TOWARD THE INFLATIONARY PARADIGM: LECTURES ON INFLATIONARY COSMOLOGY*

By M. S. Turner

Theoretical Astrophysics, Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA

and

Department of Physics and Astronomy and Astrophysics, The University of Chicago, Illinois 60637, USA

(Received February 23, 1987)

A review of the present status of inflationary cosmology is presented.

PACS numbers: 98.80.Bp

1. Overview

Guth's inflationary Universe scenario has revolutionized our thinking about the very early Universe. The inflationary scenario offers the possibility of explaining a handful of very fundamental cosmological facts — the homogeneity, isotropy, and flatness of the Universe, the origin of density inhomogeneities and the origin of the baryon asymmetry, while at the same time avoiding the monopole problem. It is based upon microphysical events which occurred early ($t \leqslant 10^{-34}$ sec) in the history of the Universe, but well after the Planck epoch ($t \geqslant 10^{-43}$ sec). While Guth's original model was fundamentally flawed, the variant based on the slow-rollover transition proposed by Linde, and Albrecht and Steinhardt (dubbed 'new inflation') appears viable. Although old inflation and the earliest models of new inflation were based upon first order phase transitions associated with spontaneous-symmetry breaking (SSB), it now appears that the inflationary transition is a much more generic phenomenon, being associated with the evolution of a weakly-coupled scalar field which for some reason or other was initially displaced from the minimum of its potential. Models now exist which are based on a wide variety of microphysics: SSB, SUSY/SUGR, compactification of extra dimensions, R^2 gravity, induced gravity, and some random, weakly-coupled scalar field. While there are several models which successfully implement the inflation, none is particularly compelling and all seem somewhat *ad hoc*. The common distasteful feature of all the successful models is the necessity

* Presented at the XXVI Cracow School of Theoretical Physics, Zakopane, Poland, June 1–13, 1986.

of a small dimensionless number in the model — usually in the form of a dimensionless coupling of order 10^{-15}. And of course, all inflationary scenarios rely upon the assumption that vacuum energy (or equivalently a cosmological term) was once dynamically very significant, whereas today there exists every evidence that it is not (although we have no understanding why it is not). For these reasons I have entitled these lectures *Toward the Inflationary Paradigm.* I have divided my lectures into the following sections: Successes of the standard cosmology; Shortcomings of the standard cosmology; New inflation — the slow-rollover transition; Scalar field dynamics; Origin of density inhomogeneities; Specific models, I. Interesting failures; Lessons learned — prescription for successful inflation; Two models that work; The inflationary paradigm; Loose ends, and Inflation confronts observation.

2. *The standard cosmology and its successes*

The hot, big bang cosmology — the co-called standard cosmology, neatly accounts for the (Hubble) expansion of the Universe, the 2.7 K microwave background radiation (see Figs 1, 2), and through primordial nucleosynthesis, the cosmic abundances of the light elements D and ^{4}He (and in all likelihood, ^{3}He and Li as well; see Fig. 3). The most distant galaxies and QSO's observed to date have redshifts in excess of 3 — the current

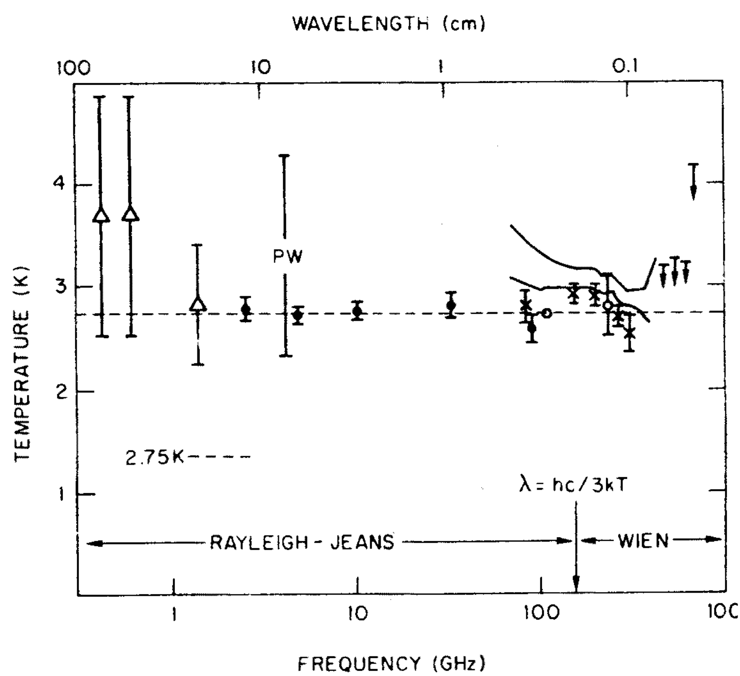


Fig. 1. Summary of microwave background temperature measurements from $\lambda \simeq 0.05$ to 80 cm (see Refs. [4]). Measurements indicate that the background radiation is well described as a 2.75 ± 0.05 K black body. PW denotes the discovery measurement of Penzias and Wilson

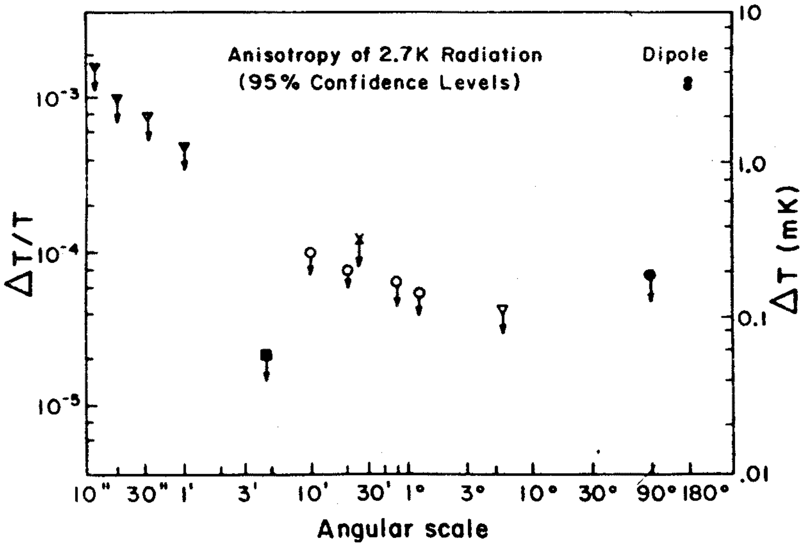


Fig. 2. Summary of microwave background anisotropy measurements on angular scales from 10″ to 180° (see Ref. [5]). With the exception of the dipole measurements, the rest are 95% confidence upper limits to the anisotropy

record holders are: for galaxies $z = 3.2$ (Ref. [1] and QSO's $z = 4.0$ (Ref. [2]). The light we observe from an object with redshift $z = 3$ left that object only 1–2 Byr after the bang. Observations of even the most distant galaxies and QSO's are consistent with the standard cosmology, thereby testing it back to times as early as 1 Byr (see, e.g., Ref. [3]). The surface of last scattering for the microwave background is the Universe at an age of a few $\times$ 50^5 yrs and temperature of about 3000 K. Measurements at wavelengths from 0.05 cm to 80 cm indicate that it is consistent with being radiation from a black body of temperature 2.75 K ± 0.05 K (see Fig. 1 and Ref. [4]). Measurements of the isotropy indicate that the temperature is uniform to a part in 1000 on angular scales ranging from 10″ to 180° — to a part in 10^4 after the dipole component is removed (see Fig. 2 and Ref. [5]). The observations of the microwave background test the standard cosmology back to times as early as 100,000 yrs. According to the standard cosmology, when the Universe was 0.01 sec–300 sec old, corresponding to temperatures of $\simeq$ 10 MeV–0.1 MeV, conditions were right for the synthesis of a number of light nuclei. The predicted abundances of D, ^{3}He, ^{4}He, and ^{7}Li are consistent with their observed abundances provided that the baryon-to-photon ratio is

$$\eta \equiv n_b/n_\gamma \approx (4-7)\times 10^{-10}. \tag{1}$$

The baryon-to-photon ratio and the fraction of critical density contributed by baryons are related by: $\Omega_b h^2/T_{2.7}^3 \simeq 3.53\times 10^7\,\eta$ where $T_{2.7}$ is the microwave temperature in units of 2.7 K and h is the present value of the Hubble constant in units of 100 km s^{-1} Mpc^{-1}. The allowed range for η corresponds to: $0.014 \lesssim \Omega_b h^2/T_{2.7}^3 \lesssim 0.025$, implying that baryons alone cannot provide the closure density. The concordance of theory and observation for D and ^{4}He is particularly compelling evidence in support of the standard cosmology as there are no known contemporary astrophysical sites which can simultaneously account for the primordial abundances of both these isotopes (see Fig. 3; see Ref. [6] for further discussion of primordial nucleosynthesis). In sum, all the available evidence indicates that the standard cosmology provides an accurate accounting of the evolution of the Universe from 0.01 sec after the bang until today, some 15 or so Byr late—quite a remarkable achievement!

I will now briefly review the standard cosmology (more complete discussions of the standard cosmology are given in Ref. [3]). Throughout I will use high energy physics units, where $\hbar = k = c = 1$. The following conversion factors may be useful.

$$\begin{aligned}
1\ \mathrm{GeV}^{-1} &= 0.197\times 10^{-13}\ \mathrm{cm},\\
1\ \mathrm{GeV}^{-1} &= 0.658\times 10^{-24}\ \mathrm{sec},\\
1\ \mathrm{GeV} &= 1.160\times 10^{13}\ \mathrm{K},\\
1\ \mathrm{GeV}^{4} &= 2.32\times 10^{17}\ \mathrm{gcm}^{-3},\\
1\ \mathrm{M}_{\odot} &= 1.99\times 10^{33}\ \mathrm{g} \simeq 1.2\times 10^{57}\ \mathrm{baryons},\\
1\ \mathrm{pc} &= 3.26\ \mathrm{light-year} \simeq 3.09\times 10^{18}\ \mathrm{cm},\\
1\ \mathrm{Mpc} &= 3.09\times 10^{24}\ \mathrm{cm},\\
G_{\mathrm{N}} &= 6.673\times 10^{-8}\ \mathrm{cm}^{3}\ \mathrm{g}^{-1}\ \mathrm{sec}^{-2} \equiv m_{\mathrm{Pl}}^{-2},\\
(m_{\mathrm{Pl}} &= 1.22\times 10^{19}\ \mathrm{GeV}).
\end{aligned}$$

On large scales ($\gg 100$ Mpc) the Universe is isotropic and homogeneous, as evidenced by the uniformity of the 2.7 K background radiation, the x-ray background, and counts of galaxies and radio sources, and so the standard cosmology is based on the maximally-symmetric Robertson–Walker line element

$$ds^2 = -dt^2 + R^2(t)\,[dr^2/(1-kr^2) + r^2 d\theta^2 + r^2 \sin^2\theta d\varphi^2], \tag{2}$$

where ds^2 is the square of the proper separation between two space-time events, k is the curvature signature (and can, by a suitable rescaling of R, be set equal to -1, 0, or $+1$), and $R(t)$ is the cosmic scale factor. The expansion of the Universe is embodied in $R(t)$ — as $R(t)$ increases all proper (i.e., physical — as measured by meter sticks) distances scale with $R(t)$. The coordinates r, θ, and φ are comoving coordinates: test particles initially at rest will have constant comoving coordinates, and the velocity of NR test particles moving with respect to the comoving coordinates decrease ($\propto R(t)^{-1}$). The distance between two objects comoving with the expansion, e.g., two galaxies, simply scales up with $R(t)$. The momentum of any freely-propagating particle decreases as $1/R(t)$. In particular, the wavelength of a photon $\lambda \propto R(t)$, i.e., is redshifted by the expansion of the Universe.

The coordinate distance at which curvature effects become noticeable is $|k|^{-1/2}$, which corresponds to the physical (or proper) distance

$$R_{\mathrm{curv}} \simeq R(t)\,|k|^{-1/2} \tag{3}$$

— which one might call the curvature radius of the Universe. Note that R_{curv} also just scales with the cosmic scale factor $R(t)$.

The evolution of the cosmic scale factor and of the stress energy in the Universe are governed by the Friedmann equations:

$$H^2 \equiv (\dot{R}/R)^2 = 8\pi G\varrho/3 - k/R^2, \tag{4}$$

$$d(\varrho R^3) = -p d(R^3), \tag{5}$$

where ϱ is the total energy density and p is the isotropic pressure. (The assumption of isotropy and homogeneity require that the stress-energy tensor take on the perfect fluid form: T^μ_ν = diagonal $(-\varrho, p, p, p)$.) Because $\varrho \propto R^{-n}$ ($n = 3$ for matter, $n = 4$ for radiation) it follows from Eq. (4) that model Universes with $k < 0$ expand forever, while those with $k > 0$ must necessarily recollapse.

The expansion rate H (also known as the Hubble parameter) sets the characteristic imescale for the growth of $R(t)$: H^{-1} is the e-folding time for R. The present value of H is

$$H = 100h \text{ km sec}^{-1} \text{ Mpc};$$

where the observational data strongly suggest that $0.4 \leqslant h \leqslant 1$ (Ref. [7]).

The sign of the spatial curvature k — and the ultimate fate of the Universe can be determined from measurements of ϱ and H:

$$k/H^2R^2 = \varrho/(3H^2/8\pi G) - 1 \equiv \Omega - 1, \tag{6}$$

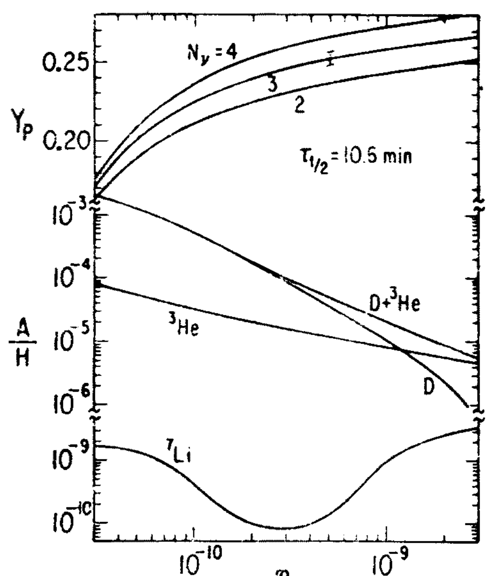


Fig. 3. Big bang nucleosynthesis predictions for the primordial abundances of D, ^{3}He, ^{4}He, and ^{7}Li. Y_p = mass fraction of ^{4}He, shown for $N_\nu = 2, 3, 4$ light neutrino species. Present observational data suggest: $0.23 \leqslant Y_p \leqslant 0.25$, $(D/H)_p \gg 1 \times 10^{-5}$, $[(D+{}^3He)/H]_p \leqslant 10^{-4}$, and $({}^7Li/H)_p \simeq (1.1 \pm 0.4) \times 10^{-10}$. Concordance requires $\eta \simeq (4-7) \times 10^{-10}$. For further discussion see Ref. [6]

where $\Omega = \varrho/\varrho_{\rm crit}$ and $\varrho_{\rm crit} = 1.88\,h^2 \times 10^{-29}$ gcm^{-3} $= 1.05 \times 10^4\,h^2$ eV cm^{-3}. The curvature radius, $R_{\rm curv}$, is related to Ω by

$$(R_{\rm curv}/H^{-1})^2 = 1/(\Omega - 1). \tag{7}$$

A reliable and definitive determination of Ω has thus far eluded cosmologists. Based upon the luminous matter in the Universe (which is relatively easy to keep track of) we can set a lower bound to Ω

$$\Omega \geqslant \Omega_{\rm LUM} \simeq 0.01.$$

Based on dynamical techniques — which all basically involve Kepler's third law in one guise or another, the observational data seem to indicate that the material that clusters with visible galaxies on scales $\leqslant$10–30 Mpc accounts for

$$\Omega_{\rm GAL} \simeq 0.1 - 0.3.$$

Although Ω can, in principle, be determined by measurements of the deceleration parameter q_0

$$q_0 \equiv -(\ddot{R}/R)/H^2 = \Omega(1+3p/\varrho)/2, \tag{8}$$

because of the difficulty of reliably determining q_0, the observations probably only restrict Ω to be less than a few [7]. (For a more thorough discussion of the amount of matter in the Universe see Ref. [8].)

The best upper limit to Ω comes from the age of the Universe. The age of the Universe is related to the Hubble time H^{-1} by

$$t_{\rm U} = f(\Omega) H_0^{-1}, \tag{9}$$

where $f(\Omega)$ is a monotonically decreasing function of Ω; $f(0) = 1$ and $f(1) = 2/3$ for a matter--dominated Universe and 1/2 for a radiation-dominated Universe. The dating of the oldest stars and the elements strongly suggest that the Universe is at least 10 Byr old — the best estimate being around 15 Byr old [9]. From Eq. (9) and $t_{\rm U} \geqslant t_{10}$ 10 Byr it follows that $\Omega f^2/t_{10}^2 \geqslant \Omega h^2$. The function Ωf^2 is monotonically increasing and bounded above by $\pi^2/4$, implying that *independent* of h, $\Omega h^2 \leqslant 2.5/t_{10}^2$. Requiring $h \geqslant 0.4$ and $t_{10} \geqslant 1$, it follows that $\Omega h^2 \leqslant 1.1$ (see Fig. 4).

The energy density of the Universe quite naturally splits up into that contributed by relativistic particles — today the microwave photons and cosmic neutrino backgrounds, and that contributed by non-relativistic particles — baryons and whatever else! The energy density contributed by non-relativistic particles decreases as $R(t)^{-3}$ — just due to the increase in the proper volume of the Universe, while that of relativistic particles varies as $R(t)^{-4}$ — the additional factor of R being due to the fact that the momenta of relativistic particles are redshifted by the expansion. (Both of these results follow directly from Eq. (5).)

