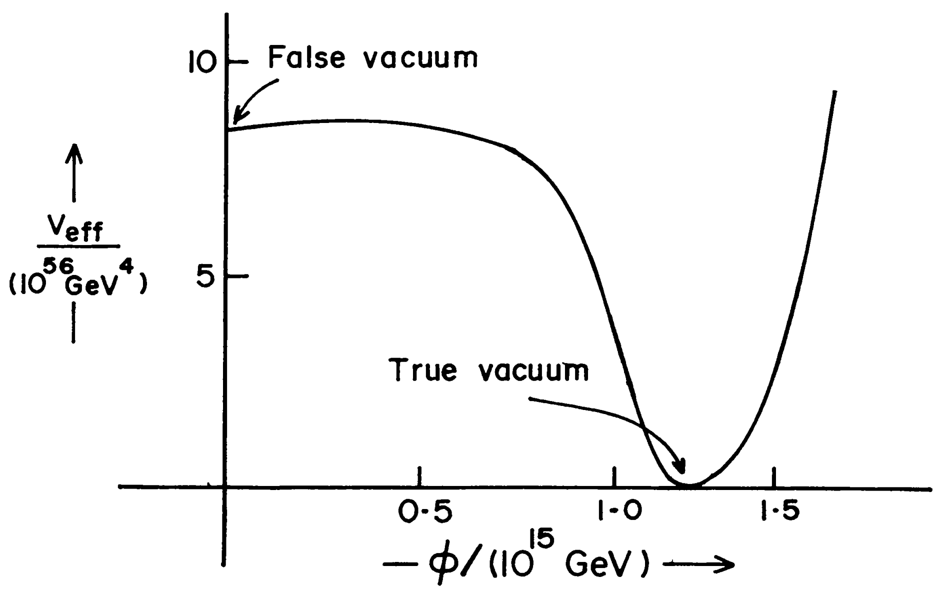

Figure 4 The Coleman-Weinberg potential that was used in the first major revision of the inflationary model.

verted into relativistic particles in a timescale $\Delta t_{\text{reheat}} \approx \Gamma^{-1} \ll H^{-1}$. This will allow the universe to be reheated to a temperature of about $T_{\text{reheat}} \approx \omega \approx 2 \times 10^{14}$ GeV $\approx T_{\text{initial}}$. The decay width of several Coleman-Weinberg models can be about $\Gamma^{-1} \approx 10^{13}$ GeV $\gg H$. This ensures good "reheating" of the universe (2, 4, 36). Because the field has already tunneled out of the false vacuum before the onset of inflation, we do not face the problems that plagued the original inflation. Instead of several bubbles having to collide, coelesce, and make up the whole observable universe of today, we have one huge bubble encompassing everything observable now.

Though it is an improvement on the original version, this model, too, is not free from problems. The field should start its slow roll over from a value $\phi_s \approx H$ to ensure sufficient inflation. The quantum fluctuations in the scalar field are about $\Delta\phi \simeq (H/2\pi)$ (75, 120). Since $\phi_s \sim \Delta\phi$, the entire analysis based on semiclassical $V(\phi)$ is of doubtful validity. The second and more serious difficulty stems from the calculation of density perturbations in this model: They are too large by a factor of about 10^6, unless the parameter B is artificially reduced by a factor 10^{-12} or so! (See Section 4.)

The original model for inflation used a strongly first order phase transition whereas the second model may be considered to be using a weakly first order (or even second order) phase transition. It is possible to construct inflationary scenarios in which no phase transition is involved. The idea of *chaotic inflation*, suggested by Linde, falls in this class (76). In this model, the potential has a very simple form: $V(\phi) = \lambda\phi^4$. Inflation results because of the rather slow motion of ϕ from some initial value ϕ_0 toward the minimum. (The initial nonzero value of the ϕ_0 is supposed to be due to "chaotic" initial conditions.) This model can also lead to sufficient inflation but suffers from two other difficulties: (*a*) To obtain the correct value for the density perturbation, it is necessary to fine-tune λ to very small values: $\lambda \approx 4 \times 10^{-14}$. (*b*) In order for the inflation to take place, the kinetic energy of the scalar field has to be small compared to its potential energy. Detailed calculation shows that this requires the field to be uniform over sizes bigger than the Hubble radius, a requirement completely against the original spirit of inflation!

A further epicycle in the saga of inflation envisages a universe whose origin was without a big bang (77). In this version, the de Sitter type inflationary phase is *self-reproducing* in a chaotic set up with the help of large scale quantum fluctuations of a scalar field φ. The bubbles of FRW models are nucleated in it at random points of space and time through quantum phase transitions.

An attractive feature of the de Sitter expansion is that because of its rapidity, the universe loses all information on initial conditions. This is a conjecture known as the *cosmic baldness hypothesis* (10). It has not been rigorously proved but looks plausible. On the basis of this hypothesis, one can assert that whatever the initial conditions, the universe will eventually reach the de Sitter state.

A solution to the bubble nucleation and coelescence problem of the original Guth model (sometimes referred to as the *graceful exit* problem) was proposed in yet another way by La & Steinhardt (71). In their *extended inflationary cosmology*, these authors used the Brans-Dicke theory of gravity (see 20) instead of general relativity as the background theory for the early universe. The inflationary phase in this model has a power law type of expansion factor instead of the exponential one, thus allowing the inflationary phase to end gracefully through bubble nucleation.

Nevertheless, this idea also ran into trouble with distortions of the MBR and was changed to *hyper-extended inflation*. The background theory of gravity for this model differs from the Brans-Dicke theory through the inclusion of higher order couplings of the scalar field with gravity (113). In a rapidly changing subject in which the half-life of a theory is one year, passing judgment on the merits of this scenario is difficult.

The schemes and shortcomings discussed above are typical of several other models suggested in the literature. The most serious constraint on inflationary scenarios arises from the study of density perturbations (discussed in detail in Section 4). No single model for inflation suggested so far can be considered completely satisfactory.

3.3 *Historical Comment*

We have already pointed out the similarity between the inflationary phase and the de Sitter spacetime. A closer similarity exists between some of the ideas invoked in inflation and the steady state model (16, 55). In the steady state model, a steady exponential expansion [with $(\dot{a}/a)$ = constant] was made compatible with a constant density of matter (ρ = constant) by invoking the continuous creation of matter. Where did this matter come from?

McCrea (78) had proposed negative stresses in vacuum to provide the required energy tensor to drive the expansion. In the 1950s, particle physics had not advanced to the sophisticated levels of today and so McCrea's phenomenological ideas were largely ignored by physicists. Today we can view them with greater sympathy.

An alternative theory based on Hoyle's earlier ideas of a cosmological field creating matter (55, 56) was investigated further by Hoyle & Narlikar (57–60). This involved a scalar C-field, whose action function was first proposed by Pryce (103; private communication). The C-field action had two terms as part of the overall Hilbert action:

$$S_c = -\frac{1}{2}f\int\int C_i C^i \sqrt{-g}d^4x + \sum_a \int C_i dx_a^i, \quad C_i \equiv \frac{\partial C}{\partial x^i}, \qquad 29.$$

of which the first, *pure field term* had a negative energy density and negative stresses for the coupling constant $f > 0$. This was responsible for driving the expansion. The second, *field-matter interaction term*, involving the effect of the C-field on matter particles (labelled by $a, b, \ldots$) and vice versa, came into operation only at the instant of creation of matter.