The energy density contributed by relativistic particles at temperature T is

$$\varrho_{\rm R} = g_*(T)\frac{\pi^2}{30}T^4, \tag{10}$$

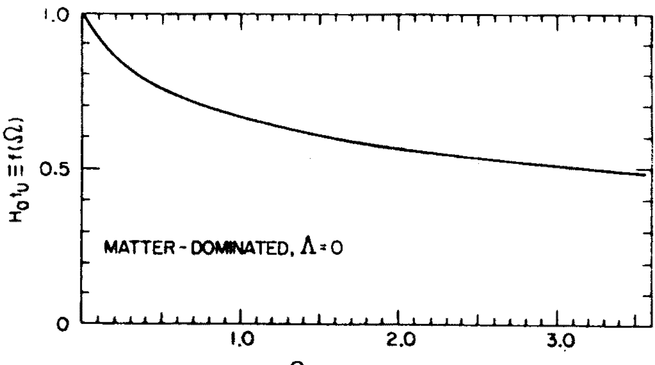


Fig. 4a. The age of a matter-dominated, $\Lambda = 0$ model universe in Hubble units, $f(\Omega) = H_0 t_U$, as a function of Ω

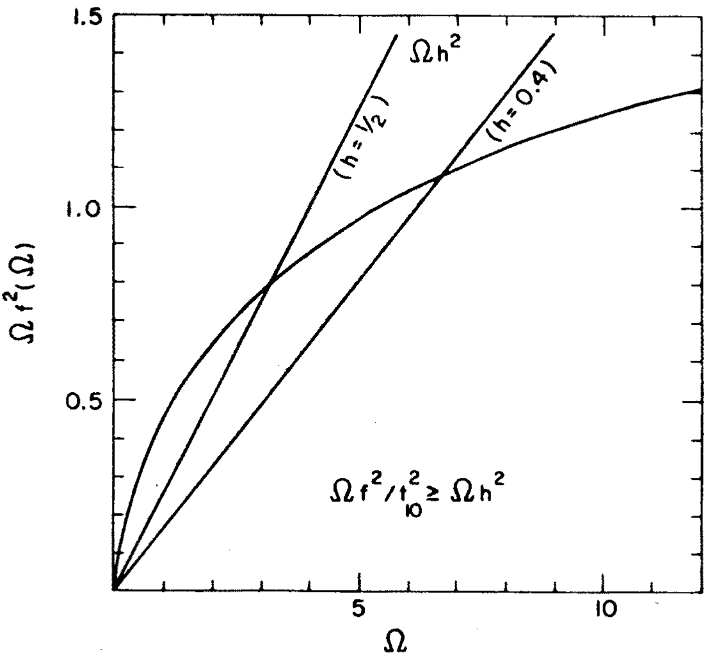


Fig. 4b. The functions Ωf^2 and Ωh^2 ($h = 0.4$ and 0.5). The function $\Omega f^2/t_{10}^2$ bounds Ωh^2 from above. For $t_{10} \gg 1$ and $h \gg 0.4(0.5)$, this implies $\Omega h^2 \ll 1.1(0.8)$. The age of the Universe $t_U = t_{10} 10$ Gyr

where $g_*(T)$ counts the effective number of degrees of freedom (weighted by their temperature) of all the relativistic particle species (those with $m \ll T$):

$$g_*(T) = \sum_{\text{Bose}} g_B (T_i/T)^4 + 7/8 \sum_{\text{Fermi}} g_F (T_i/T)^4, \tag{11}$$

here T_i is the temperature of the species i, and T is the photon temperature.

Today the energy density contributed by relativistic particles (photons and three neutrino species) is very small ($g_* = 3.36$)

$$\Omega_{\gamma 3\nu} h^2 \simeq 4\times 10^{-5} T_{2.7}^4.$$

However, because $\varrho_R \propto R^{-4}$, while $\varrho_{NR} \propto R^{-3}$, at early times the energy density contributed by relativistic particles dominated that of non-relativistic particles. To be specific, the Universe was radiation-dominated for

$$\begin{aligned} t &\leqslant t_{EQ} \simeq 4\times 10^{10}\ \text{sec}\,(\Omega h^2)^{-2} T_{2.7}^6, \\ R &\leqslant R_{EQ} \simeq 4\times 10^{-5} R_{\text{today}} (\Omega h^2)^{-1} T_{2.7}^4, \\ T &\geqslant T_{EQ} \simeq 5.8\ \text{eV}\ \Omega h^2 T_{2.7}^3. \end{aligned}$$

Therefore, at very early times Eq. (4) simplifies to

$$\begin{aligned} H &= (\dot{R}/R) = (4\pi^3 g_*/45)^{1/2} T^2/m_{Pl}, \\ &= 1.66 g_*^{1/2} T^2/m_{Pl}. \end{aligned} \tag{12}$$

(Note since the curvature term varies as $R(t)^{-2}$ it too is negligible compared to the energy density in relativistic particles.) For reference, $g_*(\text{few MeV}) = 10.75$ (γ, $e^\pm$, $3\nu\bar{\nu}$); $g_*(100\ \text{GeV}) \simeq 110$ (γ, 8G, $W^\pm Z$, 3 families of quarks and leptons, and 1 Higgs doublet).

So long as thermal equilibrium is maintained, the second Friedmann equation, Eq. (5), implies that the entropy per comoving volume, $S \propto sR^3$, remains constant. Here s is the entropy density which is dominated by the contribution from relativistic particles, and is

$$s = (\varrho + p)/T \simeq (2\pi^2/45) g_* T^3. \tag{13}$$

The entropy density is just proportional to the number density of relativistic particles. Today the entropy density is just 7.04 times the number density of photons. The constancy of S means that $s \propto R^{-3}$, or that the ratio of any number density to s is just proportional to the number of that species per comoving volume. The baryon number-to-entropy ratio is

$$n_B/s \simeq (1/7)\eta \simeq (6-10)\times 10^{-11}$$

and since today the number density of baryons is much greater than that of antibaryons, this ratio is also the net baryon number per comoving volume — which is conserved so long as the rate of baryon-number non-conserving reactions is small.

The constancy of S implies that

$$T \propto g_*(T)^{-1/3} R(t)^{-1}. \tag{14}$$

Whenever g_* is constant, this means that $T \propto R(t)^{-1}$. Together with Eq. (12) this gives

$$\begin{aligned} R(t) &= R(t_0)\,(t/t_0)^{1/2}, \\ t &\simeq 1/2\, H^{-1} \simeq 0.3 g_*^{-1/2} m_{Pl}/T^2, \\ &\simeq 2.4\times 10^{-6}\ \text{sec}\ g_*^{-1/2} (T/\text{GeV})^{-2}. \end{aligned} \tag{15}$$

Finally, let me mention one more important feature of the standard cosmology, the existence of particle horizons. In the standard cosmology the distance a photon could have traveled since the bang is finite, meaning that at a given epoch the Universe is comprised of many causally-distinct domains. Photons travel on paths characterized by $ds^2 = 0$; for simplicity and without loss of generality consider a trajectory with $d\varphi = d\theta = 0$. The coordinate distance traversed by a photon since 'the bang' is

$$\int_0^t dt'/R(t')$$

which corresponds to the physical distance (measured at time t)

$$d_H(t) = R(t) \int_0^t dt'/R(t'). \tag{16}$$

If $R(t) \propto t^n$ and $n < 1$, then the horizon distance $d_H(t)$ is finite and $d_H(t) = t/(1-n) = nH^{-1}/(1-n) \simeq t$.

Note that even if $d_H(t)$ diverges (e.g., if $R(t) \propto t^n$ with $n > 1$), the Hubble radius H^{-1} still sets the scale of the 'Physics Horizon'. All physical distances scale with $R(t)$. Thus microphysical processes operating on a timescale $\gtrsim H^{-1}$ will have their effects distorted by the expansion, strongly suggesting that a coherent microphysical process can only operate over a time interval of order H^{-1}. Then, causally-coherent microphysical processes can only operate on distances $\leqslant$ the Hubble radius, H^{-1}. The intuitive notion that the Hubble radius acts as the 'Physics Horizon' is borne out quantitatively time and time again, and so it is useful to think of H^{-1} as the maximum scale for microphysical processes.

During the radiation-dominated era $n = 1/2$ and $d_H(t) = 2t$; the entropy and baryon number within the horizon at a given time are easily computed:

$$\begin{aligned} S_{\rm HOR} &= (4\pi/3)t^3 s, \\ &\simeq 0.05 g_*^{-1/2}(m_{\rm Pl}/T)^3, \\ N_{\rm B-HOR} &= (n_{\rm B}/s)S_{\rm HOR}, \\ &\simeq 10^{-12}(m_{\rm Pl}/T)^3, \\ &\simeq 10^{-2} M_\odot (T/\text{MeV})^{-3}. \end{aligned}$$

We can compare these numbers to the entropy and baryon number contained within the present horizon volume:

$$S_{\rm U} \simeq 10^{88}, \quad N_{\rm BU} \simeq 10^{78}.$$

Evidently, in the standard cosmology the comoving volume which corresponds to the part of the Universe which is presently observable contained many, many horizon volumes at early times. This is an important point to which we shall return shortly.

3. Shortcomings of the standard cosmology

The standard cosmology is very successful — it provides us with a reliable framework for describing the history of the Universe as early as 10^{-2} sec after the bang (when the temperature was about 10 MeV) and perhaps as early as 10^{-43} sec after the bang (see Fig. 5). In sum, the standard cosmology is a great achievement. There is nothing in our present understanding of physics that would indicate that it is incorrect to extrapolate the standard cosmology back to times as early as 10^{-43} sec — the fundamental constituents of matter, quarks and leptons, are point-like particles and their known interactions should remain 'weak' up to energies as high as 10^{19} GeV — justifying the dilute gas approximation made in writing $\varrho_R \propto T^4$. (This fact was first pointed out by Collins and Perry [9a]. However, at times earlier than 10^{-43} sec, corresponding to temperatures greater than 10^{19} GeV, quantum corrections to general relativity — a classical theory, should become very significant.)

However, it is not without its shortcomings. There are a handful of very important and fundamental cosmological facts which, while it can accommodate, it in no way elucidates. I will briefly review these puzzling facts.

(1–2) Large-scale isotropy and homogeneity

The observable Universe (size $\simeq H^{-1} \simeq 10^{28}$ cm $\simeq 3000\, h^{-1}$ Mpc) is to a high degree of precision isotropic and homogeneous on the largest scales, say > 100 Mpc. (Of course, our knowledge of the Universe outside our past light cone is very limited; see Ref. [10].) The best evidence for the isotropy and homogeneity is provided by the uniformity of the cosmic background temperature (see Fig. 2): $(\delta T/T) < 10^{-3}$ (10^{-4} if the dipole anisotropy is interpreted as being due to our motion relative to the cosmic rest frame). Large-scale density inhomogeneities or anisotropic expansion would result in temperature fluctuations of comparable magnitude (see Refs [11, 12]). The smoothness of the observed Universe is puzzling if one wishes to understand it as being due to causal, microphysical processes which operated during the early history of the Universe. Our Hubble volume today contains an entropy of about 10^{88}. At decoupling ($t \simeq 6\times 10^{12}\, (\Omega h^2)^{-1/2}$ sec, $T \simeq 1/3$ eV), the last epoch when matter and radiation were known to be interacting vigorously and particle interactions might have been able to smooth the radiation, the entropy within the horizon was only about 8×10^{82}; that is, the comoving volume which contains the presently-observable Universe, then was comprised of about 2×10^5 causally-distinct regions. How is it that they came to be homogeneous? Put another way, the particle horizon at decoupling only subtends an angle of about 1/2° on the sky today — how is it that the cosmic background temperature is so uniform on angular scales much greater than this?

The standard cosmology can accommodate these facts — after all the FRW cosmology is exactly isotropic and homogeneous, but at the expense of very special initial data. In 1973 Collins and Hawking [13] showed that the set of initial data which evolve to a Universe which globally is as smooth as ours has measure zero (provided that the strong and dominant energy conditions are always satisfied).

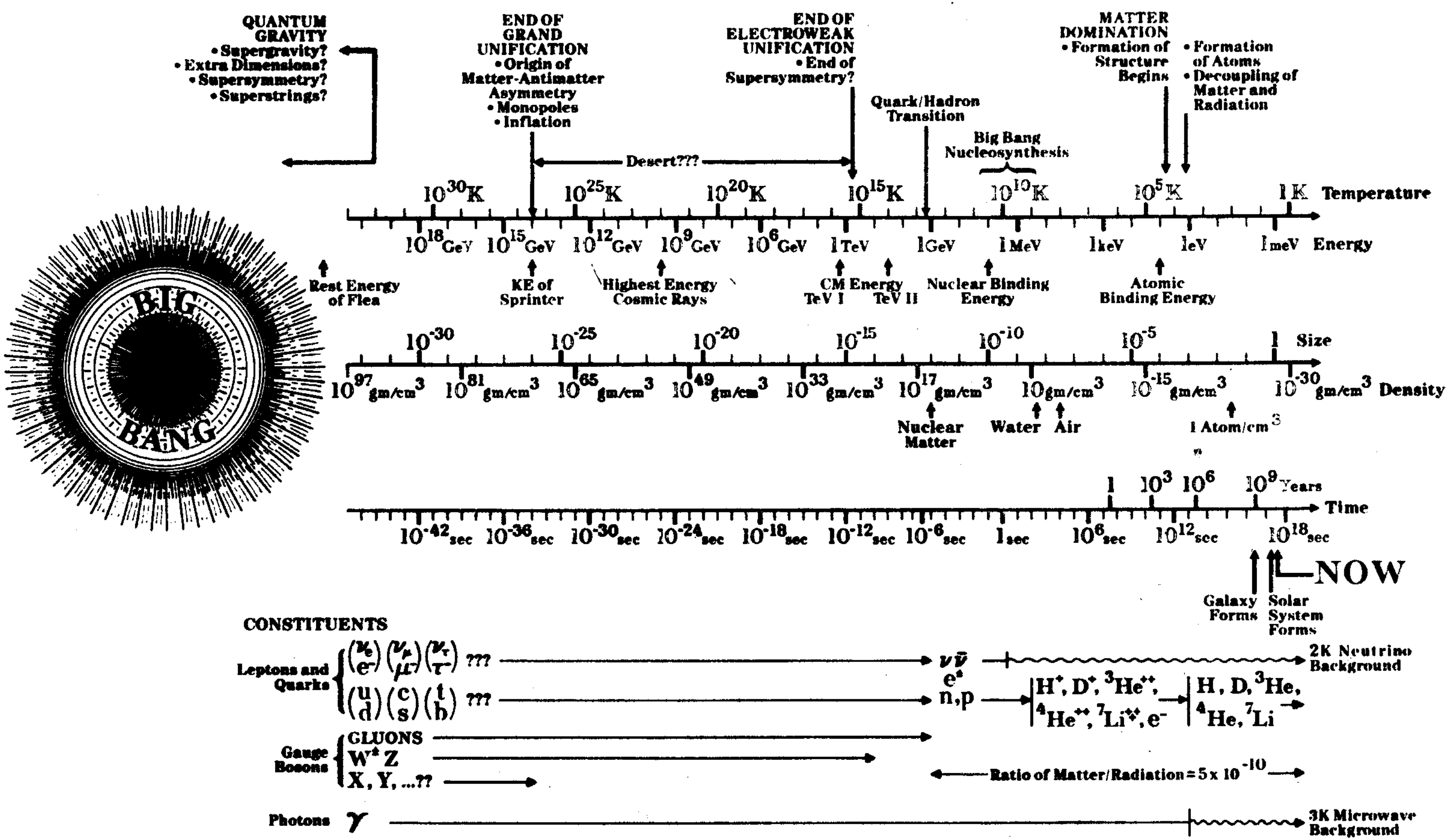


Fig. 5. History of the Universe according to the standard cosmology and currently fashionable ideas in particle theory

(3) Small-scale inhomogeneity

As any real astronomer will gladly testify, the Universe is very lumpy — stars, galaxies, clusters of galaxies, superclusters, etc. Today, the density contrast on the scale of galaxies is: $\delta\varrho/\varrho \simeq 10^5$. The fact that the microwave background radiation is very uniform even on very small angular scales ($\ll 1°$) indicates that the Universe was smooth even on the scale of galaxies at decoupling. (The relationship between the angle on the sky and mass contained within the corresponding length scale at decoupling is: $\theta \simeq 1'(M/10^{12}\,M_\odot)^{1/3} \times \Omega^{-1/3}h^{1/3}$.) On small angular scales: $\delta T/T \simeq c(\delta\varrho/\varrho)_{\rm dec}$, where the numerical constant $c \simeq 10^{-1} - 10^{-2}$ (see Ref. [12] for further details). Whence came the structure which today is so conspicuous?

Once matter decouples from the radiation and is free of the pressure support provided by the radiation, any density inhomogeneities present will grow via the Jeans (or gravitational instability) — in the linear regime, $\delta\varrho/\varrho \propto R(t)$. (If the mass density of the Universe is dominated by a collisionless particle species, e.g., a light, relic neutrino species or relic axions, density perturbations in these particles can begin to grow as soon as the Universe becomes matter-dominated.) In order to account for the present structure, density perturbations of amplitude $\sim 10^{-3}$ or so at decoupling are necessary on the scale of galaxies. The standard cosmology sheds no light as to the origin or nature (spectrum and type — adiabatic or isothermal) of the primordial density perturbations so crucial for understanding the structure observed in the Universe today. (For a review of the formation of structure in the Universe according to the gravitational instability picture, see Ref. [14].)

(4) Flatness (or oldness) of the Universe

The observational data suggest that

$$0.01 \leqslant \Omega \leqslant \text{few}.$$

Ω is related to both the expansion rate of the Universe and the curvature radius of the Universe:

$$\Omega = 8\pi G\varrho/3H^2 \equiv H_{\rm crit}^2/H^2, \tag{17}$$

$$|\Omega - 1| = (H^{-1}/R_{\rm curv})^2. \tag{18}$$

The fact that Ω is not too different from unity today implies that the present expansion rate is close to the critical expansion rate and that the curvature radius of the Universe is comparable to or larger than the Hubble radius. As the Universe expands Ω does not remain constant, but evolves away from 1

$$\Omega = 1/(1 - x(t)), \tag{19}$$

$$x(t) = (k/R^2)/(8\pi G\varrho/3), \tag{20}$$

$$x(t) \propto \begin{cases} R(t)^2 & \text{radiation-dominated} \\ R(t) & \text{matter-dominated.} \end{cases}$$

That Ω is still of order unity means that at early times it was equal to 1 to a very high degree of precision:

$$|\Omega(10^{-43}\text{ sec})-1| \simeq O(10^{-60}),$$

$$|\Omega(1\text{ sec})-1| \simeq O(10^{-16}).$$

This in turn implies that at early times the expansion rate was equal to the critical rate to a high degree of precision and that the curvature of the Universe was much, much greater than the Hubble radius. If it were not, i.e., suppose that $|(k/R^2)/(8\pi G\varrho/3)| \simeq O(1)$ at $t \simeq 10^{-43}$ sec, then the Universe would have collapsed after a few Planck times ($k > 0$) or would have quickly become curvature-dominated, ($k < 0$), in which case $R(t) \propto t$ and $t(T = 3\text{ K}) \leqslant 10^{-11}$ sec! Why was this so?

The co-called flatness problem has sometimes been obscured by the fact that it is conventional to rescale $R(t)$ so that $k = -1$, 0, or $+1$, making it seem as though there are but three FRW models. However, that clearly is not the case; there are an infinity of models, specified by the curvature radius $R_{\text{curv}} = R(t)|k|^{-1/2}$ at some given epoch, say the Planck epoch. Our model corresponds to one with a curvature radius that exceeds its initial Hubble radius by 30 orders-of-magnitude. Again, this fact can be accommodated by FRW models, but the extreme flatness of our Universe is in no way explained by the standard cosmology. (The flatness problem and the naturalness of the $k = 0$ model have been emphasized by Dicke and Peebles [14a].)

(5) Baryon number of the Universe

There is ample evidence (see Ref. [15]) for the dearth of antimatter in the observable Universe. That fact together with the baryon-to-photon ratio ($\eta \simeq 4-7\times 10^{-10}$) means that our Universe is endowed with a net baryon number, quantified by the baryon number--to-entropy ratio

$$n_B/s \simeq (6-10)\times 10^{-11},$$

which in the absence of baryon number non-conserving interactions or significant entropy production is proportional to the constant net baryon number per comoving volume which the Universe has always possessed. Until five or so years ago this very fundamental number was without explanation. Of course it is now known that in the presence of interactions that violate B, C, and CP a net baryon asymmetry will evolve dynamically. Such interactions are predicted by Grand Unified Theories (or GUTs) and 'baryogenesis' is one of the great triumphs of the marriage of grand unification and cosmology. (See Ref. [16] for a review of grand unification.) If the baryogenesis idea is correct, then the baryon asymmetry of the Universe is subject to calculation just as the primordial Helium abundance is. Although the idea is very attractive and certainly appears to be qualitatively correct, a precise calculation of the baryon number-to-entropy ratio cannot be performed until *The* Grand Unified Theory is known. (Baryogenesis is reviewed in Ref. [17].)

(6) The monopole problem

If the great success of the marriage of GUTs and cosmology is baryogenesis, then the great disappointment is 'the monopole problem'. 't Hooft-Polyakov monopoles [18] are a generic prediction of GUTs. In the standard cosmology (and for the simplest GUTs) monopoles are grossly overproduced during the GUT symmetry-breaking transition, so much so that the Universe would reach its present temperature of 3 K at the very tender age of 30,000 yrs! (For a detailed discussion of the monopole problem, see Refs [19, 20].) Although the monopole problem initially seemed to be a severe blow to the Inner Space/Outer Space connection, as it has turned out it provided us with a valuable piece of information about physics at energies of order 10^{14} GeV and the Universe at times as early as 10^{-34} sec — the standard cosmology and the simplest GUTs are definitely incompatible! As it turned out, it was the search for a solution to the monopole problem which in the end led Guth to come upon the inflationary Universe scenario [21, 22].

(7) The smallness of the cosmological constant

With the possible exception of supersymmetry/supergravity and superstring theories, the absolute scale of the scalar potential $V(\phi)$ is not specified (here ϕ represents the scalar fields in the theory, be they fundamental or composite). A constant term in the scalar potential is equivalent to a cosmological term (the scalar potential contributes a term $Vg_{\mu\nu}$ to the stress energy of the Universe [23]. At low temperatures (say temperatures below any scale of spontaneous symmetry-breaking) the constant term in the potential receives contributions from all the stages of SSB — chiral symmetry breaking, electroweak SSB, GUT SSB, etc. The observed expansion rate of the Universe ($H = 100h$ km sec^{-1} Mpc^{-1}) limits the total energy density of the Universe to be

$$\varrho_{\text{TOT}} \leqslant O(10^{-46}\ \text{GeV}^4).$$

Making the seemingly very reasonable assumption that all stress energy selfgravitates (which is dictated by the equivalence principle) it follows that the vacuum energy of our

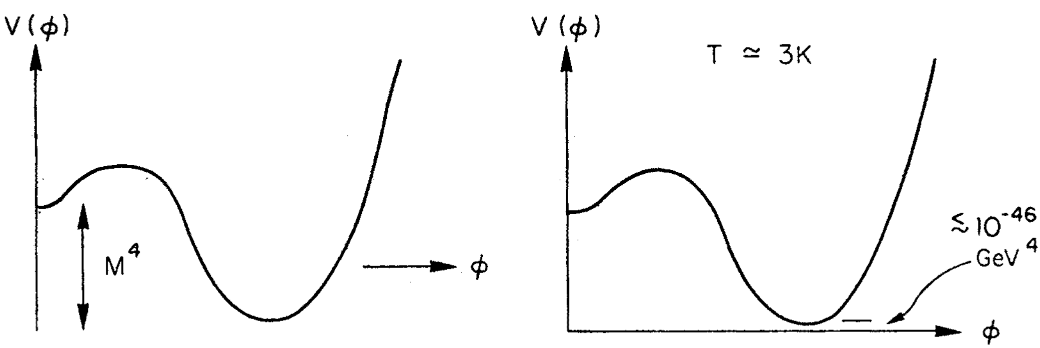


Fig. 6. In gauge theories the vacuum energy is a function of one or more scalar fields (here denoted collectively as ϕ); however, the absolute energy scale is not set. Vacuum energy behaves like a cosmological term; the present expansion rate of the Universe constrains the value of the vacuum energy today to be $\leqslant 10^{-46}$ GeV^4

SU(3) × U(1) vacuum must be less than 10^{-46} GeV^4. Compare this to the scale of the various contributions to the scalar potential: $O(M^4)$ for physics associated with a symmetry breaking scale of M

$$V_{\text{today}}/M^4 \leqslant \varrho_{\text{TOT}}/M^4 \leqslant \begin{cases} 10^{-122} & M = m_{\text{Pl}} \\ 10^{-102} & M = 10^{14}\ \text{GeV} \\ 10^{-56} & M = 300\ \text{GeV} \\ 10^{-46} & M = 1\ \text{GeV}. \end{cases}$$

At present there is no explanation for the vanishingly small value of the energy density of our very unsymmetrical vacuum. It is easy to speculate that a fundamental understanding of the smallness of the cosmological constant will likely involve an intimate link between gravity and quantum field theory.

Today we can be certain the vacuum energy is small and plays a minor role in the dynamics of the expansion of the Universe (compared to the potential role that it could play). If we accept this as an empirical determination of the absolute scale of the scalar potential $V(\phi)$, it then follows that the energy density associated with an expectation value of ϕ near zero is enormous — of order M^4 (see Fig. 6) and therefore could have played an important role in the dynamics of the very early Universe. Accepting this *empirical determination* of the zero of vacuum energy — which is a very great leap of faith, is the starting point for inflation. In fact, the rest of the journey is downhill.

All of these cosmological facts can be accommodated by the standard model, but seemingly at the expense of highly special initial data (the possible exception being the monopole problem). Over the years there have been a number of attempts to try to understand and/or explain this apparent dilemma of initial data. Inflation is the most recent attempt and I believe shows great promise. Let me begin by briefly mentioning the earlier attempts:

* Mixmaster Paradigm — Starting with a solution with a singularity which exhibits the features of the most general singular solutions known (the so-called mixmaster model) Misner and his coworkers hoped that they could show that particle viscosity would smooth out the geometry. In part because horizons still effectively exist in the mixmaster solution 'the chaotic cosmology program' has proven unsuccessful (for further discussion see Ref. [24]).

* Nature of the Initial Singularity — Penrose [25] explored the possibility of explaining the observed smoothness of the Universe by restricting the kinds of initial singularities which are permitted in Nature (those with vanishing Weyl curvature). In a sense his approach is to postulate a law of physics governing allowed initial data.

* Quantum Gravity Effects — The first two solutions involve appealing to classical gravitational effects. A number of authors have suggested that quantum gravity effects might be responsible for smoothing out the space-time geometry (de Witt [26], Parker [27]; Zel'dovich [28]; Starobinskii [29]; Anderson [30]; Hartle and Hu [31]; Fischetti et al. [32]). The basic idea being that anisotropy and/or inhomogeneity would drive gravitational particle creation, which due to back reaction effects would eliminate particle horizons

and smooth out the geometry. Recently, Hawking and Hartle [33] have advocated the Quantum Cosmology approach to actually compute the initial state. All of these approaches necessarily involve events at times $\lesssim 10^{-43}$ sec and energy densities $\gtrsim m_{\rm Pl}^4$.

* Anthropic Principle — Some (see, e.g., Ref. [34]) have suggested (or in some cases even advocated) 'explaining' many of the puzzling features of the Universe around us (and in some cases, even the laws of physics!) by arguing that unless they were as they are intelligent life would not have been able to develop and observe them! Hopefully we will not have to resort to such an explanation.

The approach of inflation is somewhat different from previous approaches. Inflation (at least from my point-of-view) is based upon well-defined and reasonably well-understood microphysics (albeit, some of it very speculative). That microphysics is:

* Classical Gravity (general relativity), at least as an effective, low-energy theory of gravitation.

* 'Modern Particle Physics' — grand unification, supersymmetry/supergravity, field theory limit of superstring theories, etc. at energy scales $\lesssim m_{\rm Pl}$.

As I will emphasize, in all viable models of inflation the inflationary period (at least the portion of interest to us) takes place well after the planck epoch, with the energy densities involved being far less than $m_{\rm Pl}^4$ (although semi-classical quantum gravity effects might have to be included as non-renormalizable terms in the effective Lagrangian). Of course, it could be that a resolution to the cosmological puzzles discussed above involves both 'modern particle physics' and quantum gravitational effects in their full glory (as in a fully ten dimensional quantum theory of strings).

I will not take the time or the space here to review the historical development of our present view of inflation; I refer the interested reader to the interesting paper on this subject by Lindley [35]. It suffices to say that Guth's very influential paper of 1981 was the 'shot heard round the world' which initiated the inflation revolution [22], and that Guth's doomed original model (see Guth and Weinberg [36], Hawking et al. [36]) was revived by Linde's [37] and Albrecht and Steinhardt's [38] variant, 'new inflation'. I will focus all of my attention on the present status of the 'slow-rollover' model of Linde [37] and Albrecht and Steinhardt [38].

4. *New inflation — the slow-rollover transition*

The basic idea of the inflationary Universe scenario is that there was an epoch when the vacuum energy density dominated the energy density of the Universe. During this epoch $\varrho \simeq V \simeq$ constant, and thus $R(t)$ grew exponentially ($\propto \exp(Ht)$), allowing a small, causally-coherent region (initial size $\lesssim H^{-1}$) to grow to a size which encompasses the region which eventually becomes our presently-observable Universe. In Guth's original scenario [22], this epoch occurred while the Universe was trapped in the false ($\phi = 0$) vacuum during a strongly, first-order phase transition. In new inflation, the vacuum-dominated, inflationary epoch occurs while the region of the Universe in question is slowly, but inevitably, evolving toward the true, SSB vacuum. Rather than considering specific models in this section, I will try to discuss new inflation in the most general context. For the moment I will however assume that the epoch of inflation is associated with a first-order, SSB phase transition, and that the Universe is in thermal equilibrium before the transition. As we shall see later new inflation is more general than these assumptions. But for definiteness (and for historical reasons), let me begin by making these assumptions.

Consider a SSB phase transition characterized by an energy scale M. For $T \gg T_c \simeq O(M)$ the symmetric ($\phi = 0$) vacuum is favored, i.e., $\phi = 0$ is the global minimum of the finite temperature effective potential $V_T(\phi)$ (= free energy density). As T approaches T_c a second minimum develops at $\phi = 0$, and at $T = T_c$, the two minima are degenerate. At temperatures below T_c the SSB ($\phi = \sigma$) minimum is the global minimum of $V_T(\phi)$ (see Fig. 7). However, the Universe does not instantly make the transition from $\phi = 0$

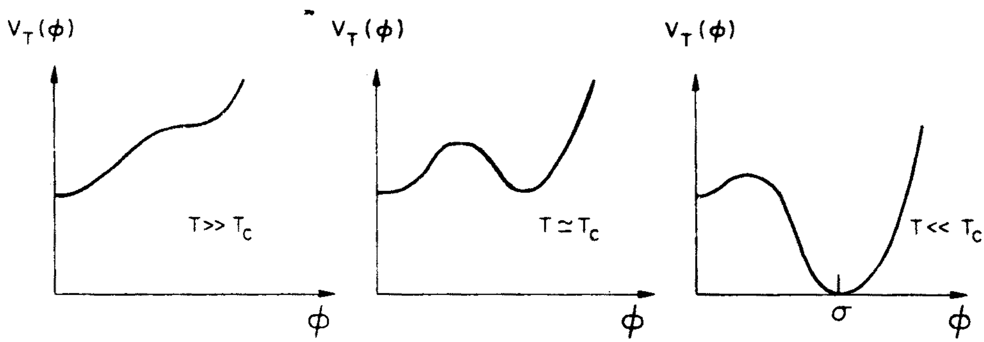


Fig. 7. The finite temperature effective potential as a function of T (schematic). The Universe is usually assumed to start out in the high temperature, symmetric minimum ($\phi = 0$) of the potential and must eventually evolve to the low temperature, asymmetric minimum ($\phi = \sigma$). The evolution of ϕ from $\phi = 0$ to $\phi = \sigma$ can prove to be very interesting — as in the case of an inflationary transition

to $\phi = \sigma$; the details and time required are a question of dynamics. (The scalar field ϕ is the order parameter for the SSB transition under discussion; in the spirit of generality ϕ might be a gauge singlet field or might have nontrivial transformation properties under the gauge group, possibly even responsible for the SSB of the GUT.) Once the temperature of the Universe drops below $T_c \simeq O(M)$, the potential energy associated with ϕ being far from the minimum of its potential, $V \simeq V(0) \simeq M^4$, dominates the energy density in radiation ($\varrho_r < T_c^4$), and causes the Universe to expand exponentially. During this exponential expansion (known as a deSitter phase) the temperature of the Universe decreases exponentially causing the Universe to supercool. The exponential expansion continues so long as ϕ is far from its SSB value. Now let's focus on the evolution of ϕ.

Assuming a barrier exists between the false and true vacua, thermal fluctuations and/or quantum tunneling must take ϕ across the barrier. The dynamics of this process determine when and how the process occurs (bubble formation, spinodal decomposition, etc.) and the value of ϕ after the barrier is penetrated. If the action for bubble nucleation remains large, $S_b \gg 1$, then the barrier will be overcome by the nucleation of Coleman-deLuccia bubbles [39]; on the other hand if the action for bubble nucleation becomes of order unity, then the Universe will undergo spinodal decomposition, and irregularly-shaped fluctuation regions will form (see Fig. 8; for a more detailed discussion of the barrier penetration process see Refs [38–40]). For definiteness suppose that the barrier is overcome when the temperature is T_{MS} and that after the barrier is penetrated the value of ϕ is ϕ_0. From this point the journey to the true vacuum is downhill (literally). For the moment let us assume that the evolution of ϕ is adequately described by semi-classical equations of motion:

$$\ddot{\phi}+3H\dot{\phi}+\Gamma\dot{\phi}+V' = 0, \tag{21}$$

where ϕ has been normalized so that its kinetic term in the Lagrangian is $1/2\partial_\mu\phi\partial^\mu\phi$, and prime indicates a derivative with respect to ϕ. The subscript T on V has been dropped; for $T \ll T_c$ the temperature dependence of V_T can be neglected and the zero temperature potential ($\equiv V$) can be used. The $3H\dot{\phi}$ term acts like a frictional force, and arises because the expansion of the Universe 'redshifts away' the kinetic energy of $\phi(\propto R^{-3})$. The $\Gamma\dot{\phi}$ term accounts for particle creation due to the time-variation of ϕ (Refs. [41, 42]). The quantity Γ is determined by the particles which couple to ϕ and the strength with which they couple ($\Gamma^{-1} \simeq$ lifetime of a ϕ particle). As usual, the expansion rate H is determined

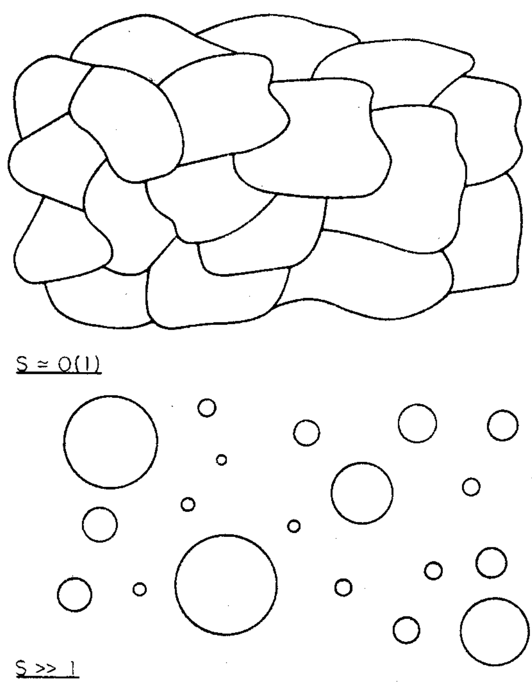


Fig. 8. If the tunneling action is large ($S \gg 1$), barrier penetration will proceed via bubble nucleation, while in the case that it becomes small ($S \simeq O(1)$), the Universe will fragment into irregularly-shaped fluctuation regions. The very large scale (scale $\gg$ bubble or fluctuation region) structure of the Universe is determined by whether $S \simeq O(1)$ — in which case the Universe is comprised of irregularly-shaped domains, or $S \gg O(1)$ — in which case the Universe is comprised of isolated bubbles

by the energy density of the Universe:

$$H^2 = 8\pi G\varrho/3, \tag{22}$$

$$\varrho \simeq 1/2\dot{\phi}^2 + V(\phi) + \varrho_r, \tag{23}$$

where ϱ_r represents the energy density in radiation produced by the time variation of ϕ. (For $T_{MS} \ll T_c$ the original thermal component makes a negligible contribution to ϱ.) The evolution of ϱ_r is given by

$$\dot{\varrho}_r + 4H\varrho_r = \Gamma\dot{\phi}^2, \tag{24}$$

where the $\Gamma\dot{\phi}^2$ term accounts for particle creation by ϕ.

In writing Eqs (21–24) I have implicitly assumed that ϕ is spatially homogeneous. In some small region (inside a bubble or a fluctuation region) this will be a good approximation. The size of this smooth region will turn out to be unimportant; take it to be of order the 'Physics Horizon', H^{-1} — certainly, it is not likely to be larger. Now follow the evolution of ϕ within the small, smooth patch of size H^{-1}.

If $V(\phi)$ is sufficiently flat somewhere between $\phi = \phi_0$ and $\phi = \sigma$, then ϕ will evolve very slowly in that region, and the motion of ϕ will be 'friction-dominated' so that $3H\dot{\phi} \simeq -V'$ (in the slow growth phase particle creation is not important [43]). If V is sufficiently flat, then the time required for ϕ to transverse the flat region can be long compared to the expansion timescale H^{-1}; for definiteness say, $\tau_\phi = 100\,H^{-1}$. During this slow growth phase $\varrho \simeq V(\phi) \simeq V(\phi = 0)$; both ϱ_r and $1/2\dot{\phi}^2$ are $\ll V(\phi)$. The expansion rate H is then just

$$H \simeq (8\pi V(0)/3m_{Pl}^2)^{1/2} \simeq O(M^2/m_{Pl}), \tag{25}$$

where $V(0)$ is assumed to be of order M^4. While $H \simeq$ constant, R grows exponentially: $R \propto \exp(Ht)$; for $\tau_\phi = 100H^{-1}$, R expands by a factor of e^{100} during the slow rolling period, and the physical size of the smooth region increases to $e^{100}H^{-1}$.

As the potential steepens, the evolution of ϕ quickens. Near $\phi = \sigma$, ϕ oscillates around the SSB minimum with frequency m_ϕ: $m_\phi^2 \simeq V''(\sigma) \simeq O(M^2) \gg H^2 \simeq M^4/m_{Pl}^2$. As ϕ oscillates about $\phi = \sigma$ its motion is damped both by particle creation and the expansion of the Universe. If $\Gamma^{-1} \ll H^{-1}$, then coherent field energy density $(V + 1/2\dot{\phi}^2)$ is converted into radiation in less than an expansion time ($\Delta t_{RH} \simeq \Gamma^{-1}$), and the patch is reheated to a temperature $T \simeq O(M)$ — the vacuum energy is efficiently converted into radiation ('good reheating'). On the other hand, if $\Gamma^{-1} \gg H^{-1}$, then ϕ continues to oscillate and the coherent field energy redshifts away with the expansion: $(V + 1/2\dot{\phi}^2) \propto R^{-3}$ — the coherent energy behaves like non-relativistic matter. Eventually, when $t \simeq \Gamma^{-1}$ the energy in radiation begins to dominate that in coherent field oscillations, and the patch is reheated to a temperature $T \simeq (\Gamma/H)^{1/2}M \simeq (\Gamma m_{Pl})^{1/2} \ll M$ ('poor reheating'). The evolution of ϕ is summarized schematically in Fig. 9. In the next section I will discuss the all-important scalar field dynamics in great detail.

For the following discussion let us assume 'good reheating' ($\Gamma \gg H$). After reheating the patch has a physical size $e^{100}H^{-1}$ ($\simeq 10^{17}$ cm for $M \simeq 10^{14}$ GeV), is at a temperature of order M, and in the approximation that ϕ was initially constant throughout the patch, the patch is exactly smooth. From this point forward the region evolves like a radiation-dominated FRW model. How have the cosmological conundrums been 'explained'?