The advantage of the formulation by Pryce was that it is based on an action principle and hence guarantees conservation of matter and energy. In the steady state solution, matter creation is compensated by augmentation of the strength of the C-field, which has negative energy. In this solution the Hubble constant is given by

$$H = \sqrt{\frac{4\pi G f}{3}}. \qquad 30.$$

There were also solutions of this theory with no creation of matter, however. What role did they have to play? Hoyle & Narlikar (59) argued that these solutions appear as "bubbles" in a highly dense steady state universe with $H \gg H_0$, the present value of Hubble's constant. The mechanism of the switch-over from a "creation" to "noncreation" mode was left undiscussed for want of a quantum field theory of the C-field.

Thus, the key idea of the inflationary scenario of the present FRW model originating as a bubble in an external de Sitter type universe was anticipated 15 years earlier. The 1966 version [like McCrea's (78) 1951 paper] appeared too far ahead of its time to be appreciated by the contemporary particle physicists.

Finally, the cosmic baldness conjecture had also been anticipated by the steady state cosmology (57), which argued that the C-field driven expansion would wipe away any initial departures from nonuniformity.

In later years, other papers discussed the idea of inflation before or around the time of Guth's paper. For example, Kazanas (67) explicitly discussed the notion of inflation as a solution to the horizon problem. Kazanas assumed a temperature-dependent energy density of the vacuum and showed that during the phase transition, the universe would expand substantially faster than the $a \propto t^{1/2}$ law. For some temperature-dependent vacuum energies Kazanas obtained an exponential expansion law.

Sato (107) also discussed the implications of a first order phase transition of vacuum in the very early universe, and obtained the exponential expansion rate. Sato calculated the bubble nucleation rate, which depends on the particle physics theory used. He found that if the nucleation rates are small and the vacuum stays at the metastable state for a long enough time, the universe would expand exponentially, with the result that the phase transition would be further stretched out in time. Sato also considered the possibility of density and velocity fluctuations created by the phase transition growing to form galaxies.

4. THE PROBLEM OF STRUCTURE FORMATION

4.1 *The Scale Invariant Spectrum*

The most attractive feature of inflation, from the point of view of an astronomer, is the possibility that inflation may provide the seed perturbations that grow to form the structures we see today. In this section we provide an overview of how this is achieved and what difficulties arise.

We have discussed in Section 3.1 how inflation may lead to seed perturbations. To construct a working model from these ideas, we must (*a*) have a scenario that produces the observed structures from seed-perturbations and decide on the form of the seed perturbations needed, and (*b*) compute explicitly the nature of perturbations produced by inflation. A comparison of the outcomes of (*a*) and (*b*) will decide the measure of success achieved by inflation.

The first task can be achieved in principle by determining $\rho(\mathbf{x}, t_0)$ today observationally and extrapolating it backwards theoretically. In practice, of course, this is an impossibly difficult task! The structures seen today are the result of very complicated nonlinear evolution in the "recent" past (say $0 < z < 50$), and we do not have a sufficiently well-defined theoretical formalism to allow us to extrapolate back in time. (Neither do we know $\rho(\mathbf{x}, t_0)$ to sufficient accuracy.) An indirect approach to this problem is, therefore, necessary.

We know from observations of MBR that the density perturbations in the universe must have been quite small ($\ll 1$) at the time of decoupling ($z \approx 10^3$), for all astrophysically relevant scales (27, 96, 128). We also notice from Equation 16 that all these relevant scales "entered the Hubble radius" in the radiation-dominated era, i.e. before decoupling. It follows that the linear approximation is valid when the density perturbations enter the Hubble radius. One can, therefore, characterize the density perturbations completely by giving the amplitude ($\ll 1$) of each perturbation when it enters the Hubble radius. In other words, we need to specify the function

$$F(\mathbf{k}) \equiv |\delta(\mathbf{k}, t)|^2_{t=t_{\text{enter}}(\mathbf{k})} = |\delta[\mathbf{k}, t_{\text{enter}}(\mathbf{k})]|^2 \qquad 31.$$

where $t_{\text{enter}}(\mathbf{k})$ is the time at which the perturbation labeled by $\mathbf{k}$ enters the

Hubble radius decided by the equation $2\pi k^{-1}a(t_{\rm enter}) = cH^{-1}(t_{\rm enter})$. [Here, most physical quantities depend only on the magnitude $k = |\mathbf{k}|$; in such a situation, we simplify notation by writing, say $\delta(k, t)$, instead of $\delta(\mathbf{k}, t)$, etc. We may also use $\delta_{\mathbf{k}}$ to denote the same quantity as was done earlier.]

The study of the linear perturbation theory allows us to evolve further a perturbation that enters the Hubble radius (see e.g. 98, 102). Perturbations with $\lambda < \lambda_{\rm eq} \simeq 13$ Mpc $(\Omega h^2)^{-1}\theta^2$, carrying mass $M_{\rm eq} = 3.19 \times 10^{14}\ M_{\odot}\ (\Omega h^2)^{-2}\theta^6$, enter the Hubble radius in the radiation-dominated era; they grow by a small factor S—about $\mathcal{O}(10)$ or so—until $t = t_{\rm eq}$ and grow in proportion with $a(t)$ afterwards. Those perturbations with $\lambda > \lambda_{\rm eq}$ enter the Hubble radius when the universe is matter-dominated and will grow as $a(t)$ if the wavelength is bigger than the Jeans length. (For the scales in which we are interested, this condition is usually satisfied). Taking $F(k) = \alpha k^n$, one can easily work out $\delta(k, t)$ for all k at $t > t_{\rm eq}$. We obtain

$$k^3|\delta(\mathrm{k}, \mathrm{t})|^2 \propto \begin{cases} k^{n+3} \cdot S^2[a(t)/a_{\rm eq}]^2 \propto M^{-(n+3)/3}; & \lambda < \lambda_{\rm eq}, M < M_{\rm eq} \\ k^{n+7}k_{\rm eq}^{-4}[a(t)/a_{\rm eq}]^2 \propto M^{-(n+7)/3}; & \lambda > \lambda_{\rm eq}, M > M_{\rm eq} \end{cases} \qquad 32.$$

where M is the mass carried by the particular perturbation. Scales with $\lambda > cH^{-1}(t)$ are still outside the Hubble radius at this time; it can be shown that they grow as $a^2(t)$ in the radiation phase and as $a(t)$ in the matter dominated era; they also satisfy the scalings given above for $\lambda > \lambda_{\rm eq}$ (see e.g. 98).

The quantity $k^3|\delta_k|^2$ is directly related to the r.m.s. fluctuation in the mass $M(R)$ contained in a size R. For a power-law spectrum $|\delta_k|^2 \propto k^n$, it can be easily seen that the average value of $\langle(\delta M/M)^2\rangle$ in a region of size R is proportional to $k^3|\delta_k|^2$ at $k = R^{-1}$. Therefore $(\delta M/M)^2 \propto M^{-(n+3)/3}$ for $M < M_{\rm eq}$ and $(\delta M/M)^2 \propto M^{-(n+7)/3}$ for $M > M_{\rm eq}$.