First, *the homogeneity and isotropy*; our observable Universe today ($\simeq 10^{28}$ cm) had a physical size of about 10 cm ($= 10^{28}$ cm $\times$ 3 K/10^{14} GeV) when T was 10^{14} GeV — thus it lies well within one of the smooth regions produced by the inflationary epoch. Put another way, inflation has resulted in a smooth patch which contains an entropy of order $(10^{17}\text{ cm})^3 \times (10^{14}\text{ GeV})^3 \simeq 10^{134}$, which is much, much greater than that within

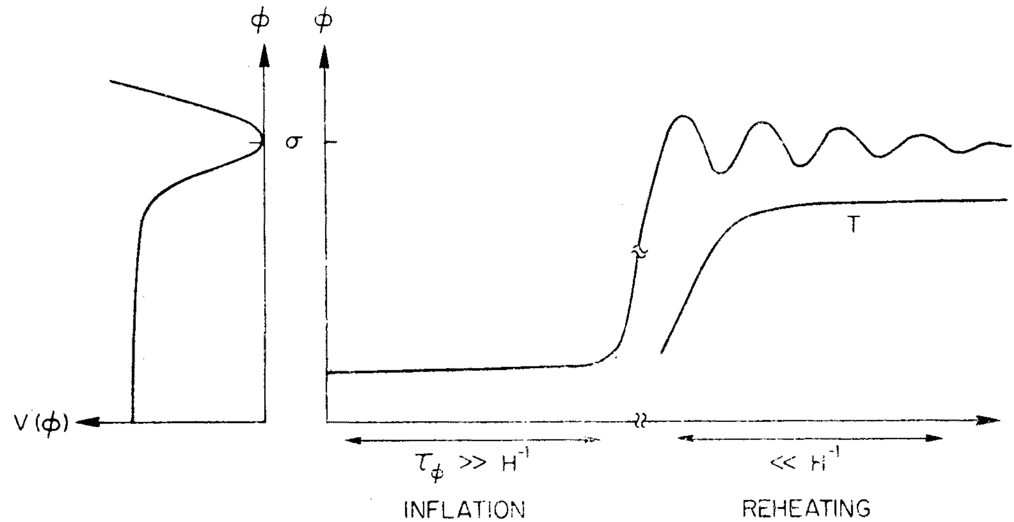


Fig. 9. Evolution of ϕ and the temperature inside the bubble or fluctuation region (schematic). Early on ϕ evolves slowly (relative to the expansion timescale), then as the potential steepens ϕ evolves rapidly (on the expansion timescale). The oscillations of ϕ are damped by particle creation, which leads to the reheating of the bubble or fluctuation region

the presently-observed Universe ($\simeq 10^{88}$). Before inflation that same volume contained only a very small amount of entropy, about $(10^{-23}\text{ cm})^3$ $(10^{14}\text{ GeV})^3 \simeq 10^{14}$. The key to inflation then is the highly nonadiabatic event of reheating (see Fig. 10). The very large-scale cosmography depends upon the state of the Universe before inflation and how inflation was initiated (bubble nucleation or spinodal decomposition); see Ref. [45] for further discussion.

Since we have assumed that ϕ is spatially constant within the bubble or fluctuation region, after reheating the patch in question is precisely uniform, and at this stage *the inhomogeneity puzzle* has not been solved. Inflation *has* produced a smooth manifold on which small fluctuations can be impressed. Due to deSitter space produced quantum fluctuations in ϕ, ϕ is not exactly uniform even in a small patch. Later, I will discuss the density inhomogeneities that result from the quantum fluctuations in ϕ.

The flatness puzzle involves the smallness of the ratio of the curvature term to the energy

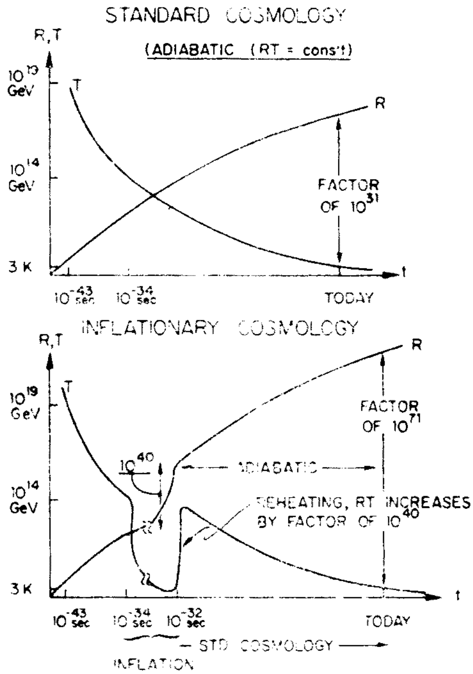


Fig. 10. Evolution of the scale factor R and temperature T of the Universe in the standard cosmology and in the inflationary cosmology. The standard cosmology is always adiabatic ($RT \simeq$ const), while the inflationary cosmology undergoes a highly non-adiabatic event (reheating) after which it is adiabatic

density term. This ratio is exponentially smaller after inflation: $x_{\text{after}} \simeq e^{-200}\, x_{\text{before}}$ since the energy density before and after inflation is $O(M^4)$, while k/R^2 has exponentially decreased (by a factor of e^{200}). Since the ratio x is reset to an exponentially small value, the inflationary scenario predicts that today Ω should be $1 \pm O(10^{-\mathrm{BIG}\neq})$.

If the Universe is reheated to a temperature of order M, a *baryon asymmetry* can evolve in the usual way, although the quantitative details may be different [17, 43]. If the Universe is not efficiently reheated ($T_{\mathrm{RH}} \ll M$), it may be possible for n_{B}/s to be produced directly in the decay of the coherent field oscillations [41–44] (which behave just like NR ϕ particles); this possibility will be discussed later. In any case, it is absolutely necessary to have baryogenesis occur after reheating since any baryon number (or any other quantum number) present before inflation is diluted by a factor of $(M/T_{\mathrm{MS}})^3 \exp(3H\tau_\phi)$ — the factor by which the total entropy increases. Note that if C, CP are violated spontaneously, then ε (and n_{B}/s) could have a different sign in different patches — leading to a Universe which on the very largest scales ($\gg e^{100}\, H^{-1}$) is baryon symmetric.

Since the patch that our observable Universe lies within was once (at the beginning of inflation) causally-coherent, the Higgs field could have been aligned throughout the patch (indeed, this is the lowest energy configuration), and thus there is likely to be $\leqslant 1$ monopole within the entire patch which was produced as a topological defect. *The glut of monopoles* which occurs in the standard cosmology does not occur. (The production of other topological defects (such as domain walls, etc.) is avoided for similar reasons.) Some monopoles will be produced after reheating in rare, very energetic particle collisions [46a]. The number produced is both exponentially small and exponentially uncertain. (In discussing the resolution of the monopole problem I am tacitly assuming that the SSB of the GUT is occurring during the SSB transition in question, or that it has already occurred in an earlier SSB transition; if not then one has to worry about the monopoles produced in the subsequent GUT transition.) Although monopole production is intrinsically small in inflationary models, the uncertainties in the number of monopoles produced are exponential and of course, it is also possible that monopoles might be produced as topological defects in a subsequent phase transition [46b] (although it may be difficult to arrange that they not be overproduced).

Finally, the inflationary scenario sheds no light upon *the cosmological constant puzzle.* Although it can potentially successfully resolve all of the other puzzles in my list, inflation is, in some sense, a house of cards built upon the cosmological constant puzzle.

5. *Scalar field dynamics*

The evolution of the scalar field ϕ is key to understanding new inflation. In this section I will focus on the semi-classical dynamics of ϕ. Later, I will return to the question of the validity of the semi-classical approach. Much of what I will discuss here is covered in more detail in Ref. [47].

Stated in the most general terms, the current view of inflation is that it involves the dynamical evolution of a very weakly-coupled scalar field (hereafter referred to as ϕ) which is, for one reason or another, initially displaced from the minimum of its potential (see Fig. 11). While it is displaced from its minimum, and is slowly-evolving toward that minimum, its potential energy density drives the rapid (exponential) expansion of the Universe, now known as inflation.

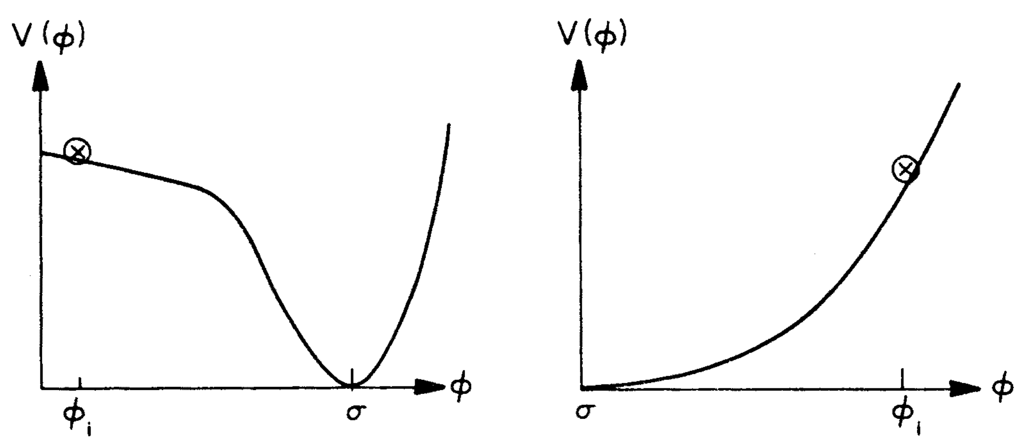


Fig. 11. Stated in the most general terms, inflation involves the dynamical evolution of a scalar field which was initially displaced from the minimum of its potential, be that minimum at $\sigma = 0$ or $\sigma \neq 0$

The usual assumptions which are made (often implicitly) in order to analyze the scalar field dynamics inflation are:

* A FRW spacetime with scale factor $R(t)$ and expansion rate

$$H^2 \equiv (\dot{R}/R)^2 = 8\pi\varrho/3m_{\rm Pl}^2 - k/R^2, \tag{26}$$

where the energy density is assumed to be dominated by the stress energy associated with the scalar field (in any case, other forms of stress energy rapidly redshift away during inflation and become irrelevant).

* The scalar field ϕ is spatially constant (at least on a scale $\gtrsim H^{-1}$) with initial value $\phi_i \neq \sigma$, where $V(\sigma) = V'(\sigma) = 0$.

* The semi-classical equation of motion for ϕ provides an accurate description of its evolution; more precisely,

$$\phi(t) = \phi_{\rm cl}(t) + \Delta\phi_{\rm QM},$$

where the quantum fluctuations (characterized by size $\Delta\phi_{\rm QM} \simeq H/2\pi$) are assumed to be a small perturbation to the classical trajectory $\phi_{\rm cl}(t)$. From this point forward I will drop the subscript 'cl'. I will return later to these assumptions to discuss how they have been or can be relaxed and/or justified.

Consider a classical scalar field (minimally coupled) with lagrangian density given by

$$\mathscr{L} = -\tfrac{1}{2}\,\partial_\mu\phi\partial^\mu\phi - V(\phi). \tag{27}$$

For now I will ignore the interactions that ϕ must necessarily have with other fields in the theory. As it will turn out they must be weak for inflation to work, so that this assumption is a reasonable one. The stress-energy tensor for this field is then

$$T_{\mu\nu} = -\partial_\mu\phi\partial_\nu\phi - \mathscr{L}g_{\mu\nu}. \tag{28}$$

Assuming that in the region of interest ϕ is spatially-constant, $T_{\mu\nu}$ takes on the perfect fluid form with energy density and pressure given by

$$\varrho = \tfrac{1}{2}\,\dot{\phi}^2 + V(\phi)\,(+(\nabla\phi)^2/2R^2), \tag{29a}$$

$$p = \tfrac{1}{2}\,\dot{\phi}^2 - V(\phi)\,(-(\nabla\phi)^2/6R^2), \tag{29b}$$

where I have included the spatial gradient terms for future reference. (Note, once inflation begins the spatial gradient terms decrease rapidly, $(\nabla\phi)^2/R^2 \propto R^{-2}$, for wavelengths $\gtrsim H^{-1}$, and quickly become negligible.) That the spatial gradient term in ϕ be unimportant is crucial to inflation; if it were to dominate the pressure and energy density, then $R(t)$ would grow as t (since $p = -\tfrac{1}{3}\varrho$) and not exponentially.

The equations of motion for ϕ can be obtained either by varying the action or by using $T^{\mu\nu}_{;\nu} = 0$. In either case the resulting equation is:

$$\ddot{\phi} + 3H\dot{\phi}(+\Gamma\dot{\phi}) + V'(\phi) = 0. \tag{30}$$

I have explicitly included the $\Gamma\dot{\phi}$ term which arises due to particle creation. The $3H\dot{\phi}$ friction term arises due to the expansion of the Universe; as the scalar field gains momentum, that momentum is redshifted away by the expansion.

This equation, which is analogous to that for a ball rolling with friction down a hill with a valley at the bottom, has two qualitatively different regimes, each of which has a simple, approximate, analytic solution. Fig. 12 shows schematically the potential $V(\phi)$.

(1) The slow-rolling regime

In this regime the field rools at terminal velocity and the $\ddot{\phi}$ term is negligible. This occurs in the interval where the potential is very flat, the conditions for sufficient flatness being [44]:

$$|V''| \leqslant 9H^2, \tag{31a}$$

$$|V' m_{\rm Pl}/V| \leqslant (48\pi)^{1/2}. \tag{31b}$$

Condition (31a) usually subsumes condition (31b), so that condition (31a) generally suffices. During the slow-rolling regime the equation of motion for ϕ reduces to

$$\dot{\phi} \simeq -V'/3H. \tag{32}$$

During the slow-rolling regime particle creation is exponentially suppressed [43] because the timescale for the evolution of ϕ (which sets the energy/momentum scale of the particles created) is much greater than the Hubble time (which sets the physics horizon), i.e., any particles radiated would have to have wavelengths much larger than the physics horizon,

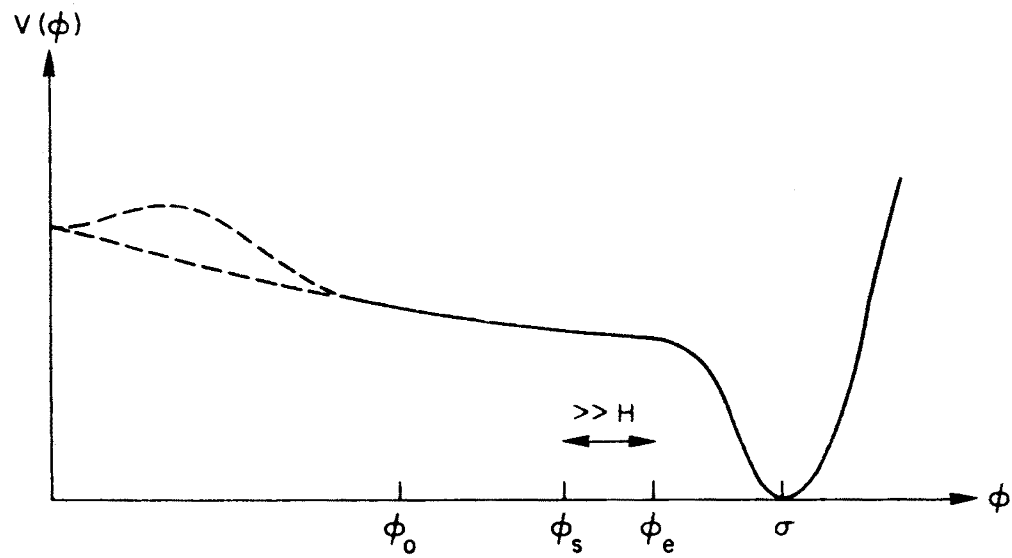


Fig. 12. Schematic plot of the potential required for inflation. The shape of the potential for $\phi \ll \sigma$ determines how the barrier between $\phi = 0$ and $\phi = \sigma$ (if one exists) is penetrated. The value of ϕ after barrier penetration is taken to be ϕ_0; the flat region of the potential is the interval $[\phi_s, \phi_e]$

which results in the exponential suppression of particle creation during this epoch. Thus, the $\Gamma\dot{\phi}$ term can be neglected during the 'slow roll'.

Suppose the interval where conditions (31a, b) are satisfied is $[\phi_s, \phi_e]$, then the number of e-folds of expansion which during the time ϕ is evolving from $\phi = \phi_s$ to $\phi = \phi_e$ ($\equiv N$) is

$$N \equiv \int H dt \simeq -3 \int_{\phi_s}^{\phi_e} H^2 d\varphi / V'(\phi) \simeq -(8\pi/m_{\rm Pl}^2) \int V(\phi) d\varphi / V'(\phi). \tag{33}$$

(Note that $R_e/R_s \equiv \exp(N)$ since $\dot{R}/R \equiv H$.) Taking H^2/V' to be roughly constant over this interval and approximating V' as $\simeq \phi V''$ (which is approximately true for polynomial potentials) it follows that

$$N \approx 3H^2/V'' \geqslant 3.$$

If there is a region of the potential where the evolution is friction-dominated, then N will necessarily be greater than 1 (by condition (31a)).

(2) Coherent field oscillations

In this regime

$$|V''| \gg 9H^2,$$

and ϕ evolves rapidly, on a timescale $\ll$ the Hubble time H^{-1}. Once ϕ reaches the bottom of its potential, it will oscillate with an angular frequency equal to $m_\phi \equiv V''(\sigma)^{1/2}$. In this regime it proves useful to rewrite Eq. (30) for the evolution of ϕ as

$$\dot{\varrho}_\phi = -3H\dot{\phi}^2 - \Gamma\dot{\phi}^2, \tag{34}$$

where

$$\varrho_\phi \equiv 1/2\, \dot{\phi}^2 + V(\phi).$$

Once ϕ is oscillating about $\phi = \sigma$, $\dot{\phi}^2$ can be replaced by its average over a cycle

$$\langle \dot{\phi}^2 \rangle_{\rm cycle} = \varrho_\phi,$$

and Eq. (34) becomes

$$\dot{\varrho}_\phi = -3H\varrho_\phi - \Gamma\varrho_\phi \tag{35}$$

which is nothing else but the equation for the evolution of the energy density of zero momentum, massive particles with a decay width Γ.

Referring back to Eq. (29) we can see that the cycle average of the pressure (i.e., space-space components of $T_{\mu\nu}$) is zero — as one would expect for NR particles. The coherent ϕ oscillations are in every way equivalent to a very cold condensate of ϕ particles. The decay of these oscillations due to quantum particle creation is equivalent to the decay of zero-momentum ϕ particles.

The complete set of semi-classical equations for the reheating of the Universe is

$$\dot{\varrho}_\phi = -3H\varrho_\phi - \Gamma\varrho_\phi, \tag{36a}$$

$$\dot{\varrho}_r = -4H\varrho_r + \Gamma\varrho_\phi, \tag{36b}$$

$$H^2 = 8\pi G(\varrho_r + \varrho_\phi)/3, \tag{36c}$$

where $\varrho_r = (\pi^2/30)g_* T^4$ is the energy density in the relativistic particles produced by the decay of the coherent field oscillations. (I have tacitly assumed that the decay products of ϕ rapidly thermalize; Eq. (36b) is correct whether or not the decay products thermalize, so long as they are relativistic.) The evolution for the energy density in the scalar is easy to obtain

$$\varrho_\phi = M^4(R/R_e)^{-3} \exp\left[-\Gamma(t-t_e)\right], \tag{37}$$

where I have set the initial energy equal to M^4, the initial epoch being when the scalar field begins to evolve rapidly (at $R = R_e$, $\phi = \phi_e$, and $t = t_e$).