Some constraints on the form and amplitude of this spectrum can be obtained from the bounds on the temperature anisotropies of MBR. A perturbation of size λ will produce anisotropies in the MBR at angular scale $\theta \simeq 0.55'\Omega h$ (λ/1 Mpc) (see e.g. 98). A scale of $\lambda_H = 65(\Omega h^2)^{-1/2}$ Mpc, subtending an angle of $\theta_H \simeq 0.87^0\Omega^{1/2}(z_{\rm dec}/1100)^{-1/2}$, will be entering the Hubble radius at the time of decoupling $t = t_{\rm dec}$. Temperature anisotropies in MBR at larger angles ($\theta > \theta_H \approx 1^0$) arise from perturbations that are still outside the Hubble radius at $t = t_{\rm dec}$ and are due to Sachs-Wolfe effect: $(\Delta T/T) \approx 0.5\delta(k, t_{\rm dec})$ (106). Fluctuations at smaller scales are due to baryons that are coupled to photons: $(\Delta T/T) \approx 0.33\ \delta_{\rm baryon}(k, t_{\rm dec})$ (see e.g. 66 and references cited therein). Bounds on $(\Delta T/T)$ from (1°–30°) imply that $\delta \lesssim 10^{-4}$ at scales corresponding to $(65\text{–}3000)h^{-1}$

Mpc today. Such a uniformity suggests that $(n+3) \gtrsim -0.1$. One can also show that if $\delta \approx \mathcal{O}(1)$ at $M \approx 10^{15} g \simeq 3 \times 10^{-19} M_{\odot}$, there will be far too many compact black holes today, thus suggesting $(n+3) \lesssim 0.2$ (21, 22). These considerations therefore favor $(n+3) = 0$, i.e. a $n = -3$ spectrum with the amplitude of about 10^{-6}–10^{-4} when each perturbation enters the Hubble radius.

Zeldovich and, independently, Harrison had argued, based on theoretical considerations, that at the time of entering the Hubble radius, the perturbations should have the index $n = -3$ (51, 134). In other words, $F(\mathbf{k}) \equiv |\delta(\mathbf{k}, t_{\text{enter}}(\mathbf{k})|^2 \propto k^{-3}$ or $k^3 F(\mathbf{k})$ should be a constant. In that case $\langle(\delta M/M)^2\rangle$ will be independent of the scale R at $t = t_{\text{enter}}(\mathbf{k})$, giving equal power at all scales at the instant of entering the horizon.

Three points need to be stressed regarding the above discussion. First, note that $k^3|\delta(\mathbf{k}, t)|^2$ is not expected to be scale-invariant (in fact, it will not be); it is only the quantity $k^3|\delta[\mathbf{k}, t_{\text{enter}}(\mathbf{k})]|^2$ that is expected to be independent of $\mathbf{k}$. That is, each scale enters the Hubble radius with an amplitude which is independent of $\mathbf{k}$; but, of course, each scale enters at a different time $t_{\text{enter}}(\mathbf{k})$. Second, the excess mass at scale R is in general related to the *integrated* power spectrum and depends on $|\delta(\mathbf{k})|^2$ at all $0 < k < R^{-1}$ rather than just to power at $k = R^{-1}$. Finally, theoretical calculations are expected to fix only $|\delta(\mathbf{k}, t)|^2$; the phase of $\delta(\mathbf{k}, t)$ is not known. This is equivalent to supposing that $[\rho(\mathbf{x}, t) - \bar{\rho}(t)]$ is a random variable with zero mean and some specified dispersion. What is determined by the theory is the power spectrum of this random variable, which is the Fourier transform of the two-point correlation function $\langle\rho(\mathbf{x}+\mathbf{y}, t)\rho(\mathbf{y}, t)\rangle$ where the brackets denote statistical averaging over the random phases of $\delta(\mathbf{k}, t)$.

4.2 *Origin of Density Perturbations*

Having decided the form of the density perturbations that is required, we now turn to the actual mechanism by which these are produced. The most natural choice, in the context of inflation, comes from the quantum fluctuations in the scalar field $\phi(t, \mathbf{x})$ driving the inflation. The computation of classical perturbations generated by a quantum field is a difficult and technically involved issue. Several questions of principle are still unresolved in this calculation (see e.g. 17, 94). Since this review is primarily intended for the astronomer, we limit discussion to the physical idea rather than to the technical aspects of the calculation.

During inflation, the universe was assumed to be, on the average, in a FRW state with small inhomogeneities. This implies that the source—which is a classical scalar field $\phi(t, \mathbf{x})$—can be split as $\phi_0(t) + f(t, \mathbf{x})$, where $\phi_0(t)$ denotes the average, homogeneous part and $f(t, \mathbf{x})$ represents the spatially dependent, fluctuating part. Since the energy density due to a scalar field is $\rho c^2 \cong \frac{1}{2}\dot{\phi}^2$, we obtain

$$\delta\rho(t, \mathbf{x}) = \rho(\mathbf{x}, t) - \bar{\rho}(t) \cong \dot{\phi}_0(t)\dot{f}(t, \mathbf{x})/c^2 \tag{33.}$$

(where $\bar{\rho}(t)c^2 = \frac{1}{2}\dot{\phi}_0(t)^2$ and we have assumed $f \ll \phi_0$). The Fourier transform will now give

$$\delta\rho(\mathbf{k}, t)c^2 \cong \dot{\phi}_0(t)\dot{Q}_k(t), \tag{34.}$$

where we have put

$$f(t, \mathbf{x}) \equiv \int \frac{d^3\mathbf{k}}{(2\pi)^3} Q_k(t)e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{35.}$$

Since the average energy density during inflation is dominated by the constant term V_0, we have the density contrast

$$\delta(\mathbf{k}, t) = \frac{\delta\rho c^2}{V_0} = \frac{\dot{\phi}_0(t)\dot{Q}_k(t)}{V_0}. \tag{36.}$$

It might now appear that all we have to do is to compute the quantities $\phi_0(t)$ and $Q_k(t)$ from the equation of motion for the scalar field. For $\phi_0(t)$ we can use the mean evolution of the scalar field during the slow roll-over phase and determine $\dot{\phi}_0(t)$ from the classical solution. The fluctuating field $f(t, \mathbf{x})$ is supposed to be some classical object mimicking the quantum fluctuations. Such a quantity is conceptually difficult to visualize and justify. What is usually done is to choose some convenient quantum mechanical measure for fluctuations and define Q_k in terms of this quantity.

In quantum theory, the field $\hat{\phi}(t, \mathbf{x})$ and its Fourier coefficients $\hat{q}_k(t)$ will become operators related by

$$\hat{\phi}(t, \mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} \hat{q}_k(t)e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{37.}$$

The quantum state of the field can be specified by giving the quantum state $\psi_k(q_k, t)$ of each of the modes $\hat{q}_k$. (One can think of q_k as coordinates of a particle and $\psi_k(q_k, t)$ as the wavefunction describing this particle.)

The fluctuations in q_k can be characterized by the dispersion

$$\sigma_k^2(t) = \langle\psi|q_k^2(t)|\psi\rangle - \langle\psi|q_k(t)|\psi\rangle^2 = \langle\psi|q_k^2(t)|\psi\rangle \tag{38.}$$

in this quantum state. (The mean value of the scalar field operator $\langle\hat{\phi}(t, \mathbf{x})\rangle = \phi_0(t)$ is homogeneous; therefore, we have set $\langle\hat{q}_k\rangle$ to zero in the above expression. Note that we are interested only in the $\mathbf{k} \neq 0$ modes.) Expressing $\hat{q}_k$ in terms of $\hat{\phi}(t, \mathbf{x})$, it is easy to see that

$$\sigma_k^2(t) = \int d^3\mathbf{x} \langle \psi | \hat{\phi}(t, x+y) \hat{\phi}(t, y) | \psi \rangle e^{i\mathbf{k}\cdot\mathbf{x}}. \quad 39.$$

In other words, the power spectrum of fluctuations σ_k^2 is related to the Fourier transform of the two-point-correlation function of the scalar field. Since $\sigma_k^2(t)$ appears to be a good measure of quantum fluctuations, we may attempt to define $Q_k(t)$ as

$$Q_k(t) = \sigma_k(t). \quad 40.$$

This is equivalent to defining the fluctuating classical field $f(t, \mathbf{x})$ to be

$$f(t, \mathbf{x}) \equiv \int \frac{d^3k}{(2\pi)^3} \sigma_k(t) e^{i\mathbf{k}\cdot\mathbf{x}}. \quad 41.$$