From $t = t_e$ until $t \simeq \Gamma^{-1}$, the energy density of the Universe is dominated by the coherent sloshings of the scalar field ϕ, set into motion by the initial vacuum energy associated with $\phi \ll \sigma$. During this phase

$$R(t) \propto t^{2/3},$$

that is, the Universe behaves as if it were dominated by NR particles — which it is!

Interestingly enough it follows from Eq. (36) that during this time the energy density in radiation is actually decreasing ($\varrho_r \propto R^{-3/2}$ — see Fig. 13). (During the first Hubble time after the end of inflation ϱ_r does increase.) However, the all important entropy per comoving volume is increasing $S \propto R^{15/8}$. When $t \simeq \Gamma^{-1}$, the coherent oscillations begin to decay exponentially, and the entropy per comoving volume levels off — indicating the end of the reheating epoch. The temperature of the Universe at this time is,

$$T_{RH} \simeq g_*^{-1/4}(\Gamma m_{Pl})^{1/2}. \tag{38}$$

If Γ^{-1} is less than H^{-1}, so that the Universe reheats in less than an expansion time, then all of the vacuum is converted into radiation and the Universe is reheated to a temperature

$$T_{RH} \simeq g_*^{-1/4} M \quad (\text{if } \Gamma \gg H) \tag{38'}$$

the so-called case of good reheating.

To summarize the evolution of the scalar field ϕ: early on ϕ evolves very slowly, on a timescale $\gg$ the Hubble time H^{-1}; then as the potential steepens (and $|V''|$ becomes $> 9H^2$) ϕ begins to evolve rapidly, on a timescale $\ll$ the Hubble time H^{-1}. As ϕ oscillates about the minimum of its potential the energy density in these oscillations dominates the energy density of the Universe and behaves like NR matter ($\varrho_\phi \propto R^{-3}$); eventually when $t \simeq \Gamma^{-1}$, these oscillations decay exponentially, 'reheating' the Universe to a temperature of $T_{RH} \simeq g_*^{-1/4}(\Gamma m_{Pl})^{1/2}$ (if $\Gamma > H$, so that the Universe does not e-fold in the time it takes the oscillations to decay, then $T_{RH} \simeq g_*^{-1/4} M$). Saying that the Universe reheats when $t \simeq \Gamma^{-1}$ is a bit paradoxical as the temperature has actually been *decreasing* since shortly after the ϕ oscillations began. However, the fact that the temperature of the Universe was actually once greater than $T_{\rm RH}$ for $t < \Gamma^{-1}$ is probably of no practical use since the entropy per comoving volume increases until $t \simeq \Gamma^{-1}$ — by a factor of $(M^2/\Gamma m_{\rm Pl})^{5/4}$, and any interesting objects that might be produced (e.g., net baryon number, monopoles, etc.) will be diluted away by the subsequent entropy production. By any reasonable measure, $T_{\rm RH}$ is the reheat temperature of the Universe. The evolution of ϱ_ϕ, ϱ_r, and S are summarized in Fig. 13.

Armed with our detailed knowledge of the evolution of ϕ we are ready to calculate the precise number of e-folds of inflation necessary to solve the horizon and flatness problems and to discuss direct baryon number production. First consider the requisite number of e-folds required for sufficient inflation. To solve the homogeneity problem we need to insure that a smooth patch containing an entropy of at least 10^{88} results from inflation. Suppose the initial bubble or fluctuation region has a size $H^{-1} \simeq m_{\rm Pl}/M^2$ — certainly it is not likely to be significantly larger than this. During inflation it grows by a factor of $\exp(N)$. Next, while the Universe is dominated by coherent field oscillations it grows by a factor of

$$(R_{\rm RH}/R_e) \simeq (M^4/T_{\rm RH}^4)^{1/3},$$

where $T_{\rm RH}$ is the reheat temperature. Cubing the size of the patch at reheating (to obtain ts volume) and multiplying its volume by the entropy density ($s \approx T_{\rm RH}^3$), we obtain

$$S_{\rm patch} \simeq e^{3N} m_{\rm Pl}^3/(M^2 T_{\rm RH}).$$

Insisting that $S_{\rm patch}$ be greater than 10^{88}, it follows that

$$N \geqslant 53 + \tfrac{2}{3}\ln(M/10^{14}\ {\rm GeV}) + \tfrac{1}{3}\ln(T_{\rm RH}/10^{10}\ {\rm GeV}). \qquad (39)$$

Varying M from 10^{19} GeV to 10^8 GeV and $T_{\rm RH}$ from 1 GeV to 10^{19} GeV this lower bound on N only varies from 36 to 68.

The flatness problem involves the smallness of the ratio

$$x = (k/R^2)/(8\pi G\varrho/3)$$

equired at early times. Taking the pre-inflationary value of x to be x_i and remembering that

$$x(t) \propto \begin{cases} R^{-2} & \varrho = \text{const} \\ R & \varrho \propto R^{-3} \\ R^2 & \varrho \propto R^{-4} \end{cases}$$

it follows that the value of x today is

$$x_{\rm today} = x_i e^{-2N}(M/T_{\rm RH})^{4/3}(T_{\rm RH}/10\ {\rm eV})^2(10\ {\rm eV}/3\ {\rm K}).$$

Insisting that $x_{\rm today}$ be at most of order unity implies that

$$N \geqslant 53 + \ln(x_i) + \tfrac{2}{3}\ln(M/10^{14}\ {\rm GeV}) + \tfrac{1}{3}\ln(T_{\rm RH}/10^{10}\ {\rm GeV})$$

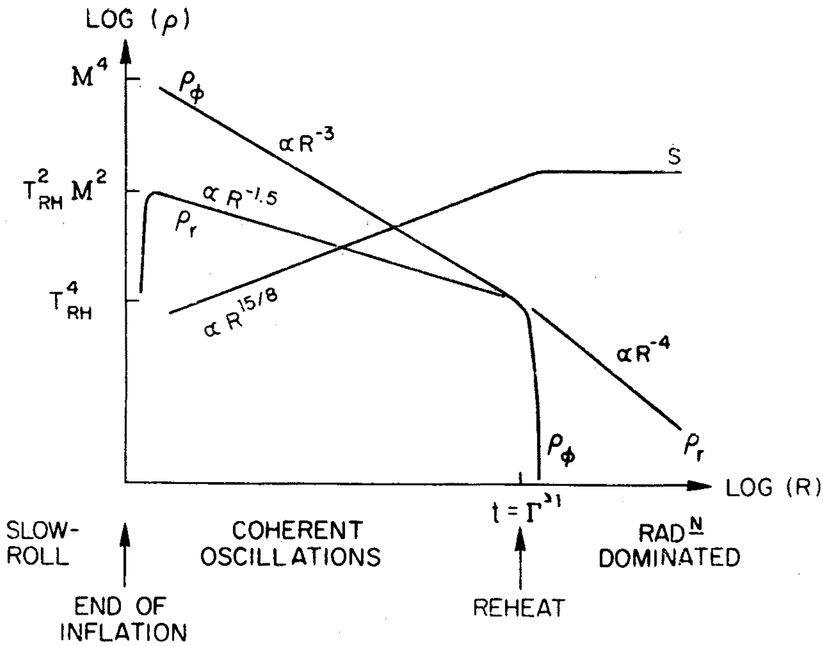


Fig. 13. The evolution of ϱ_ϕ, ϱ_r, and S during the epoch when the Universe is dominated by coherent ϕ-oscillations. The reheat temperature $T_{RH} \simeq g^{-1/4}(\Gamma m_{Pl})^{1/2}$. The maximum temperature achieved after inflation is actually greater, $T_{max} \simeq (T_{RH}M)^{1/2}$

— up to the term $\ln(x_i)$, precisely the same bound as we obtained to solve the homogeneity problem. Solving the isotropy problem depends upon the initial anisotropy present; during inflation isotropy decreases exponentially (see Ref. [48]).

Finally, let's calculate the baryon asymmetry that can be directly produced by the decay of the ϕ particles themselves. Suppose that the decay of each ϕ particle results in the production of net baryon number ε. This net baryon number might be produced directly by the decay of a ϕ particle (into quarks and leptons) or indirectly through an intermediate state ($\phi \to X\overline{X}$; X, $\overline{X} \to$ quarks and leptons; e.g., X might be a superheavy, color triplet Higgs [49]). The baryon asymmetry produced per volume is then

$$n_B \simeq \varepsilon n_\phi.$$

On the other hand we have

$$(g_* \pi^2/30)T^4_{RH} \simeq n_\phi m_\phi.$$

Taken together it follows that [42, 50]

$$n_B/s \simeq (3/4)\varepsilon T_{RH}/m_\phi. \tag{40}$$

This then is the baryon number per entropy produced by the decay of the ϕ particles directly. If the reheat temperature is not very high, baryon number non-conserving interactions will not subsequently reduce the asymmetry significantly. Note that the baryon asymmetry produced only depends upon the *ratio* of the reheat temperature to the ϕ particle mass. This is important, as it means that a very low reheat temperature can be tolerated, so long as the ratio of it to the ϕ particle mass is not too small.

6. *Origin of density inhomogeneities*

To this point I have assumed that ϕ is precisely uniform within a given bubble or fluctuation region. As a result, each bubble or fluctuation region resembles a perfectly isotropic and homogeneous Universe after reheating. However, because of deSitter space produced quantum fluctuations, ϕ cannot be exactly uniform, even within a small region of space. It is a well-known result that a massless and noninteracting scalar field in deSitter space has a spectrum of fluctuations given by (see, e.g., Ref. [51])

$$(\Delta\phi)^2 \equiv (2\pi)^{-3}k^3|\delta\phi_k|^2 = H^2/16\pi^3, \tag{41}$$

where

$$\delta\phi = (2\pi)^{-3}\int d^3k\delta\phi_k^{-ikx}, \tag{42}$$

and $\vec{x}$ and $\vec{k}$ are comoving quantities. This result is applicable to inflationary scenarios as the scalar field responsible for inflation must be very weakly-coupled and nearly massless. (That Universe is not precisely deSitter during inflation, i.e., $\varrho+p = \dot{\phi}^2 \neq 0$, does not affect this result significantly; this point is addressed in Ref. [52]). These deSitter space produced quantum fluctuations result in a calculable spectrum of adiabatic density perturbations. These density perturbations were first calculated by the authors of Refs [53–56]; they have also been calculated by the authors of Ref. [57] who addressed some of the technical issues in more detail. All the calculations done to date arrive at the same result. I will briefly describe the calculation in Ref. [56]; my emphasis here will be to motivate the result. I refer the reader interested in more details to the aforementioned references.

It is conventional to expand density inhomogeneities in a Fourier expansion

$$\delta\varrho/\varrho = (2\pi)^{-3}\int \delta_k e^{-ikx}d^3k. \tag{43}$$

The physical wavelength and wavenumber are related to comoving wavelength and wave number, λ and k, by

$$\lambda_{\rm ph} = (2\pi/k)R(t) \equiv \lambda R(t),$$

$$k_{\rm ph} = k/R(t).$$

The quantity most people refer to as $\delta\varrho/\varrho$ on a given scale is more precisely the RMS mass fluctuation on that scale

$$(\delta\varrho/\varrho)_k \equiv \langle(\delta M/M)^2\rangle_k \simeq \Delta_k^2 \equiv (2\pi)^{-3}k^3|\delta_k|^2, \tag{44}$$

which is just related to the Fourier component δ_k on that scale.

The cosmic scale factor is often normalized so that $R_{\text{today}} = 1$; this means that given Fourier components are characterized by the physical size that they have today (neglecting the fact that once a given scale goes non-linear ($\delta\varrho/\varrho \gtrsim 1$) objects of that size form bound 'lumps' that no longer participate in the universal expansion and remain roughly constant in size). The mass (in NR matter) contained within a sphere of radius $\lambda/2$ is

$$M(\lambda) \simeq 1.5\times 10^{11} M_{\odot}(\lambda/\text{Mpc})^3 \Omega h^2.$$

Although physics depends on physical quantities (k_{ph}, λ_{ph}, etc.), the comoving labels k, M, and λ are the most useful way to label a given component as they have the affect of the expansions already scaled out.

I want to emphasize at the onset that the quantity $\delta\varrho/\varrho$ is not gauge invariant (under general coordinate transformations). This fact makes life very difficult when discussing Fourier components with wavelengths larger than the physics horizon (i.e., $\lambda_{\text{ph}} \gg H^{-1}$). The gauge non-invariance of $\delta\varrho/\varrho$ is not a problem when $\lambda_{\text{ph}} \ll H^{-1}$, as the analysis is essentially Newtonian. The usual approach is to pick a convenient gauge (e.g., the synchronous gauge where $g_{00} = -1$, $g_{0i} = 0$) and work very carefully (see Refs [58, 59]). The more

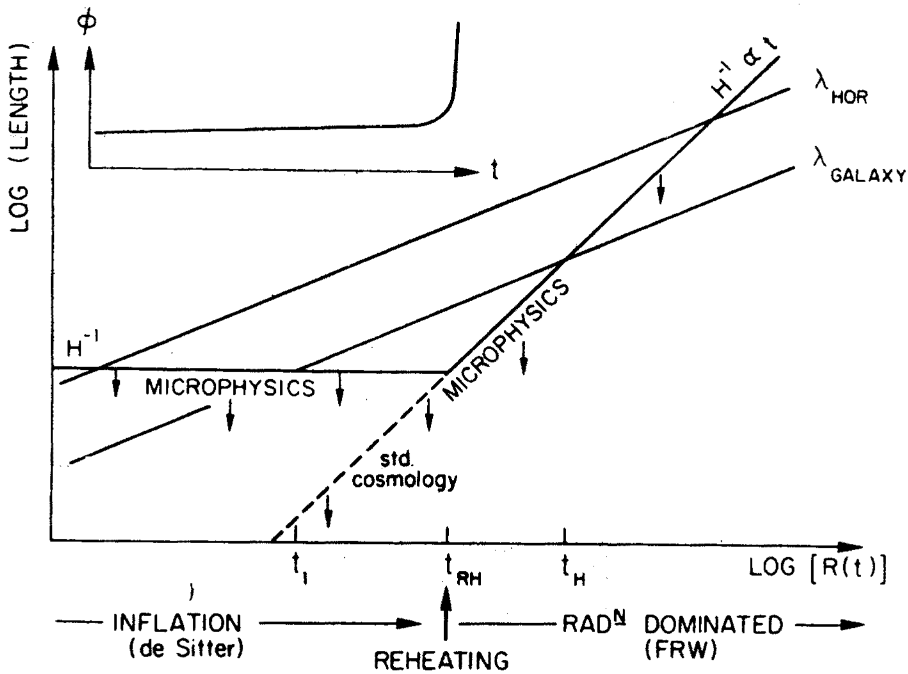


Fig. 14. The evolution of the physical size of galactic- and (present) horizon-sized perturbations ($\lambda_{\text{ph}} \propto R$) and the size of the physics horizon H^{-1}. Causally-coherent microphysics operates only on scales $< H^{-1}$. In the standard cosmology a perturbation crosses the horizon but once as $H^{-1} \propto R^n (n > 1)$, making it impossible for microphysics to create density perturbations at early times. In the inflationary cosmology a perturbation crosses the horizon twice (since $H^{-1} \simeq$ const during inflation), and so microphysics can produce density perturbations at early times

elegant approach is to focus on gauge-invariant quantities; see Ref. [60]. I will gloss over the subtleties of gauge invariance in my discussion, which is aimed at motivating the correct answer.

The evolution of a given Fourier component (in the linear regime — $\delta\varrho/\varrho \ll 1$) separates into two qualitatively different regimes, depending upon whether or not the perturbation is inside or outside the physics horizon ($\simeq H^{-1}$). When $\lambda_{\rm ph} \leqslant H^{-1}$, microphysical processes can affect its evolution — such processes include: quantum mechanical effects, pressure support, free-streaming of particles, 'Newtonian gravity', etc. In this regime the evolution of the perturbation is very dynamical. When a perturbation is outside the physics horizon, $\lambda_{\rm ph} \geqslant H^{-1}$, microphysical processes do not affect its evolution; in a very real sense its evolution is kinematic — it evolves as a wrinkle in the fabric of space-time.

In the standard cosmology, a given Fourier component crosses the horizon only once, starting outside the horizon and crossing inside at a time (see Fig. 14)

$$t \simeq (M/M_\odot)^{2/3}\ \mathrm{sec}$$

(valid during the radiation-dominated epoch). For this reason it is not possible to create adiabatic (more precisely, curvature) perturbations by causal microphysical processes which operate at early times [59, 60]. In the standard cosmology, if adiabatic perturbations are present, they must be present *ab initio*. The smallness of the particle horizon at early times relative to the comoving volume occupied by the observable Universe today strikes again!

(It is possible for microphysical processes to create isothermal, more precisely isocurvature, perturbations. Once such perturbations cross inside the horizon they are characterized by a spectrum

$$(\delta\varrho/\varrho) \propto (M/M_{\rm H})^{-1/2}$$

or steeper. Here $M_{\rm H}$ is the horizon mass when the perturbations were created. Thus the earlier the processes operate, the smaller the perturbations are on interesting scales. By an appropriate choice of gauge it is possible to view these isothermal perturbations as adiabatic perturbations with a very steep spectrum, $\delta\varrho/\varrho \propto M^{-7/6}$; however, as must be the case, they cross the horizon with the amplitude mentioned above. For more details, see Refs. [59, 60].)

Because the distance to the physics horizon remains approximately constant during inflation, the situation is very different in the inflationary Universe. All interesting scales start inside the horizon, cross outside the horizon and once again come inside the horizon (at the usual epoch); see Fig. 14. This means that causal microphysical processes can set up curvature perturbations on astrophysically-interesting scales. (This point seems to have been first appreciated by Press [61].)

Consider the evolution of a given Fourier component k. Early during the inflationary epoch $\lambda_{\rm ph} \leqslant H^{-1}$, and quantum fluctuations in ϕ give rise to density perturbations on this scale. As the scale passes outside the horizon, say at $t = t_1$, microphysical processes become impotent, and $\delta\varrho/\varrho$ freezes out at a value,

$$(\delta\varrho/\varrho)_k \simeq O(\dot{\phi}H\Delta\phi/M^4),$$
$$\simeq O(\dot{\phi}H^2/M^4), \tag{45}$$

as the scale leaves the horizon. From that point forward, the QM fluctuation is assumed to 'freeze in' and thereafter evolve classically. Note in the approximation that H and $\dot{\phi}$ are constant during the inflationary epoch the value of $\delta\varrho/\varrho$ as the perturbation leaves the horizon is independent of k. This scale independence of $\delta\varrho/\varrho$ when perturbations cross outside the horizon is of course traceable to the time translation invariance of deSitter space.