This leads to the result

$$\delta(\mathbf{k}, t) = \frac{\dot{\phi}_0(t)}{V_0} \dot{\sigma}_k(t). \quad 42.$$

The procedure may be summarized as follows:

1. In quantum theory, the field $\hat{\phi}(t, \mathbf{x})$ and its Fourier coefficient $\hat{q}_k(t)$ become operators. In any quantum state, the variables will have a mean value and fluctuations around this mean value.
2. Since the mean evolution of the scalar field is described by a homogeneous part $\phi_0(t)$, we expect the mean values of $\hat{q}_k$ to vanish (for $\mathbf{k} \neq 0$); $\langle \psi | \hat{q}_k(t) | \psi \rangle = 0$. The fluctuations around these mean values, however, characterized by $\sigma_k^2(t) = \langle \psi | \hat{q}_k^2 | \psi \rangle$, do not vanish.
3. We incorporate these quantum fluctuations in a semiclassical manner by taking the scalar field to be $\phi(t, \mathbf{x}) = \phi_0(t) + f(t, \mathbf{x})$, where $f(t, \mathbf{x})$ is related to $\sigma_k(t)$ by (41).
4. The density perturbations are calculated by treating $\phi(t, \mathbf{x})$ as a classical object.

The expression derived above gives the value of $\delta(\mathbf{k}, t)$ in the inflationary phase: $t_i < t < t_f$. To compare this with observations, we need to know the value of $\delta(\mathbf{k}, t)$ at $t = t_{\text{enter}}(k)$, that is, when the perturbations enter the Hubble radius. Fortunately, an approximate conservation law relates the value $\delta(\mathbf{k}, t_{\text{enter}})$ with $\delta(\mathbf{k}, t_{\text{exit}})$, where $t_{\text{exit}}(k)$ is the time at which the relevant perturbation "leaves" the Hubble radius in the inflationary epoch (8, 19, 46). This law can be stated as

$$\frac{\delta[\mathbf{k}, t_{\text{exit}}(k)]}{1 + W(t_{\text{exit}})} = \frac{\delta[\mathbf{k}, t_{\text{enter}}(\mathbf{k})]}{1 + W(t_{\text{enter}})}, \quad 43.$$

where $W(t)$ is the ratio between pressure $p(t)$ and the energy density $\rho(t)c^2$ of the background (mean) medium: $W(t) = p(t)/\rho(t)c^2$. In the inflationary phase with the scalar field,

$$p(t) = \frac{1}{2}\dot{\phi}_0^2 - V_0; \quad \rho(t)c^2 = \frac{1}{2}\dot{\phi}_0^2 + V_0; \quad 1 + W(t) \cong \frac{\dot{\phi}_0^2}{V_0}, \qquad 44.$$

where we have used the fact $\dot{\phi}_0^2 \ll V_0$. In the radiation-dominated phase (at $t = t_{\text{enter}}$), $1 + W = 4/3$. Therefore

$$\delta(\mathbf{k}, t_{\text{enter}}) = \delta(\mathbf{k}, t_{\text{exit}}) \cdot \frac{4}{3}\left(\frac{V_0}{\dot{\phi}_0^2}\right); \qquad 45.$$

or using Eq. 42,

$$\delta(\mathbf{k}, t_{\text{enter}}) = \frac{4}{3}\left(\frac{\dot{\sigma}_k}{\dot{\phi}_0}\right)_{t=t_{\text{exit}}} \cong \left(\frac{\dot{\sigma}_k}{\dot{\phi}_0}\right)_{t=t_{\text{exit}}}. \qquad 46.$$

This is the final result.

The problem now reduces to computing $\sigma_k(t)$ and $\phi_0(t)$, which can be done once the potential $V(\phi)$ is known. For a Coleman-Weinberg potential (see Section 3.2), detailed calculations give (see e.g. 17) the final result

$$\delta(\mathbf{k}, t_{\text{enter}}) \approx \lambda^{1/2} N^{3/2} k^{-3/2} \approx 10^2 k^{-3/2}, \qquad 47.$$

where we have taken the effective e-folding time $N \approx 50$ and $\lambda \approx 0.1$. We see that the density perturbations have the correct spectrum but too high an amplitude. To bring it down to the acceptable value of about 10^{-4}, we need to take the dimensionless parameter λ to be about 10^{-13}! This requires an extreme fine-tuning for a dimensionless parameter, especially since we have no other motivation for such a value.

This has been the most serious difficulty faced by all realistic inflationary models: They produce too large an inhomogeneity. The qualitative reason for this result can be found from Eq. 46. To obtain *slow* roll-over and sufficient inflation, we need to keep $\dot{\phi}_0$ small, and this tends to increase the value of δ. We could have saved the situation if it were possible to keep σ_k arbitrarily small; unfortunately, the inflationary phase induces a fluctuation of about $(H/2\pi)$ on any quantum field due to field theoretical reasons (see e.g. 18, 72). This lower bound prevents us from getting sensible values for δ unless we fine-tune the dimensionless parameters of $V(\phi)$ (for a general discussion, see 92). Several solutions have been suggested in the literature to overcome this difficulty but none of them appears compelling (see e.g. 42, 54, 62, 86, 90, 94, 108).

4.3 *A Critique of Inflation: Expectations and Performance*

The concept of inflation enjoyed considerable popularity in the first half of the last decade, though this enthusiasm seems to have died down somewhat in recent years. Taking stock of the expectations raised by inflationary scenarios in the light of its actual performance is therefore worthwhile.

Such a discussion falls conveniently into two types of questions:

1. How serious are the original problems that the inflation was invoked to solve? What alternative explanations are possible for these difficulties and how well does inflation perform vis à vis the other solutions?
2. What are new features, good and bad, that inflationary scenarios have introduced into cosmology?

Let us start with the first set of questions. In the original version, inflation was suggested as a possible solution to the horizon, flatness, and monopole problems. Of these, the monopole problem has received considerably less attention in recent literature. The currently favored unified field theories do not lead to a monopole problem. Thus, we concentrate on the flatness and horizon problems.

It is indeed true that inflation does solve these problems if these problems were stated in a particular form: inflation can suppress the value of (kc^2/a^2) term; it can also make the observed region of the last scattering surface (LSS) a causally connected domain. One should not ignore, however, certain disturbing features in the inflationary solution to these problems:

Both these problems deal with the initial conditions for our universe. Since Einstein's equations permit $k = -1$, 0 or $+1$ in a FRW solution, the value of k needs to be supplied as an extra input to classical theory. But the creation of the universe (i.e. physics at $t < t_p$) needs to be understood quantum mechanically rather than classically! [It is quite possible that quantum gravitational effects make a $k = 0$ model highly probable (see e.g. 87).] Inflation thus tries to provide a classical answer to an inherently quantum gravitational question.

This anomaly is quite striking when we consider the horizon problem. The horizon problem exists because the integral

$$r(t) = \int_0^t \frac{dx}{a(x)} \qquad 48.$$

is finite. But do we know that it is finite? To make such a claim we have to assume that there was a singularity at $t = 0$ and that we know the behavior of $a(t)$ arbitrarily close to $t = 0$. For $t < t_p$, quantum gravitational effects will modify the behavior of $a(t)$ and will probably eliminate the singularity problem. Then, for almost all $a(t)$ (except for a class of functions of measure zero), the above integral will diverge, automatically solving the horizon problem. In other words, flatness and horizon problems owe their existence to our using classical physics beyond its domain of validity. There are several quantum gravitational models in which these problems are solved as an offshoot of elimination of the singularity problem (see e.g. 82, 83, 88). This is a possibility that did not exist in classical cosmology.