While outside the horizon the evolution of a perturbation is kinematical, independent of scale, and gauge dependent. There is a gauge independent quantity ($\equiv \zeta$) which remains constant while the perturbation is outside the horizon, and which at horizon crossing ($t = t_1$ and t_H) is given by

$$\zeta|_{\text{horizon crossing}} \simeq \delta\varrho/(\varrho+p),$$
$$\zeta(t_1) = \zeta(t_H),$$
$$\Rightarrow [\delta\varrho/(\varrho+p)]_{t=t_H} \simeq [\delta\varrho/(\varrho+p)]_{t=t_1} \tag{46}$$

(see Refs [56] and [62] for more details). When the perturbation crosses back inside the horizon: $(\varrho+p) = n\varrho(n = 4/3$, radiation-dominated; $n = 1$, matter-dominated) so that up to a numerical factor $(\delta\varrho/\varrho)|_{t_H} \simeq [\delta\varrho/(\varrho+p)]|_{t_H}$. During inflation, however, $\varrho+p = \dot{\phi}^2 \ll \varrho \simeq M^4$ so that $(\delta\varrho/\varrho)|_{t_1} \simeq (\dot{\phi}^2/M^4)\,[\delta\varrho/\varrho(+p)]|_{t_1}$. Note, $M^4/\dot{\phi}^2$ is typically a very large number. Eqs (45, 46) then imply

$$(\delta\varrho/\varrho)_H \equiv (\delta\varrho/\varrho)_{t=t_H} \simeq (M^4/\dot{\phi}^2)\,(\delta\varrho/\varrho)_{t_1} \simeq H^2/\dot{\phi}. \tag{47}$$

Note that in the approximation that $\dot{\phi}$ and H are constant during inflation the amplitude of $\delta\varrho/\varrho$ at horizon crossing ($= (\delta\varrho/\varrho)_H$) is independent of scale. This fact is traceable to the time-translation invariance of the nearly-deSitter inflationary epoch and the scale--independent evolution of $(\delta\varrho/\varrho)$ while the perturbation is outside the horizon. The so--called scale-invariant or Zel'dovich spectrum of density perturbations was first discussed, albeit in another context, by Harrison [63] and Zel'dovich [64]. Scale-invariant adiabatic density perturbations are a generic prediction of inflation. (Because H and $\dot{\phi}$ are not precisely constant during inflation, the spectrum is not quite scale-invariant. However the scales of astrophysical interest, say $\lambda \simeq 0.1$ Mpc — 100 Mpc, cross outside the horizon during a very short interval, $\Delta N \simeq 6.9$, during which H, $\dot{\phi}$, and ϕ are very nearly constant. For most models of inflation the deviations are not expected to be significant; for further discussion see Refs [65, 66].) Although the details of structure formation are not presently sufficiently well understood to say what the initial spectrum of perturbations must have been, the Zel'dovich with an amplitude of about 10^{-4}–10^{-5} is certainly a viable possibility.

Before moving on, let me be very precise about the amplitude of the inflation-produced adiabatic density perturbations. Perturbations which re-enter the horizon while the Universe is still radiation-dominated ($\lambda \leqslant \lambda_{eq} \simeq 13h^{-2}$ Mpc), do so as a sound wave in the photons and baryons with amplitude

$$(\delta\varrho/\varrho)_H \equiv k^{3/2}|\delta_k|/(2\pi)^{3/2} \simeq H^2/(\pi^{3/2}\dot{\phi}). \tag{48a}$$

Perturbations in non-interacting, relic particles (such as massive neutrinos, axions, etc.), which by the equivalence principle must have the same amplitude at horizon crossing, do not oscillate, but instead grow slowly ($\propto \ln R$). By the epoch of matter-radiation equivalence they have an amplitude of 2–3 times that of the initial baryon-photon sound wave, or

$$(\delta\varrho/\varrho)_{MD} \simeq (2-3)\,(\delta\varrho/\varrho)_H \simeq (2-3)H^2/(\pi^{3/2}\dot{\phi}). \tag{48b}$$

It is this amplitude which must be of order 10^{-5}–10^{-4} for successful galaxy formation.

Perturbations which re-enter the horizon when the Universe is already matter-dominated (scales $\lambda \geqslant \lambda_{eq} \simeq 13h^{-2}$ Mpc) do so with amplitude

$$(\delta\varrho/\varrho)_H \simeq k^{3/2}|\delta_k|/(2\pi)^{3/2} \simeq (H^2/10)/(\pi^{3/2}\dot{\phi}). \tag{49}$$

Once inside the horizon they continue to grow (as $t^{3/2}$ since the Universe is matter-dominated).

When the structure formation problem is viewed as an initial data problem, it is the spectrum of density perturbations at the epoch of matter domination which is the relevant input spectrum. The shape of this spectrum has been carefully computed by the authors of Ref. [67]. Roughly speaking, on scales less than λ_{eq} the spectrum is almost flat, varying as $\lambda^{-3/4} \propto M^{-1/4}$ for scales around the galaxy scale ($\simeq 1$ Mpc). On scales much greater than λ_{eq}, $(\delta\varrho/\varrho) \propto \lambda^{-2} \propto M^{-2/3}$ (in the synchronous gauge where adiabatic perturbations grow as t^n; $n = 2/3$ matter dominated, $n = 1$ radiation dominated. Since these scales have yet to re-enter the horizon they have not yet achieved their horizon-crossing amplitude).

In order to compute the amplitude of the inflation-produced adiabatic density perturbations we need to evaluate $H^2/\dot{\phi}$ when the astrophysically-relevant scales crossed outside the horizon. Recall, in the previous Section we computed when the comoving scale corresponding to the present Hubble radius crossed outside the horizon during inflation — up to 'ln terms' $N \simeq 53$ or so e-folds before the end of inflation, cf., Eq. (39). The present Hubble radius corresponds to a scale of about 3000 Mpc; therefore the scale λ must have crossed the horizon $\ln(3000\ \mathrm{Mpc}/\lambda)$ e-folds later:

$$N_\lambda \simeq N_{HOR} - 8 + \ln(\lambda/\mathrm{Mpc}) \simeq 45 + \ln(\lambda/\mathrm{Mpc})$$
$$+\tfrac{2}{3}\ln(M/10^{14}\ \mathrm{GeV}) + \tfrac{1}{3}\ln(T_{RH}/10^{10}\ \mathrm{GeV}).$$

Typically $H^2/\dot{\phi}$ depends upon N_λ to some power [65]; since N_λ only varies logarithmically ($\Delta N/N \simeq 0.14$ in going from 0.1 Mpc to 3000 Mpc), the scale dependence of the spectrum is almost always very minimal.

As mentioned earlier, a generic prediction of the inflationary Universe is that today Ω should be equal to one to a high degree of precision. Equivalently, that means

$$|(k/R^2)/(8\pi G\varrho/3)| \ll 1$$

since

$$\Omega = 1/[1-(k/R^2)/(8\pi G\varrho/3)].$$

Therefore one might conclude that an accurate measurement of Ω would have to yield 1 to extremely high precision. However, because of the adiabatic density perturbations produced during inflation that is not the case. Adiabatic density fluctuations correspond to fluctuations in the local curvature

$$\delta\varrho/\varrho \simeq \delta(k/R^2)/(8\pi G\varrho/3).$$

This means that should we be able to very accurately probe the value of Ω (equivalently the curvature of space) on the scale of our Hubble volume, say by using the Hubble diagram, we would necessarily obtain a value for Ω which is dominated by the curvature fluctuations on the scale of the present horizon,

$$\Omega_{\text{obs}} \simeq 1+\delta(k/R^2)/(8\pi G\varrho/3) \simeq 1 \pm O(10^{-4}-10^{-5}),$$

and so we would obtain a value different from 1 by about a part in 10^4 or 10^5.

Finally, let me briefly mention that isothermal density perturbations can also arise during inflation. (Isothermal density perturbations are characterized by $\delta\varrho = 0$, but $\delta(n_i/n_\gamma) \neq 0$ in some components. They correspond to spatial fluctuations in the local pressure due to spatial fluctuations in the local equation of state.) Such perturbations can arise from the deSitter produced fluctuations in other quantum fields in the theory.

The simplest example occurs in the axion-dominated Universe [68, 69, 70]. Suppose that Peccei-Quinn symmetry breaking occurs before or during inflation. Until instanton effects become important ($T \simeq$ few 100 MeV) the axion field $a = f_a\theta$ is massless and θ is in general not aligned with the minimum of its potential: $\theta = \theta_1 \neq 0$ (I have taken the minimum of the axion potential to be $\theta = 0$). Once the axion develops a mass (equivalently, its potential develops a minimum) θ begins to oscillate; these coherent oscillations correspond to a condensate of very cold axions, with number density $\propto \theta^2$. (For further discussion of the coherent axion oscillations see Refs [71–73].) During inflation deSitter space produced quantum fluctuations in the axion field gave rise to spatial fluctuations in θ_1:

$$\delta\theta \simeq \delta a/f_a \simeq H/f_a.$$

Once the axion field begins to oscillate, these spatial fluctuations in the axion field correspond to fluctuations in the local axion to photon ratio

$$\delta(n_a/n_\gamma)/(n_a/n_\gamma) \simeq 2\delta\theta/\theta_1 \simeq 2H/(f_a\theta_1).$$

More precisely

$$(\delta n_a/n_a) = k^{3/2}|\delta a(k)|/(2\pi)^{3/2} = H/(2\pi^{3/2} f_\lambda \theta_1), \tag{50}$$

where f_λ is the expectation value of f_a when the scale λ leaves the horizon (in some models the expectation value of the field which breaks PQ symmetry evolves as the Universe is inflating so that f_λ can be $< f_a$). It is possible that these isothermal axion fluctuations can be important for galaxy formation in an axion-dominated, inflationary Universe [69].

7. *Specific models — part I. Interesting failures*

(1) 'Old inflation'

By old inflation I mean Guth's original model of inflation. In his original model the Universe inflated while trapped in the $\phi = 0$ false vacuum state. In order to inflate enough the vacuum had to be very metastable; however, that being the case, the bubble nucleation probability was necessarily small — so small that the bubbles that did nucleate never percolated, resulting in a Universe which resembled swiss cheese more than anything else [36]. The interior of an individual bubble was not suitable for our present Universe either. Because he was not considering flat potentials, essentially all of the original false vacuum energy resided in bubble walls rather than in vacuum energy inside the bubbles themselves. Although individual bubbles would grow to a very large size given enough time, their interiors would contain very little entropy (compared to the 10^{88} in our observed Universe). In sum, the Universe inflated all right, but did not 'gracefully exit' from inflation back to a radiation-dominated Universe — close, Alan, but no cigar!

(2) Coleman–Weinberg SU(5)

The first model of new inflation studied was the Coleman–Weinberg SU(5) GUT [37, 38]. In this model the field which inflates is the 24-dimensional Higgs which also breaks SU(5) down to SU(3) × SU(2) × U(1). Let ϕ denote its magnitude in the SU(3) × SU(2) × U(1) direction. The one-loop, zero-temperature Coleman–Weinberg [74] potential is

$$\begin{aligned} V(\phi) &= 1/2\, B\sigma^4 + B\phi^4\{\ln(\phi^2/\sigma^2) - 1/2\}, \\ B &= 25\alpha_{\mathrm{GUT}}^2/16 \simeq 10^{-3}, \\ \sigma &\simeq 2\times 10^{15}\ \mathrm{GeV}. \end{aligned} \tag{51}$$

Due to the absence of a mass term ($m^2\phi^2$), the potential is very flat near the origin (SSB arises due to one-loop radiative corrections [74]); for $\phi \ll \sigma$:

$$\begin{aligned} V(\phi) &\simeq B\sigma^4/2 - \lambda\phi^4/4, \\ \lambda &\simeq |4B \ln(\phi^2/\sigma^2)| \simeq 0.1. \end{aligned} \tag{52}$$

The finite temperature potential has a small temperature dependent barrier (height $O(T^4)$) near the origin ($\phi \simeq O(T)$). The critical temperature for this transition is $O(10^{14}$–10^{15} GeV), however the $\phi = 0$ vacuum remains metastable. When the temperature of the Universe drops to $O(10^9$ GeV) or so, the barrier becomes low enough that the finite temperature action for bubble nucleation drops to order unity and the $\phi = 0$ false vacuum becomes unstable [38]. In analogy with solid state phenomenon it is expected that at this the temperature of the Universe will undergo 'spinodal decomposition', i.e., will break up into irregularly shaped regions within which ϕ is approximately constant (so-called fluctuation regions). Approximating the potential by Eq. (52) it is easy to solve for the evolution of ϕ in the slow-rolling regime $[|V''| \leqslant 9\,H^2$ for $\phi^2 \leqslant \phi_e^2 \simeq \sigma^2(\pi\sigma^2/m_{Pl}^2|\ln(\phi^2/\sigma^2)|)]$

$$(H/\phi)^2 \simeq \frac{2\lambda}{3} N(\phi), \tag{53}$$

$$H^2 \simeq \frac{4\pi}{3}\frac{B\sigma^4}{m_{Pl}^2}, \tag{54}$$

where $N(\phi) \equiv \int_{\phi}^{\phi_e} H dt$ is the number of e-folds of inflation the Universe undergoes while ϕ evolves from ϕ to ϕ_e. Clearly, the number of e-folds of inflation depends upon the initial value of $\phi(\equiv \phi_0)$; in order to get sufficient inflation ϕ_0 must be $\simeq O(H)$. Although one might expect ϕ_0 to be of this order in the fluctuation regions since $H \simeq 5\times 10^9$ GeV $\simeq$ (temperature at which the $\phi = 0$ false vacuum loses its metastability), there is a fundamental difficulty. In using the semi-classical equations of motion to describe the evolution of ϕ one is implicitly assuming

$$\phi \simeq \phi_{cl} + \Delta\phi_{QM}, \qquad \Delta\phi_{QM} \ll \phi_{cl}.$$

The deSitter space produced quantum fluctuations in ϕ are of order H. More specifically, it has been shown that [75, 76]

$$\Delta\phi_{QM} \simeq (H/2\pi)(Ht)^{1/2}.$$

Therein lies the difficulty — in order to achieve enough inflation the initial value of ϕ must be of the order of the quantum fluctuations in ϕ. At the very least this calls into question the semiclassical approximation.

The situation gets worse when we look at the amplitude of the adiabatic density perturbations:

$$(\delta\varrho/\varrho)_H \simeq (H^2/\pi^{3/2}\dot{\phi}) \tag{55}$$

$$\simeq (3/\pi^{3/2})(H^3/\lambda\phi^3), \tag{56}$$

$$(\delta\varrho/\varrho)_H \simeq (2/\pi)^{3/2}(\lambda/3)^{1/2}N^{3/2}. \tag{57}$$

For galactic-scale perturbations $N \simeq 50$, implying that $(\delta\varrho/\varrho)_H \simeq 30$! Again, it is clear that the basic problem is traceable to the fact that during inflation $\phi \leqslant H$.

The decay width of the ϕ particle is of order $\alpha_{GUT}\sigma \simeq 10^{13}$ GeV which is much greater than H (implying good reheating), and so the Universe reheats to a temperature of order 10^{14} GeV or so.

From Eq. (53, 57) it is clear that by reducing λ both problems could be remedied — however $\lambda \lesssim 10^{-13}$ is necessary [56]. Of course, as long as the inflating field is a gauge non--singlet λ is set by the gauge coupling strength. From this interesting failure it is clear that one should focus on weakly-coupled, gauge singlet fields for inflation.

(3) Geometric hierarchy model

The first model proposed to address the difficulty mentioned above, was a supersymmetric GUT [77, 78]. In this model ϕ is a scalar field whose potential at tree level is absolutely flat, but due to radiative corrections develops curvature. In the model ϕ is also responsible for the SSB of the GUT. The potential for ϕ is of the form

$$V(\phi) \simeq \mu^4[c_1 - c_2 \ln(\phi/m_{Pl})], \tag{58}$$

where $\mu \simeq 10^{12}$ GeV is the scale of supersymmetry breaking, and c_1 and c_2 are constants which depend upon details of the theory. This form for the potential is only valid away from its SSB minimum ($\sigma \simeq m_{Pl}$) and for $\phi \gg \mu$. The authors presume that higher order effects will force the potential to develop a minimum for $\phi \simeq m_{Pl}$. Since $V' \propto \phi^{-1}$ the potential gets flatter for large ϕ — which already sounds good.

The inflationary scenario for this potential proceeds as follows. The shape of the potential is not determined near $\phi = 0$; depending on the shape ϕ gets to some initial value, say $\phi = \phi_0$, either by bubble nucleation or spinodal decomposition. Then it begins to roll. During slow roll which begins when $|V''| \simeq 9H^2$ and $\phi_s \simeq (c_2/24\pi c_1)^{1/2} m_{Pl}$,

$$H^2 \simeq \frac{8\pi}{3m_{Pl}^2} c_1 \mu^4, \tag{59a}$$

$$(1-\phi^2/m_{Pl}^2) \simeq (c_2/4\pi c_1) N(\phi), \tag{59b}$$

$$(\delta\varrho/\varrho)_H \simeq (H^2/\pi^{3/2}\dot{\phi}), \tag{60a}$$

$$\simeq 8(8/3)^{1/2}(c_1^{3/2}/c_2)\mu^2\phi/m_{Pl}^3. \tag{60b}$$

Note that during the slow roll ($\phi \geqslant \phi_s$)

$$\frac{\phi}{H} \geqslant \frac{\phi_s}{H} \simeq \frac{c_2^{1/2}}{c_1} \frac{1}{8\pi} \frac{m_{Pl}^2}{\mu^2},$$

$$\simeq 10^{13} c_2^{1/2}/c_1 \gg 1,$$

thereby avoiding the difficulty encountered in the Coleman–Weinberg where $\phi \leqslant H$ was required to inflate. For $c_1 \simeq O(1)$, $c_2 \simeq 10^{-8}$ — acceptable values in the model, $(\delta\varrho/\varrho)_H \simeq 10^{-5}$ and $N(\phi_s) \simeq 4\pi c_1/c_2 \simeq 10^9$. The number of e-folds of inflation is very large — 10^9. This is quite typical of the very flat potentials required to achieve $(\delta\varrho/\varrho) \simeq 10^{-4}$–$10^{-5}$.

Now for the bad news. In this model ϕ is very weakly coupled — it only couples to ordinary particles through gravitational strength interactions. Its decay width is

$$\Gamma \simeq O(\mu^6/m_{\mathrm{Pl}}^5), \tag{61}$$

which is much less than H (implying poor reheating) and leads to a reheat temperature of

$$T_{\mathrm{RH}} \simeq O[(\Gamma m_{\mathrm{Pl}})^{1/2}], \tag{62a}$$

$$\simeq O(\mu^3/m_{\mathrm{Pl}}^2), \tag{62b}$$

$$\simeq 10\ \mathrm{MeV}. \tag{62c}$$

Such a reheat temperature safely returns the Universe to being radiation-dominated before primordial nucleosynthesis, and produces a smooth patch containing an enormous entropy — for $c_2 \simeq 10^{-8}$, $c_1 \simeq 1$, $S_{\mathrm{patch}} \simeq (m_{\mathrm{Pl}}^3/\mu^2 T_{\mathrm{RH}})e^{3N} \simeq 10^{35} \exp(3\times 10^9)$, but does not reheat it to a high enough temperature for baryogenesis. Poor reheating is a problem which plagues almost all potentially viable models of inflation. Achieving $(\delta\varrho/\varrho)_{\mathrm{H}} \lesssim 10^{-4}$ requires the scalar potential to be very flat, which necessarily means that ϕ is very weakly--coupled, and therefore $T_{\mathrm{RH}}(\propto \Gamma^{1/2})$ tends to be very low.

(4) CERN SUSY/SUGR models [79]

Early on members of the CERN theory group recognized that supersymmetry might be of use in protecting the very small couplings necessary in inflationary potentials from being overwhelmed by radiative corrections. They explored a variety of SUSY/SUGR models [79] (and dubbed their brand of inflation 'primordial inflation'). In the process, they encountered a difficulty which plagues almost all supersymmetric models of inflation based upon minimal supergravity theories.

It is usually assumed that at high temperatures the expectation value of an inflating field is at the minimum of its finite temperature effective potential (near $\phi = 0$); then as the Universe cools it becomes trapped there, and then eventually slowly evolves to the low temperature minimum (during which time inflation takes place). In SUSY models $\langle\phi\rangle_T$ is not necessarily zero at high temperatures. In fact in essentially all of their models $\langle\phi\rangle_T > 0$ and the high temperature minimum smoothly evolves into the low temperature minimum (as shown in Fig. 15) [80]. As a result the Universe in fact would never have inflated!

There are two obvious remedies to this problem: (*i*) arrange the model so that $\langle\phi\rangle_T \leqslant 0$ (as shown in Fig. 15), then ϕ necessarily gets trapped near $\phi = 0$; or (*ii*) assume that ϕ is never in thermal equilibrium before the phase transition so that ϕ is not constrained to be in the high temperature minimum of its finite temperature potential at high temperatures. Variants of the CERN models [79] based on these two remedies have been constructed by Ovrut and Steinhardt [81] and Holman, Ramond, and Ross [82].

8. Lessons learned — a prescription for successful new inflation

The unsuccessful models discussed above have proven to be very useful in that they have allowed us to 'write a prescription' for the kind of potential that will successfully implement inflation [65]. The following prescription incorporates these lessons, together with other lessons which have been learned (sometimes painfully). As we will see all but the last of the prescribed features, that the potential be part of a sensible particle physics model, are relatively easy to arrange.

(1) The potential should have an interval which is sufficiently flat so that ϕ evolves slowly (relative to the expansion timescale H^{-1}) — that is, flat enough so that a slow-rollover transition ensues. As we have seen, that means an interval

$$[\phi_s, \phi_e]$$

where

$$|V''| \leqslant 9H^2, \quad |V' m_{Pl}/V| \leqslant (48\pi)^{1/2}.$$

(2) The length of the interval where ϕ evolves slowly should be much greater than $H/2\pi$, the scale of the quantum fluctuations, so that the semi-classical approximation makes sense. (Put another way the interval should be long enough so that quantum fluctuations do not quickly drive ϕ across the interval.) Quantitatively, this calls for

$$|\phi_e - \phi_s| \gg (H\Delta t)^{1/2}(H/2\pi),$$

where Δt is the time required for ϕ to evolve from $\phi = \phi_s$ to $\phi = \phi_e$. (More precisely, the semi-classical change in ϕ in a Hubble time, $\Delta\phi_{\rm Hubble} \simeq -V' m_{Pl}^2/8\pi V$, should be much greater than the increase in $\langle\phi^2\rangle_{\rm QM}^{1/2} \simeq H/2\pi$, due to the addition of another quantum mode; see Ref. [56].)

(3) In order to solve the flatness and homogeneity problems the time required for ϕ to roll from $\phi = \phi_s$ to $\phi = \phi_e$ should be greater than about 60 Hubble times

$$N \equiv \int_{\phi_s}^{\phi_e} H dt \simeq \int_{\phi_s}^{\phi_e} 3H^2 d\phi/(-V') \simeq 3H^2/V'' \geqslant 60.$$

The precise formula for the minimum value of N is given in Eq. (39).

(4) The scalar field ϕ should be smooth in a sufficiently large patch (say size L) so that the energy density and pressure associated with the $(\nabla\phi)^2$ term is negligible:

$$1/2\,(\nabla\phi)^2 \simeq (\phi_0/L)^2 \ll V(\phi_0) \simeq M^4.$$

(Otherwise the $(\nabla\phi)^2$ term will dominate ϱ and p, so that $R(t) \propto t$ — that is, inflation does not occur.) Usually this condition is easy to satisfy, as all it requires is that

$$L \gg \phi_0/M^2 \simeq (\phi_0/m_{Pl})H^{-1};$$

since ϕ_0 is usually $\ll m_{Pl}$, $(\phi_0/m_{Pl})H^{-1} \ll H^{-1}$ — that is, ϕ only need be smooth on a patch comparable to the physics horizon H^{-1}. (I will discuss a case where it is not easy to satisfy —

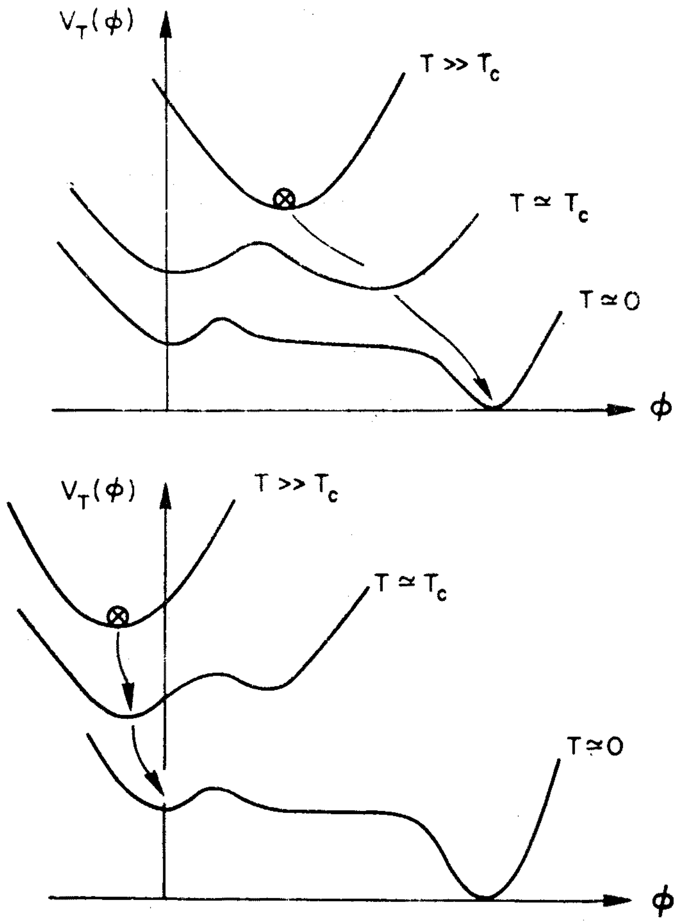


Fig. 15. In SUSY/SUGR models $\langle\phi\rangle_T$ is not necessarily equal to zero. If $\langle\phi\rangle_T > 0$, then there is the danger that $\langle\phi\rangle_T$ smoothly evolves into the zero temperature minimum of the potential, thereby eliminating the possibility of inflation (upper figure). A sure way of preventing this is to design the potential so that $\langle\phi\rangle_T \leqslant 0$ (lower figure)

Linde's chaotic inflation.) Once inflation does begin, any initial inhomogeneities in ϕ are rapidly smoothed by the exponential expansion.

(5a) In order to insure a viable scenario of galaxy formation (and microwave anisotropies of an acceptable magnitude) the amplitude of the adiabatic density perturbations must be of order 10^{-5}–10^{-4}. (In a Universe dominated by weakly-interacting relic particles such as neutrinos or axions, $(\delta\varrho/\varrho)_{\rm MD}$ must be a few $\times$ 10^{-5}.) This in turn results in the constraint

$$\text{few} \times 10^{-5} \simeq (\delta\varrho/\varrho)_{\rm MD} \simeq (2-3)\,(\delta\varrho/\varrho)_{\rm H} \simeq (2-3)\,(H^2/\pi^{3/2}\dot{\phi})_{\rm Galaxy},$$

$$(H^2/\dot{\phi})_{\rm Galaxy} \simeq 10^{-4}.$$

In general, this is by far the most difficult of the constraints (other than sensible particle physics) to satisfy and leads to the necessity of extremely flat potentials. I should add, if one has another means of producing the density perturbations necessary for galaxy formation (e.g., cosmic strings or isothermal perturbations), then it is sufficient to have

$$(H^2/\dot{\phi})_{\text{Galaxy}} \leqslant 10^{-4}$$

(5b) Isothermal perturbations produced during inflation, e.g., as discussed for the case of an axion-dominated Universe, also lead to microwave anisotropies and possibly structure formation. The smoothness of the microwave background dictates that

$$(\delta\varrho/\varrho)_{\text{ISO}} \lesssim 10^{-4},$$

while if they are to be relevant for structure formation

$$(\delta\varrho/\varrho)_{\text{ISO}} \simeq 10^{-5}-10^{-4}.$$

In the case of isothermal axion perturbations this is easy to arrange to have $(\delta\varrho/\varrho)_{\text{ISO}} \ll 10^{-4}$ unless the scale of PQ symmetry is larger than about 10^{18} GeV.

(6a) Sufficiently high reheat temperature so that the Universe is radiation-dominated at the time of primordial nucleosynthesis ($t \simeq 10^{-2}$–10^2 sec, $T \simeq 10$ MeV–0.1 MeV). Only in the case of poor reheating is T_{RH} likely to be anywhere as low as 10 MeV, in which case $T_{\text{RH}} \simeq (\Gamma m_{\text{Pl}})^{1/2}$ and the condition that T_{RH} be $\geqslant 10$ MeV then implies

$$\Gamma \geqslant 10^{-23}\ \text{GeV} \simeq (6.6\times 10^{-2}\ \text{sec})^{-1}.$$

(6b) The more stringent condition on the reheat temperature is that it be sufficiently high for baryongenesis. If baryongenesis proceeds in the usual way [17], the out-of-equilibrium decay of a supermassive particle whose interactions violate B, C, P conservation, then T_{RH} must be greater than about 1/10 the mass of the particle whose out-of-equilibrium decays are responsible for producing the baryon asymmetry. Assuming that this particle couples to ordinary quarks and leptons, its mass must be greater than 10^9 GeV or so to insure a sufficiently longlived proton, implying that the reheat temperature must be greater than about 10^8 GeV (at the very least). On the other hand if the baryon asymmetry can be produced by the decays of the ϕ particles themselves, then

$$n_{\text{B}}/s \simeq (0.75)\,(T_{\text{RH}}/m_\phi)\varepsilon$$

and a very low reheat temperature may be tolerable

$$T_{\text{RH}} \simeq 10^{-10}\varepsilon^{-1} m_\phi,$$

where as usual ε is the net baryon number produced per ϕ-decay.

(7) If ϕ is not a gauge singlet field, as in the case of the original Coleman–Weinberg SU(5) model, one must be careful that 'ϕ rolls in the correct direction'. It was shown that for the original Coleman–Weinberg SU(5) models ϕ might actually begin to roll toward the SU(4) × U(1) minimum of the potential even though the global minimum of the poten-

tial was the $SU(3) \times SU(2) \times U(1)$ minimum [83]. This is because near $\phi = 0$ the $SU(4) \times U(1)$ direction is usually the direction of steepest descent. This is the so-called problem of 'competing phases'. As mentioned earlier, the extreme flatness required to obtain sufficiently small density perturbations probably precludes the possibility that ϕ is a gauge non-singlet, so the problem of competing phases does not usually arise. (Although in SUSY/SUGR models ϕ is often complex and one has to make sure that it rolls in the correct direction.)

(8) In addition to the scalar density perturbations discussed earlier, tensor or gravitational wave perturbations also arise (these correspond to tensor perturbations in the metric $g_{\mu\nu}$) [84]. The amplitude of these perturbations is easy to estimate. The energy density in a given gravitational wave mode (characterized by wavelength λ) is

$$\varrho_{GW} \simeq m_{Pl}^2 h^2/\lambda^2, \tag{63}$$

where h is the dimensionless amplitude of the wave. As each gravitational wave mode crosses outside the horizon during inflation deSitter space produced fluctuations lead to

$$(\varrho_{GW})_{\lambda \simeq H^{-1}} \simeq H^4, \text{ or } h \simeq H/m_{Pl}. \tag{64}$$

While outside the horizon the dimensionless amplitude h remains constant (h behaves like a minimally coupled scalar field), and so each mode enters the horizon with a dimensionless amplitude

$$h \simeq H/m_{Pl}. \tag{65}$$

Gravitational wave perturbations with wavelength of order the present horizon lead to a quadrupole anisotropy in the microwave temperature of amplitude h. The upper limit to the quadrupole anisotropy of the microwave background ($\delta T/T \lesssim 10^{-4}$) leads to the constraint

$$H/m_{Pl} \leqslant 10^{-4}, \qquad M \leqslant O(10^{17}\ \text{GeV}).$$

In turn this leads to a constraint on the reheat temperature (using $g_* \simeq 10^3$)

$$T_{RH} \leqslant g_*^{-1/4} M \leqslant \text{few} \times 10^{16}\ \text{GeV}.$$

(9) One has to be mindful of various particles which may be produced during the reheating process. Of particular concern are stable or very long-lived, NR particles (including other scalar fields which may be set into oscillation and thereafter behave like NR matter). Since $\varrho_{NR}/\varrho_R \propto R(t)$ and today $\varrho_{NR}/\varrho_R \simeq 3 \times 10^4$ or so one has to be careful that ϱ_{NR}/ϱ_R is sufficiently small at early times

$$\varrho_{NR}/\varrho_R \leqslant \begin{cases} 3 \times 10^4 & \text{today} \\ 10^{-8} & T = 1\ \text{GeV} \\ 10^{-18} & T = 10^{10}\ \text{GeV} \end{cases}$$

which is not always easy — just ask any experimentalist about suppressing some effect by 18 orders-of-magnitude!

Of particular concern in supersymmetric models are gravitinos which, if produced, can decay shortly after nucleosynthesis and photodissociate the light elements produced (particularly D and ^{7}Li) [85]. (In fact, the constraint that gravitinos not be overproduced during the reheating process leads to the very restrictive bound: $T_{RH} \leqslant 10^9$ GeV or so.) In supersymmetric models where SUSY breaking is done ala Polonyi [86], the Polonyi field can be set into oscillation [87] and these oscillations which behave like NR matter can come to dominate the energy density of the Universe too early (leading to a Universe which if far too youthful when it cools to 3 K) or even worse decay into dread gravitinos! In sum, one has to be mindful of the weakly-interacting, longlived particles which may be produced during reheating as they may eventually lead to an energy crisis.

(10) In SUSY/SUGR models where the scalar field responsible for inflation is in thermal equilibrium before the inflationary transition, one has to make sure that $\langle\phi\rangle_T$ does not smoothly evolve into the zero temperature minimum of the potential. A sure way of doing this is to arrange to have

$$\langle\phi\rangle_T \leqslant 0,$$

this is the so-called thermal constraint [80].

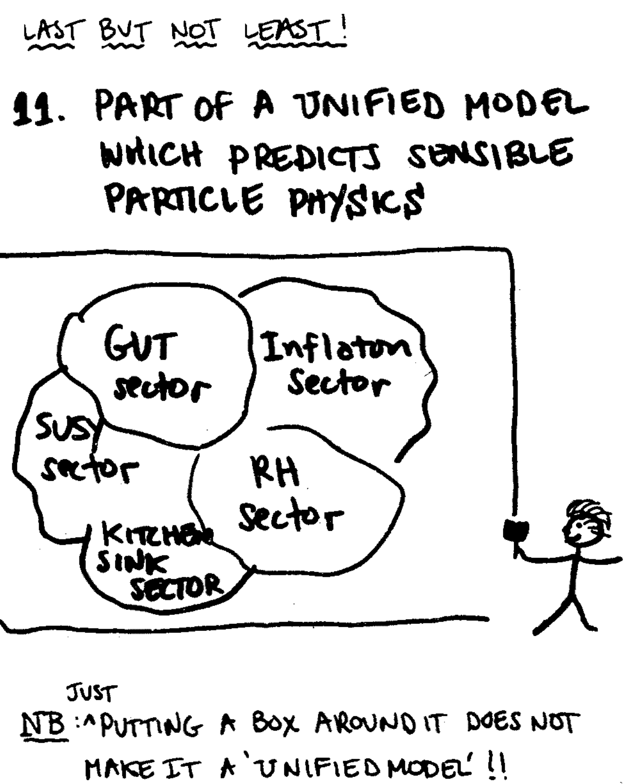


Fig. 16. Constraint (11) in "The Prescription for Successful Inflation"

(11) Last (in my probably incomplete list) but certainly not least, the scalar potential necessary for successful inflation should be but a part of a 'sensible, perhaps even elegant, particle physics model' (see Fig. 16). We do not want cosmology to be the tail that wags the dog!

These conditions are spelled out in more detail in Ref. [65]. In general they lead to a potential which is 'short and squat' and has a dimensionless coupling of order 10^{-15} somewhere. In order that radiative corrections not spoil the flatness, it is all but mandatory that ϕ be a gauge singlet field which couples very weakly to other fields in the theory.

(Suppose that ϕ has a nice, flat potential which will successfully implement inflation and has a ϕ^4 term whose coefficient $\lambda \simeq O(10^{-15})$ (as is usually the case). Now suppose that ϕ couples to another scalar field ψ or to a fermion field f through terms like: $\lambda'\psi^2\phi^2$ and $h\bar{f}f\phi$. One-loop corrections to the $\lambda\phi^4$ term in the scalar potential arise due to the coupling of ϕ to ψ of f: $(\lambda'^2 \ln + h^4 \ln)\phi^4$. In order that they not spoil the flatness of $V(\phi)$ by these 1-loop corrections, the couplings λ' and h must be small: $\lambda' \lesssim \ln^{-1/2}\lambda^{1/2}$; $h \lesssim \ln^{-1/4}\lambda^{1/4}$.)

To give an idea of the kind of potential which we are seeking consider

$$V = V_0 - a\phi^2 - b\phi^3 + \lambda\phi^4.$$

The constraints discussed above are satisfied for the following sets of parameters

$$\text{SET 1}\left\{\begin{aligned} \lambda &\leqslant 4\times 10^{-16} \\ b &\simeq 4\times 10^{7}\lambda^{3/2} m_{\rm Pl} \\ a &\leqslant H^2/40 \simeq 10^{28}\lambda^3 m_{\rm Pl}^2 \\ \sigma &\simeq 3\times 10^{7}\lambda^{1/2} m_{\rm Pl} \\ M &\simeq V_0^{1/4} \simeq 3\times 10^{7}\lambda^{3/4} m_{\rm Pl} \simeq \lambda^{1/4}\sigma \end{aligned}\right.$$

$$\text{SET 2}\left\{\begin{aligned} V &= \lambda(\phi^2-\sigma^2)^2 \; (b = 0,\, a = 2\lambda\sigma^2,\, V_0 = \lambda\sigma^4) \\ \sigma/m_{\rm Pl} &= 1/2,\, 2,\, 3,\, 10 \\ \lambda &= 2\times 10^{-44};\, 5\times 10^{-20},\, 10^{-15},\, 2\times 10^{-15},\, 3\times 10^{-16} \\ M &\simeq \lambda^{1/4}\sigma. \end{aligned}\right.$$

9. *Two simple models that work*

To date a handful of models that satisfy the prescription for successful inflation have been constructed [81, 82, 88–91, 95–97]. Here, I will discuss two particularly simple and illustrative models. The first is an SU(5) GUT model proposed by Shafi and Vilenkin [89] and refined by Pi [90]. (Note, there is nothing special about SU(5); it could just as well be an E6 model.) I will discuss Pi's version of the model. In her model the inflating field $\vec{\phi}$ is a complex gauge singlet field whose potential is of the Coleman–Weinberg form [74]

$$V(\phi) = B[\phi^4 \ln(\phi^2/\sigma^2) + (\sigma^4 - \phi^4)/2]/4, \tag{66}$$

where $\phi = |\vec{\phi}|$ and B arises due to 1-loop radiative corrections from other scalar fields in the theory and is set to be $O(10^{-14})$ in order to successfully implement inflation. (Note, in Eq. (66) I have only explicitly shown the part of the potential relevant for inflation.) Since the 1-loop corrections due to other fields in the model are of order $(\lambda^2 \ln)\phi^4$ (λ is a typical quartic coupling, e.g., $\lambda\phi^2\psi^2$) the dimensionless couplings of ϕ to other fields in the theory must be of order 10^{-7} or so. In her model, $\vec{\phi}$ is the field responsible for Peccei–Quinn symmetry breaking; the vacuum expectation value of $|\vec{\phi}|$ breaks the PQ symmetry and the argument of $\vec{\phi}$ is the axion degree of freedom. In addition, the vacuum expectation value of $|\vec{\phi}|$ induces SU(5) SSB as it leads to a negative mass-squared term for the 24-dimensional Higgs in the theory which breaks SU(5) down to SU(3) × SU(2) × U(1). In order to have the correct SU(5) breaking scale, the vacuum expectation value of $|\vec{\phi}|$ must be of order 10^{18} GeV. In addition to the usual adiabatic density perturbations there are isothermal axion fluctuations of a similar magnitude [69]. The model reheats (barely) to a high enough temperature for baryogenesis. So far the model successfully implements inflation, albeit at the cost of a very small number ($B \simeq 10^{-14}$) whose origin is not explained and whose value is not stabilized (e.g., by supersymmetry).