Even within the context of the inflationary models, the solutions to these problems work only for a limited time (40, 41, 93). For example, Ellis has shown how the flatness problem will resurface in the late-time behavior of a $k \neq 0$ universe. [This should be clear from the fact that at sufficiently late times, both the $k = +1$ and $k = -1$ models will behave very differently irrespective of the present value of the (kc^2/a^2) term.] In the case of the horizon problem, it can be shown that LSS will appear homogeneous only for times $t < t_{\rm crit}$ when

$$t_{\rm crit} \cong \frac{t_i^2}{t_f} A^2 = \frac{t_i^2}{t_f} \exp 2H(t_f - t_i). \qquad 49.$$

[For $t_i \approx 10^{-34}$ s, $t_f \approx 100t_i$, $H \approx (10^9$ GeV), $t_{\rm crit} \approx 3 \times 10^{23}$ s, which is far larger than $t_0 \approx 3 \times 10^{17}$ s; so, right now, $t_0 < t_{\rm crit}$.] This quantity $t_{\rm crit}$ is determined once and for all by microscopic physics at $t \approx 10^{-34}$ s! If we wait long enough, t will be larger than $t_{\rm crit}$, and the horizon problem will resurface. Thus inflation offers only a temporary, though long, relief from these problems, and one has to invoke the anthropic principle (11) to justify its success. In contrast, solutions based on quantum gravitational models solve these problems permanently.

Let us now consider the second question, that is, the new features that inflation has brought into cosmology. The major success of inflation in our opinion lies in these factors.

First, inflation has provided a mechanism that allows one, in principle, to compute the spectrum of density inhomogeneities from fundamental physics. This must be considered a success because this is the first time we have a computable mechanism for producing density inhomogeneities.

Also, a definite prediction emerges from inflationary models. The same mechanism that produces the density inhomogeneities also produces gravitational wave perturbations. These perturbations also have a scale invariant power spectrum and an r.m.s. amplitude of about $(H/10^{19}$ GeV). The energy density of the gravitational waves contributes a fraction $\Omega_{\rm grav} \approx 10^{-5}(H/m_{\rm p})^2 h^{-2}$ to the critical density, where $m_{\rm p}$ is the Planck mass (1, 5, 43, 105, 129). Such perturbations can induce a quadrupole anisotropy in the MBR. The present bounds on this anisotropy ($\lesssim 10^{-4}$) suggest that $H < 10^{15}$ GeV. The value of Ω_{grav} can also be restricted by the timing measurements of the millisecond pulsar; the present bound is $\Omega_{grav}(\lambda \sim 1pc) \leq 3 \times 10^{-7}$. A positive detection of quadrupole anisotropy in MBR or a direct detection of relic gravitational radiation will certainly go a long way toward boosting confidence in inflation. [The Laser Interferometer Gravity Wave Observatory (LIGO) and similar projects can, in principle, reach a sensitivity of $\Omega_{grav} \sim 10^{-11}$.]

Second, several investigations have suggested that inflation could be a generic feature of cosmology; in other words, for almost any kind of $V(\phi)$ and for a large class of initial conditions, the universe will undergo a phase of exponential expansion (63, 64, 112, 116). The amount of inflation, time of occurrence, etc. can all be quite varied, but the physical phenomenon is probably here to stay (see, however, the discussion in Section 6).

The above two features also have a negative side, discussed in Section 4.2: Inflation produces too much density inhomogeneity. Every unworkable model for inflation rules out a parameter range for the potential $V(\phi)$ and can, in principle, constrain particle physics models. This is probably the most important unsolved problem in the physics of the early universe.

5. INFLATION AND DARK MATTER

Observations indicate that our universe may contain a large amount of nonluminous, "dark" matter. Inflation adds an interesting new dimension to this issue.

In the noninflationary cosmologies, there is no preferred value for the density parameter Ω. This quantity was treated by most astronomers as an input from observations into the theory. Inflationary models, however, make a very definite prediction about Ω: In the usual scenarios, $\Omega = 1 + \mathcal{O}(10^{-4})$ [the correction of $\mathcal{O}(10^{-4})$ arising because of the fluctuations in the curvature induced by density inhomogeneities]. This prediction can be tested in two ways: First, this constraint, added to the fact that the universe is matter-dominated today, implies that $H_0 t_0 = (2/3)$. If, based on stellar and galactic ages, we take $t_0 = (12\text{–}20)$ Gyr, we have to rule out values of h greater than 0.65. An independent determination of H_0 leading to, say, $h \geq 0.65$ will rule out inflation.

More directly, inflationary models can be ruled out if observations conclusively point out a value of $\Omega < 1$. Since luminous matter in the universe contributes far less than unity, one may say that the inflationary idea can be correct only if there exists some form of dark matter in sufficient quantities.

The observational status of the value of Ω is not very certain. The following claims have been made in the literature (for a review, see 13, Ch. 10, 100, 115):

1. The mean density in the solar neighborhood gives about $0.003h^{-1}$.
2. Studies based on the Magellanic stream and timing arguments in local groups give a higher value of about $0.06h^{-1}$.
3. The mass density in groups of galaxies contributes about 0.16 and that in large clusters give about 0.25.
4. The Virgo centric fall also suggests a mass density of 0.25.
5. The constraints from primordial nucleosynthesis imply a constraint on the baryonic contribution to mass density: $\Omega_B = (0.014-0.026)h^{-2}$; or if we take $0.5 < h < 1$, we get $0.014 < \Omega_B < 0.104$. (Authors differ somewhat on the upperbound on Ω_B and the cited values are in the range 0.1 to 0.2; 69.)

Two features stand out in the above estimates if they are all correct. First, there seems to be a tendency for Ω to increase with the scale over which it is measured. (This conclusion is somewhat tentative and quite controversial.) If we use gravitational effects occurring in a system of size L to measure Ω, we will miss out on matter distributed smoothly over sizes significantly larger than L; thus, if the universe has a significant fraction of mass distributed smoothly over scales larger than, say 50 Mpc, we can still reconcile the above observations with $\Omega = 1$. Second, the observations are just marginally consistent with a fully baryonic universe with $\Omega_B \approx 0.2$ if (and only if) $h = 0.4$. Thus, these observations alone probably do not rule out a completely baryonic universe (yet!).

Inflation makes definite claims regarding the above situation. First, inflation is not compatible with a purely baryonic universe (since $\Omega_B < 0.2$). Thus not only does inflation demand dark matter, it demands nonbaryonic dark matter. Second, inflation requires most of the dark matter (a fraction of about 0.75) to be distributed smoothly at scales larger than about 20 Mpc or so, to escape small-scale bounds on Ω.

Non-baryonic dark matter may be needed for a completely different reason. As discussed above, perturbations in the baryonic matter are constrained to be less than about 10^{-4} at decoupling. This does not give sufficient time for these perturbations to grow into the structures we see today. The constraints on non-baryonic matter, which does not couple directly to radiation, are often less severe. Dark matter perturbations can grow by a factor of at least (T_{eq}/T_{dec}) from the epoch of matter domination till the time of decoupling. Baryonic matter can be allowed to "catch up" with dark matter perturbations after the decoupling. This can reduce the $(\Delta T/T)$ value by a factor of about 50–100 or so in many models (e.g. 109).

A model for galaxy formation incorporating these constraints is not very easy to build, however. If we assume that the dark matter is made of weakly interacting massive particles (WIMPs), we can broadly distinguish between two scenarios for structure formation depending on the mass m_x of the WIMP.

The growth of structures in a universe dominated by a non-baryonic dark matter particle of mass m_x proceeds as follows: These particles will become nonrelativistic at a temperature $T_{NR} = m_x$ corresponding to the time $t_{NR} \approx g^{-1/2}$ $(m_p^3/T_{NR}^2) \simeq 1.2 \times 10^7\,\text{s}\,(m_x/1$ KeV$)^{-2}r^2$, where the r is the ratio of the g-factors for the x-particle and photons. This implies that these particles were moving with relativistic velocities from the time t_D they decoupled from other matter until t_{NR}. Since these particles can cover a distance of $c(t_{NR}-t_D) \approx ct_{NR}$ ("free streaming") in this time, all perturbations at scales less than ct_{NR} would have been wiped out (14, 15, 99). This corresponds (today) to the wavelength $\lambda_{FS} \approx (ct_{NR})$ $(T_{NR}/T_0)r \approx 1$ Mpc $(m_x/$ KeV$)^{-1}r$, containing a mass of $M_{FS} \simeq 1.5 \times 10^{11}$ $M_\odot$ $(\Omega h^2)(m_x/1$ KeV$)^{-3}r^3$. This drastically reduces the power in the spectrum $k^3|\delta_k|^2$ for $M < M_{FS}$. Thus the spectrum has significant power only in the range $M_{FS} < M < M_{eq}$; within this range, the spectrum is also reasonably flat.

Candidates for dark matter with $M_{FS} \lesssim M_{eq}$ (i.e. for $m_x \lesssim 100$ eV) are called *hot dark matter*; their power spectrum will be peaked around $M_{FS} \approx M_{eq} \simeq 10^{14}$–$10^{15}$ $M_\odot$. Candidates with $m_x \gtrsim 10$ KeV are called *cold dark matter*; they will have a relatively flat, gently declining, power spectrum from about 10^8 $M_\odot$ to 10^{15} $M_\odot$.

For $m_x \lesssim 100$ eV (hot dark matter), the first structures that form contain masses of about 10^{15} $M_\odot$; these structures collapse, forming "pancakes" and "filaments" around which baryons cluster (132, 133). Numerical simulations suggest that in order to reproduce the observed galaxy-galaxy correlation function, the collapse should have occurred rather recently ($z \approx 1$–2), a requirement that is difficult to reconcile with the existence of quasars of higher redshift (23, 126, 127). Many astronomers, therefore, view the hot dark matter scenarios with disfavor.

If WIMPs have considerably higher mass [say $m_x \simeq \mathcal{O}(1$ GeV) or so], it is possible to form galactic size objects first and evolve larger structures by gravitational clustering. Numerical simulations with these scenarios produce results that are in better agreement with observation, provided that $\Omega h \approx 0.2$ or so (29–33). This, of course, is *a priori* inconsistent with inflation. In contrast, HDM can easily accommodate $\Omega = 1$.

Several attempts have been made to reconcile cold dark matter (CDM) scenarios with inflation. This can be done (for example) by invoking a very smooth contribution $\Omega_{smooth} = 1 - \Omega_{CDM} \approx 0.8$ due to some relativistic particle (85, 95, 117), or, more brazenly, by postulating a cosmological constant of the required magnitude. An alternative would be to produce a scenario in which galaxy formation is "biased"—i.e. only very overdense regions (3σ peaks) lead to the galaxies we see (9, 65). The more numerous, 1σ-peaks will be distributed in a relatively smoother fashion. No final consensus has emerged regarding these ideas and none of them appears to have a compelling simplicity or naturalness.

In conclusion, it is worth pointing out one loophole in the conclusion $\Omega = 1$ in an inflationary scenario. If there is a residual Λ-term, then $\Omega = 1$ need not correspond to $k = 0$. Should it turn out that the FRW models with $k = 0, \Omega = 1$ are inconsistent with data, the inflationary scenario can still survive with a nonzero Λ, although such a formula for survival would have been purchased dearly at the cost of simplicity and Occam's razor.

We next discuss the role of the Λ-term in the context of inflation.

6. THE COSMOLOGICAL CONSTANT

The source term for Einstein's equations is any conserved stress tensor. If this stress tensor is isotropic with strength Λ (corresponding to an equation of state $p = -\rho c^2 = -\Lambda$, implying either pressure or density to be negative), then the quantity Λ is called the *cosmological constant*. Such a term was originally postulated by Einstein and has an interesting history (see e.g. 124).

The value of such a constant is severely constrained by cosmological considerations. Since Λ contributes $\Omega_\Lambda = (8\pi G\Lambda/3H_0^2)$ to the critical density, one can safely conclude (in spite of any astronomical uncertainties!) that

$$|\Lambda| \lesssim 10^{-29}\,\mathrm{g\,cm^{-3}} \approx 10^{-47}\,(\mathrm{GeV})^4. \qquad 50.$$

As discussed below, this small value is a deep mystery.

To begin with, nothing prevents the existence of a Λ-term (say, Λ_0) in Einstein's equations (as postulated by him). This will make the gravitational part of the Lagrangian dependent on two fundamental constants Λ_0 and G, which differ widely in scale; the dimensionless combination made out of these fundamental constants, $(G\hbar/c^3)\Lambda_0$, has a value less than 10^{-126}!

The surprise in the smallness of $|\Lambda|$ is mainly due to the following: We do not know of any symmetry mechanism that requires it to be zero. In fact, we know of several independent, unrelated phenomena that contribute to Λ. To produce such a small Λ, these terms have to be fine-tuned to a bizarre accuracy.

Quantum field theory provides a wide variety of contributions to Λ. For example, consider a scalar field with a potential $V(\phi)$. The particle physics predictions do not change (except possibly in theories with exact supersymmetry, a condition not realized in nature) if we add a constant term V_0 to this potential. A potential like

$$V_1(\phi) = -\frac{1}{2}\mu^2\phi^2 + \frac{\lambda}{4}\phi^4$$

and

$$V_2 = \frac{\lambda}{4}\left(\phi^2 - \frac{\mu^2}{\lambda}\right)^2$$

will give the same effects, though they differ by the constant term $(\mu^4/4\lambda)$, but such a shift in the energy-density will contribute to Λ. According to currently accepted scenarios, the value of this constant changes in every phase transition by E^4 where E is the energy scale at which the phase transition occurs: at the GUTs transition, it is 10^{56} (GeV)^4; at the Salam-Weinberg transition, it changes by 10^{10} (GeV)^4. These are enormous numbers compared to the present value of 10^{-47} (GeV)^4. How a physical quantity can change by such a large magnitude and finally adjust itself to be zero at such fantastic accuracy is not clear.

Finally, one should not forget that the "zero-point energy" of quantum fields will also contribute to gravity (91, 131). Each degree of freedom contributes an amount

$$\Lambda \cong \int_0^{k_{max}} \frac{4\pi k^2 dk}{(2\pi)^3} \sqrt{k^2+m^2} \cong \frac{k_{max}^4}{8\pi^2}, \qquad 51.$$

where k_{max} is an ultraviolet cut-off. If we take general relativity to be valid up to Planck energies, then we may take $k_{max} \approx 10^{19}$ GeV and Λ will be 10^{76} (GeV)^4.

If we assume that all the contributions are indeed there, then they have to be fine-tuned to cancel each other, for no good reason. Before the entry of GUTs into cosmology, we needed to worry only about the first and last contribution, both of which could be tackled in an ad hoc manner. One arbitrarily sets $\Lambda_0 = 0$ in the Lagrangian defining gravity and tries to remove the zero-point contribution by complicated regularization schemes. (Neither argument is completely water-tight but both seem plausible.) With the introduction of GUTs and inflationary scenarios however, the cosmological constant becomes a dynamical entity and the situation becomes more serious. Notice that it is precisely the large change in the $V(\phi)$ that leads to a successful inflation; it has to be large to inflate the universe and change to a small value in the end for a graceful exit from the inflationary phase.

Several mechanisms have been suggested in the literature to make the cosmological constant zero: supersymmetry (26, 79, 80, 135), complicated dynamical mechanisms (37, 97, 110, F. Wilczek and A. Zee, unpublished), probabilistic arguments from quantum gravity (12, 24, 53, 89), and anthropic principle (123) are only a few of them. None of these seems to provide an entirely satisfactory solution. A somewhat different approach in which scale invariance was used to set $\Lambda = 0$ was provided by the conformal theory of gravity (61).

The smallness of the cosmological constant is probably the most important single problem that needs to be settled in cosmology. We have no idea as to what this mechanism is, but if it is based on some general symmetry consideration, it may demand vanishing of Λ at all epochs. This can wipe out the entire inflationary picture.

7. OUTLOOK

The inflationary model epitomizes the deep influence very high energy particle physics can have on cosmology. The idea is an attractive one, and, given the current momentum on the frontier of particle physics and cosmology, further epicycles are not ruled out. In the last analysis, however, astronomers should judge it by its impact on present day observations. Does it leave relics that are observed today? The formation of structures, the values of the density parameter, and the cosmological constant are examples of such relics. The present evidence on these issues is at best ambivalent and at worst embarrassing for the current theories of inflation. This may be because the particle physicists have not yet been clever enough to hit the jackpot of the correct unification theory.

Literature Cited

1. Abbott, L., Wise, M. 1984. *Nucl. Phys. B* 244: 541
2. Abbott, L., Farhi, E., Wise, M. 1982. *Phys. Lett. B* 117: 29
3. Albrecht, A., Steinhardt, P. J. 1982. *Phys. Rev. Lett.* 48: 1220
4. Albrecht, A., Steinhardt, P. J., Turner, M. S., Wilczek, F. 1982. *Phys. Rev. Lett.* 48: 1437
5. Allen, B. 1988. *Phys. Rev. D* 37: 2078
6. Bahcall, N. A., Soneira, R. M. 1983. *Ap. J.* 270: 20
7. Bardeen, J. M. 1980. *Phys. Rev. D* 22: 1882
8. Bardeen, J., Steinhardt, P., Turner, M. S. 1983. *Phys. Rev. D* 28: 679
9. Bardeen, J. M., Bond, J. R., Kaiser, N., Szalay, A. S. 1986. *Ap. J.* 300: 15
10. Barrow, J. D., Stein-Schabes, J. 1984. *Phys. Lett. A* 103: 315
11. Barrow, J. D., Tipler, F. J. 1986. *The Anthropic Cosmological Principle*. New York: Clarendon Press
12. Baum, E. 1983. *Phys. Lett. B* 133: 185
13. Binney, J., Tremaine, S. 1987. *Galactic Dynamics*. Princeton: Princeton Univ. Press
14. Bond, J. R., Szalay, A. S. 1984. *Ap. J.* 274: 443

15. Bond, J. R., Efstathiou, G., Silk, J. 1980. *Phys. Rev. Lett.* 45: 1980
16. Bondi, H., Gold, T. 1984. *Mon. Not. R. Astron. Soc.* 108: 252
17. Brandenberger, R. H. 1985. *Rev. Mod. Phys.* 57: 1
18. Brandenberger, R., Khan, R. 1982. *Phys. Lett. B* 119: 75
19. Brandenberger, R., Khan, R. 1984. *Phys. Rev. D* 29: 2172
20. Brans, C., Dicke, R. H. 1961. *Phys. Rev.* 24: 925
21. Carr, B. J. 1975. *Ap. J.* 201: 1
22. Carr, B. J. 1976. *Ap. J.* 206: 8
23. Centrella, J., Melott, A. 1982. *Nature* 305: 196
24. Coleman, S. 1988. *Nucl. Phys. B* 310: 643
25. Coleman, S., Weinberg, E. 1973. *Phys. Rev. D* 7: 1888
26. Cremmer, E., Ferrara, S., Konnas, C., Nanopoulos, D. V. 1983. *Phys. Lett. B* 133: 61
27. Davies, R. D., Lasenby, A. N., Watson, R. A., Daintree, E. J., Hopkins, J., Beckman, J., Sanchez-Almeida, J., Rebolo, R. 1987. *Nature* 326: 462
28. Davies, M., Peebles, P. J. E. 1983. *Ap. J.* 267: 465
29. Davis, M., Efstathiou, G., Frenk, C., White, S. D. M. 1985a. *Ap. J.* 292: 371
30. Davis, M., Efstathiou, G., Frenk, C., White, S. D. M. 1985b. *Ap. J. Suppl.* 57: 241
31. Davis, M., Efstathiou, G., Frenk, C., White, S. D. M. 1985c. *Nature* 317: 595
32. Davis, M., Efstathiou, G., Frenk, C., White, S. D. M. 1987. *Nature* 330: 451
33. Davis, M., Efstathiou, G., Frenk, C., White, S. D. M. 1988. *Ap. J.* 327: 507
34. De Lapparent, Geller, M., Huchra, J. 1986. *Ap. J.* 302: L11
35. De Sitter, W. 1917. *Proc. Akad. Weteusch. Amsterdam* 19: 1217
36. Dolgov, A., Linde, A. D. 1982. *Phys. Lett. B* 116: 329
37. Dolgov, A. D. 1982. In *The Very Early Universe*, ed. G. W. Gibbons, S. W. Howking, S. T. C. Siklos. Cambridge: Cambridge Univ. Press
38. Einstein, A. 1917. *Preuss. Akad. Wiss. Berlin. Sitzber* 142
39. Ellis, J. F. R. 1988. Univ. Cape Town. Preprint
40. Ellis, J. F. R. 1988. *Class. Quantum Grav.* 5: 891
41. Ellis, J. F. R., Stoger, W. 1988. *Class. Quantum Grav.* 5: 207
42. Ellis, J., Enqvist, K., Nanopoulous, D., Olive, K., Srodnicki, M. 1985. *Phys. Lett. B* 152: 175
43. Fabri, R., Pollock, M. 1983. *Phys. Lett. B* 125: 445
44. Friedman, A. 1922. *Z. Phys.* 10: 377
45. Friedman, A. 1924. *Z. Phys.* 21: 326
46. Frieman, J. A., Turner, M. S. 1984. *Phys. Rev. D* 30: 265
47. Giovanelli, R., Haynes, M. P. 1982. *Ap. J.* 87: 1355
48. Guth, A. H. 1981. *Phys. Rev. D* 23: 347
49. Guth, A., Pi, S.-Y. 1982. *Phys. Rev. Lett.* 49: 1110
50. Guth, A., Weinberg, E. 1983. *Nucl. Phys. B* 212: 321
51. Harrison, E. R. 1970. *Phys. Rev. D* 1: 2726
52. Hawking, S. W. 1982. *Phys. Lett. B* 115: 295
53. Hawking, S. W. 1984. *Phys. Lett. B* 134: 403
54. Holman, R., Ramond, P., Ross, G. G. 1984. *Phys. Lett. B* 137: 343
55. Hoyle, F. 1948. *Mon. Not. R. Astron. Soc.* 108: 372
56. Hoyle, F. 1949. *Mon. Not. R. Astron. Soc.* 109: 365
57. Hoyle, F., Narlikar, J. V. 1963. *Proc. R. Soc. London Ser. A* 273
58. Hoyle, F., Narlikar, J. V. 1966. *Proc. R. Soc. London Ser. A* 290: 143
59. Hoyle, F., Narlikar, J. V. 1966. *Proc. R. Soc. London Ser. A* 290: 162
60. Hoyle, F., Narlikar, J. V. 1966. *Proc. R. Soc. London Ser. A* 290: 177
61. Hoyle, F., Narlikar, J. V. 1966. *Proc. R. Soc. London Ser. A* 294: 138
62. Jensen, L., Olive, K. 1986. *Nucl. Phys. B* 263: 731
63. Jensen, L., Stein-Schabes, J. 1986. *Phys. Rev. D* 34: 931
64. Jensen, L., Stein-Schabes, J. 1987. *Phys. Rev. D* 35: 1146
65. Kaiser, N. 1985. *Ap. J.* 284: L9
66. Kaiser, N., Silk, J. 1986. *Nature* 324: 529
67. Kazanas, D. 1980. *Ap. J.* 241: L59
68. Kirschner, R. F., Oemler, A., Shecter, P. L., Schectman, S. A. 1983. *Early Evolution of the Universe and Its Present Structure, I.A.U. Symp.* 104: 197
69. Kolb, E. W., Turner, M. S. 1990. *The Early Universe*. Menlo Park, CA: Addison-Wesley
70. Koo, D., Kron, R., Szalay, A. 1986. *13th Texas Symp. Relativistic Astrophys*, ed. P. Ülmer. Singapore: World Scientific
71. La, D., Steinhardt, P. J. 1989. *Phys. Rev. Lett.* 62: 376
72. Lapedes, A. 1978. *J. Math. Phys.* 19: 2289
73. Lifshitz, E. M., Khalatnikov, I. 1963. *Adv. Phys.* 12: 185
74. Linde, A. 1982. *Phys. Lett. B* 108: 389
75. Linde, A. D. 1982. *Phys. Lett. B* 116: 335

76. Linde, A. D. 1983. *Phys. Lett. B* 129: 177
77. Linde, A. D. 1987. *Phys. Scripta T* 15: 169
78. McCrea, W. H. 1951. *Proc. R. Soc. London Ser. A* 206: 562
79. Moore, G. 1987. *Nucl. Phys. B* 293: 139
80. Moore, G. 1987. Preprint
81. Narlikar, J. V. 1983. *Introduction to Cosmology*. Boston: Jones and Bartlett
82. Narlikar, J. V. 1984. *Found. Phys.* 14: 443
83. Narlikar, J. V., Padmanabhan, T. 1983. *Ann. Phys.* 150: 289
84. Narlikar, J. V., Padmanabhan, T. 1986. *Gravity, Gauge Theories and Quantum Cosmology*. Dordrecht: Reidel
85. Olive, K., Seckel, D., Vishniac, E. T. 1985. *Ap. J.* 292: 1
86. Ovrut, B., Steinhardt, P. J. 1983. *Phys. Lett. B* 133: 161
87. Padmanabhan, T. 1983. *Phys. Lett. A* 96: 110
88. Padmanabhan, T. 1983. *Phys. Rev. D* 28: 756
89. Padmanabhan, T. 1984. *Phys. Lett. A* 104: 196
90. Padmanabhan, T. 1988. *Phys. Rev. Lett.* 60: 2229
91. Padmanabhan, T. 1989. *Int. J. Mod. Phys. A* 4: 4735
92. Padmanabhan, T., Seshadri, T. R. 1986. *Phys. Rev. D* 34: 951
93. Padmanabhan, T., Seshadri, T. R. 1988. *Class Quantum Grav.* 5: 221
94. Padmanabhan, T., Seshadri, T. R., Singh, T. P. 1989. *Phys. Rev. D* 39: 2100
95. Padmanabhan, T., Vasanthi, M. M. 1987. *Ap. J.* 315: 411
96. Partridge, R. B. 1988. *Rep. Prog. Phys.* 51: 647
97. Peccei, R. D., Solá, J., Wetterich, C. 1987. *Phys. Lett. B* 195: 183
98. Peebles, P. J. E. 1980. *Large Scale Structure of the Universe*. Princeton: Princeton Univ. Press
99. Peebles, P. J. E. 1982. *Ap. J. Lett.* 263: L1
100. Peebles, P. J. E. 1986. *Nature* 321: 27
101. Press, W., Vishniac, E. T. 1980. *Ap. J.* 239: 1
102. Primack, J. R. 1984. In *Proc. Int. Sch. Phys. "Enrico Fermi" Course 92*, ed. N. Cabiboo. Amsterdam: North Holland
103. Pryce, M. H. L. 1960. Preprint
104. Robertson, H. P. 1935. *Ap. J.* 82: 248
105. Rubakov, V. A., Sazhin, M., Veryaskin, A. 1982. *Phys. Lett. B* 115: 189
106. Sachs, R. K., Wolfe, A. M. 1967. *Ap. J.* 147: 73
107. Sato, K. 1981. *Mon. Not. R. Astron. Soc.* 195: 467
108. Shafi, Q., Vilenkin, A. 1984. *Phys. Rev. Lett.* 52: 691
109. Silk, J. 1986. In *Inner Space/Outer Space*, ed. E. W. Kolb et al. Chicago: Univ. Chicago Press
110. Singh, T. P., Padmanabhan, T. 1988. *Int. J. Mod. Phys. A* 3: 1593
111. Starobinsky, A. A. 1982. *Phys. Lett. B* 117: 175
112. Starobinsky, A. A. 1983. *JETP Lett.* 37: 66
113. Steinhardt, P. J., Accetta, F. S. 1990. *Phys. Rev. Lett.* 64: 2740
114. Strauss, M. A., Davis, M. 1988. *Proc. Vatican Study Week on Large-scale Motions in the Universe*, ed. G. Coyne, V. Rubin. Vatican City: Pontifical Sci. Acad.
115. Trimble, V. 1988. *Contemp. Phys.* 29: 373
116. Turner, M. S., Widrow, L. M. 1986. *Phys. Rev. Lett.* 57: 2237
117. Turner, M. S., Steigman, G., Krauss, L. 1984. *Phys. Rev. Lett.* 52: 2090
118. Turner, M. S., Cohen, A., Kaplan, D. 1989. *Phys. Lett. B* 216: 20
119. Uson, J., Wilkinson, D. 1984. *Ap. J.* 277: L1
120. Vilenkin, A., Ford, L. 1982. *Phys. Rev. D* 26: 1231
121. Walker, A. G. 1936. *Proc. London Math. Soc.* 42(2): 90
122. Weinberg, S. 1972. *Gravitation and Cosmology*. New York: Wiley
123. Weinberg, S. 1987. *Phys. Rev. Lett.* 59: 2607
124. Weinberg, S. 1989. *Rev. Mod. Phys.* 61: 1
125. Weinberg, S. 1980. *Phys. Scripta* 21: 773
126. White, S. D. M., Frenk, C., Davis, M. 1983a. *Ap. J.* 274: L1
127. White, S. D. M., Frenk, C., Davis, M. 1983b. *Ap. J.* 271: 417
128. Wilkinson, D. T. 1987. *13th Texas Symp. on Relativistic Astrophys.* ed. M. P. Ulmer. Singapore: World Scientific
129. Yagnik, U. A. 1990. *Phys. Lett. B* 234: 271
130. Yang, J., Turner, M. S., Steigman, G., Schramm, D: N., Olive, K. 1984. *Ap. J.* 281: 493
131. Zeldovich, Ya. B. 1967. *JETP Lett.* 6: 316
132. Zeldovich, Ya. B. 1970. *Astrofizika* 6: 319
133. Zeldovich, Ya. B. 1970. *Astron. Astrophys.* 5: 84
134. Zeldovich, Ya. B. 1972. *Mon. Not. R. Astron. Soc.* 160: 1p
135. Zumino, B. 1975. *Nucl. Phys. B* 89: 535