The second model is a SUSY/SUGR model proposed by Holman, Ramond, and Ross [82] which is based on a very simple superpotential. They write the superpotential for the full theory as

$$W = I + S + G, \tag{67}$$

where I, S, G pieces are the inflation, SUSY, and GUT sectors respectively. For the I piece of the superpotential they choose the very simple form

$$I = (\Delta^2/M)(\phi - M)^2, \tag{68}$$

where $M = m_{\rm Pl}/(8\pi)^{1/2}$, Δ is an intermediate scale, and ϕ is the field responsible for inflation. Their potential has one free parameter: the mass scale Δ. This superpotential leads to the following scalar potential

$$\begin{aligned} V(\phi) &= \exp(|\phi|^2/M^2)\,[|\partial I/\partial\phi + \phi^* I/M^2|^2 - 3|I|^2/M^2] \\ &= \Delta^4 \exp(\phi^2/M^2)\,[\phi^6/M^6 - 4\phi^5/M^5 + 7\phi^4/M^4 - 4\phi^3/M^3 - \phi^2/M^2 + 1]. \end{aligned}$$

Expanding the exponential one obtains

$$V_I(\phi) = \Delta^4(1 - 4\phi^3/M^3 + 6.5\phi^4/M^4 - 8\phi^5/M^5 + \ldots), \tag{69a}$$

$$V_I' = \Delta^4(-12\phi^2/M^3 + 26\phi^3/M^4 - 40\phi^4/M^5 + \ldots). \tag{69b}$$

It is sufficient to keep just the first two terms in $V_I(\phi)$ to solve the equations of motion

$$\phi/M \simeq [12(N(\phi) + 1/3)]^{-1}, \tag{70a}$$

$$H^2/\dot{\phi} \simeq (12\sqrt{3})^{-1}(\Delta/M)^2(\phi/M)^{-2} \simeq (12/\sqrt{3})\,(\Delta/M)^2 N^2. \tag{70b}$$

By choosing $\Delta/M \simeq 9 \times 10^{-5}$ density perturbations of an acceptable magnitude result (and about 2×10^6 e-folds of inflation!). Taking $\Delta/M \simeq 9 \times 10^{-5}$ corresponds to an intermediate scale in the theory of about $\Delta \simeq 2 \times 10^{14}$ GeV — a very suggestive value.

The ϕ field couples to other fields in the theory only by gravitational strength interactions and

$$\Gamma \simeq m_\phi^3/M^2 \simeq \Delta^6/M^5, \tag{71}$$

where $m_\phi^2 \simeq 8e\Delta^4/M^2$.

The resulting reheat temperature is

$$T_{RH} \simeq (\Gamma m_{Pl})^{1/2} \simeq (\Delta/M)^3 M \simeq 10^6 \text{ GeV}. \tag{72}$$

The baryon asymmetry in this model is produced directly by ϕ-decays ($\phi \to H_3\overline{H}_3$; $H_3\overline{H}_3 \to$ q′sl′s; H_3 = color triplet Higgs

$$n_B/s \simeq (0.75\varepsilon)T_{RH}/m_\phi$$

$$\simeq 10^{-1}\varepsilon(\Delta/M).$$

Since $10^{-1}\,\Delta/M \simeq 10^{-5}$, a C, CP violation of about $\varepsilon \simeq 10^{-5}$ is required to explain the observed baryon asymmetry of the Universe ($n_B/s \simeq 10^{-10}$).

Their model satisfies all the constraints for successful inflation except the thermal constraint. They argue that the thermal constraint is not relevant as the interactions of the ϕ field are too weak to put it into thermal equilibrium at early times. They therefore must take the initial value of ϕ ($\equiv \phi_0$) to be a free parameter and assume that in some regions of the Universe ϕ_0 is sufficiently far from the minimum so that inflation occurs ($\phi_0 \lesssim 10^{-3}\,M$). This model is somewhat *ad hoc* in that it contains a special sector of the theory whose sole purpose is inflation. Once again the model contains a small dimensionless coupling (the coefficient of the ϕ^4-term $\simeq 3\times10^{-16}$) or equivalently, a small mass ratio

$$(\Delta/M)^4 \simeq 10^{-16}.$$

Since the model is supersymmetric that small number is stabilized against radiative corrections. Although the small ratio is not explained in their model, its value when expressed as a ratio of mass scales suggests that it might be related to one of the other small dimensionless numbers in particle physics (which also beg explanation)

$$(m_{GUT}/m_{Pl}) \simeq 10^{-4},$$

$$(m_W/m_{Pl}) \simeq 10^{-17},$$

$$g_e \simeq m_e/300 \text{ GeV} \simeq 10^{-6}.$$

While neither of these models is particularly compelling and both have been somewhat contrived to successfully implement inflation, they are at the very least ‘proof of existence’ models which demonstrate that it is possible to construct a simple model which satisfies all the known constraints. Fair enough!

10. Toward the inflationary paradigm

Guth's original model of inflation was based upon a strongly, first order phase transition associated with SSB of the GUT. The first models of new inflation were based upon Coleman–Weinberg potentials, which exhibit weakly-first order transitions. It now appears that the key feature needed for inflation is a very flat potential and that even potentials which lead to second order transitions (i.e., the $\phi = 0$ state is never metastable) will work just as well [92]. Most of the models for inflation now do not involve SSB, at least directly, they just involve the evolution of a scalar field which is initially displaced from the minimum of its potential. (There is a downside to this; in many models inflation is a sector of the theory all by itself.) Since the fields involved are very weakly coupled, thermal corrections can no longer be relied upon to set the initial value of ϕ. Inflation has become much more than just a scenario — it has become an early Universe paradigm!

On the horizon now are models which inflate, but are even more far removed from the original idea of a strongly-first order, GUT SSB phase transition; I'll discuss three of them here. Inflation — that is the rapid growth of our three spatial dimensions, appears to be a very generic phenomenon associated with early Universe microphysics.

(1) Chaotic inflation

Linde [92] has proposed the idea that inflation might result from a scalar field with a very simple potential, say

$$V(\phi) = \lambda\phi^4,$$

which due to 'chaotic initial conditions' (which thus far have not been well-defined) is displaced from the minimum of its potential — in this case $\phi = 0$ (see Fig. 17). With the initial condition $\phi = \phi_0$ this potential is very easy to analyze:

$$N(\phi_0) \equiv \int_{\phi_0}^{0} H dt \simeq \pi(\phi_0/m_{\rm Pl})^2,$$

$$(\delta\varrho/\varrho)_{\rm H} \simeq (H^2/\dot\phi) \simeq (32/3)^{1/2}\lambda^{1/2}N(\phi)^{3/2}.$$

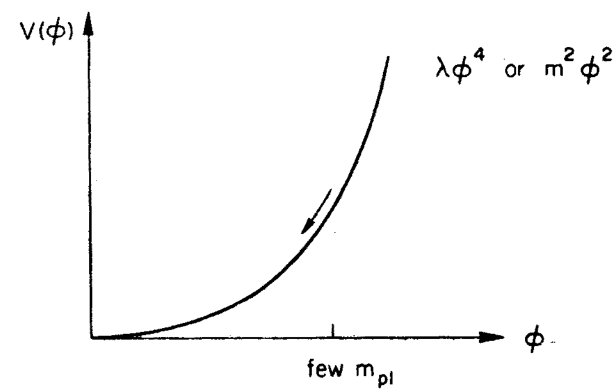


Fig. 17. A potential for 'chaotic inflation'. In Linde's chaotic inflation, due to initial conditions, ϕ is displaced from the minimum of its potential ($\phi = 0$) and inflation occurs as it evolves to $\phi = 0$

In order to obtain density perturbations of the proper amplitude ($\delta\varrho/\varrho \simeq 10^{-4}$) λ must be very small

$$\lambda \simeq 10^{-14}$$

— as usual! In order to obtain sufficient inflation, the initial value of ϕ must be

$$N(\phi_0) \simeq \pi(\phi_0/m_{Pl})^2 \gtrsim 60 \Rightarrow \phi_0 \gtrsim 4.4 m_{Pl}.$$

Both of these two conditions are rather typical of potentials which successfully implement inflation. However, when one talks about truly chaotic initial conditions one wonders if a large enough patch exists where ϕ is approximately constant. Remember the key constraint is that the gradient energy density be small compared to the potential energy

$$(\nabla\phi)^2/2 \ll \lambda\phi_0^4.$$

Labeling the typical dimension of the patch L, the above requirement translates to

$$L \gg \lambda^{-1/2}(m_{Pl}/\phi_0)m_{Pl}^{-1} \simeq 2(\phi_0/m_{Pl})H^{-1},$$

which requires that L be rather large compared to the Hubble radius, therefore seeming to require rather special initial conditions. Still the simplicity of Linde's idea is very appealing.

(Note that the potential $V = \frac{1}{2}m^2\phi^2$ (corresponding to a massive scalar field) is also suitable for inflation. In this case

$$N(\phi_0) \simeq 2\pi(\phi_0/m_{Pl})^2, \tag{74a}$$

$$(\delta\varrho/\varrho)_H \simeq H^2/\dot\phi \simeq 4(\pi/3)^{1/2}(m/m_{Pl})N. \tag{74b}$$

Sufficient inflation requires: $\phi_0 \gtrsim 3m_{Pl}$, and density perturbations of an acceptable magnitude requires: $(m/m_{Pl}) \simeq 10^{-4}/(4N) \simeq 4\times10^{-7}$. This potential has been analyzed by I. Moss (private communication) and L. Jensen (private communication), and more recently by the authors of Ref. [93].)

(2) Induced gravity inflation

Consider the Ginzburg–Landau theory of induced gravity based upon the effective Lagrangian [94].

$$\mathscr{L} = -\varepsilon\phi^2 R/2 - \partial_\mu\phi\partial^\mu\phi/2 - \lambda(\phi^2 - v^2)^2/8, \tag{75}$$

where ε, λ are dimensionless couplings, R is the Ricci scalar, and $v \equiv \varepsilon^{-1/2}(8\pi G)^{-1/2}$. The equation of motion for ϕ is

$$\ddot\phi + 3H\dot\phi + \dot\phi^2/\phi + [V' - 4V/\phi]/(1+6\varepsilon) = 0 \tag{76a}$$

supplemented by

$$H^2[1+(2\dot\phi/\phi)/H] = (3\varepsilon\phi^2)^{-1}[\dot\phi^2/2 + Vt\phi)]. \tag{76b}$$

Successful inflationary scenarios can be constructed for $\phi_0 \ll v$ and for $\phi_0 \gg v$ (ϕ_0 = the initial value of ϕ), so long as $\varepsilon \leqslant 10^{-2}$ and $\lambda \simeq O(10^{-12}\text{–}10^{-16})$ [95, 96]. The small dimensionless coupling constant required in the scalar potential is by now a very familiar condition.

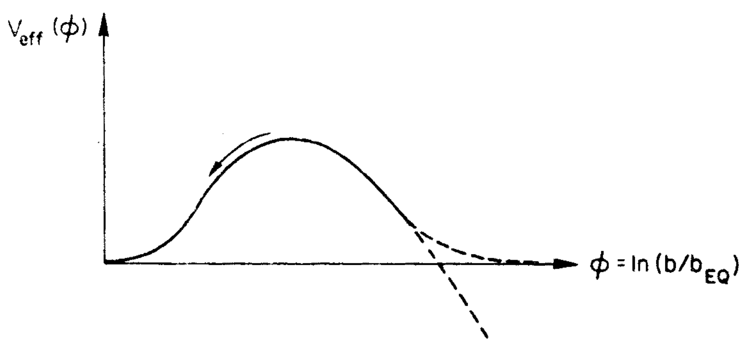
(3) The compactification transition

Ever increasing numbers of physicists are pursuing the idea that unification of the forces may require additional spatial dimensions (or as the optimist would say, unification of the forces is evidence for extra dimensions!), e.g., Kaluza–Klein theories, supergravity theories, and most recently, superstring theories. We know experimentally that these extra dimensions must be very small ($\ll 10^{-17}$ cm) and indeed in most theories the extra dimensions form a compact manifold of characteristic size 10^{-34} cm or so. If our space-time is truly more than four dimensional, then we have yet another problem to add to our list of puzzling cosmological facts — the extreme smallness of the extra spatial dimensions, some $62 \simeq \log(10^{28}\ \text{cm}/10^{-34}\ \text{cm})$ or so orders of magnitude smaller than the more familiar three spatial dimensions. The possible use of inflation to explain this largeness problem has not escaped the attention of researchers in this field.

In these theories there is a natural candidate for the 'inflating field' (which is also automatically a gauge singlet) — the radius of the extra dimensions. If there are extra dimensions there must be some dynamics which determine their present, equilibrium size ($\equiv b_{eq}$), and in principle one should be able to construct an effective potential associated with the size of the extra dimensions

$$V_{eff} = V(\phi), \qquad \phi = \ln(b/b_{eq})$$

(see Fig. 18). (The substitution $\phi = \ln(b/b_{eq})$ results in the usual kinetic term for ϕ when the higher dimensional Einstein equations are written down.) If the extra dimensions are



... ...ories with additional spatial dimensions there must be an effective potential associated with the size of the extra dimensions (shown here schematically). One might expect that very early on ($t < 10^{-43}$ sec) the size of the extra dimensions is displaced from its equilibrium value ($\equiv b_{eq}$), due to finite temperature corrections, initial conditions, or whatever. It is speculated that inflation might occur as the size of the extra dimensions evolves to its equilibrium value, thereby solving both the usual cosmological puzzles and the puzzle of why the extra dimensions are so small compared to our three familiar spatial dimensions displaced from their equilibrium value — an idea which seems not at all unreasonable since very early on ($t < 10^{-43}$ sec) one might expect all the dimensions to be on equal footing, then while they are evolving to their equilibrium value ($\phi = 0$) the Universe will be endowed with a large potential energy (and may very well inflate), thereby explaining the largeness of our three spatial dimensions as well as the usual cosmological puzzles. Inflationary models involving the compactification transition have already been investigated and the results are encouraging [97].

11. Loose ends

Inflation offers the possibility of making the present state of the Universe (on scales as large as our Hubble radius) insensitive to the initial data for the Universe. Since we have little hope of ever knowing what the initial data were this is a very attractive proposition. It has by no means yet achieved that lofty goal. There are a number of loose ends (and perhaps even a loose thread which may unravel the entire tapestry). I will briefly mention a few of them here.

(*i*) 'Who is ϕ?'

Inflationary models exist in which the scalar field ϕ: effects SSB of the GUT [89, 90] effects SSB of SUSY [81], induces Newton's constant (in a Landau-Ginzburg model of induced gravity) [95, 96], is $\sim \ln(r_X/r_{XEQ})$ (where r_X is the radius of compactified extra dimensions) in theories with extra dimensions which become compactified [97], is $\propto$ (scalar curvature)$^{1/2}$ in R^2 theories of gravity [98], is just some 'random' scalar field [92], or is merely in the theory to effect inflation [79, 82]. Given the number of different kinds of inflationary scenarios which exist, it seems as though inflation is generic to early Universe microphysics, occurring whenever a weakly-coupled scalar field finds itself displaced from the minimum of its potential. Clearly, a key question at this point is just how 'the inflation sector' of the theory fits into the Big Picture!

(*ii*) What determines the initial value of ϕ?

One thing is certain, and that is that ϕ must be very weakly-coupled, as quantified by its small dimensionless coupling constant. Because of this fact, it is almost certain that ϕ was not initially in thermal contact with the rest of the Universe and so the initial value of $\phi(\equiv \phi_0)$ is unlikely to be determined by thermal considerations (in the earliest models of new inflation, ϕ_0 was determined by thermal considerations, however these models resulted in density perturbations of an unacceptably large amplitude). At present ϕ_0 must be taken as initial data. Some have argued that ϕ_0 might be determined in an anthropic--like way, as regions of the Universe where ϕ_0 is sufficiently far displaced from equilibrium will undergo inflation and eventually occupy most of the physical volume of the Universe. Perhaps the wave-function of the Universe approach will shed some light on the initial distribution of the scalar field ϕ. Or it could be that due to 'as-of-yet unknown dynamics' ϕ was indeed in thermal equilibrium at a very early epoch. It goes without saying that it is crucial that ϕ be initially displaced from its minimum.

(*iii*) Validity of the semi-classical equations of motion for ϕ

While it may seem perfectly plausible that ϕ evolves according to its semi-classical equations of motion, the validity of this assumption has troubled inflationists from the 'dawn of new inflation'. While a full quantum field theory treatment of inflation is very difficult and has not been effected, a number of specific issues have been addressed. Several authors have studied the role of inhomogeneities in ϕ, and have found that for the very weakly-coupled fields one is dealing with, mode coupling is not important and the individual modes are quickly smoothed by the exponential expansion of their physical wavelengths [99]. I already mentioned the necessity of having ϕ smooth over a sufficiently large region so that the gradient terms in the stress energy do not dominate.

The effect of quantum fluctuations on the evolution of ϕ has been studied in some detail by Guth and Pi [91], Fischler et al. [98], Linde [75], Vilenkin and Ford [76], Semenoff and Weiss [101], and Evans and McCarthy [102]. The basic conclusion that one draws from the work of these authors is that the use of the semi-classical equations of motion is valid so long as $\phi_{cl} \gg \Delta\phi_{QM} \simeq N^{1/2}\, H/2\pi$, which is almost always satisfied for the very flat potentials of interest to inflationists (at least for the last 50 or so e-folds which affect our present Hubble volume). (More precisely, the semiclassical change in ϕ in a Hubble time, $\Delta\phi_{Hubble} \simeq -V'/3H^2 \simeq -V'm_{Pl}^2/(8\pi V)$, should be much greater than the increase in $\langle\phi^2\rangle_{QM}^{1/2}$, $\simeq H/2\pi$, due to the addition of another quantum mode; see Bardeen et al. [56]. At present the validity of the semi-classical equations of motion seems to be reasonably well established.

(*iv*) No hair conjectures

While inflation has been touted from the very beginning as making the present state of the Universe insensitive to the initial spacetime geometry, not much has been done to justify this claim until very recently. As I mentioned earlier, inflation is nearly always analyzed in the context of a flat, FRW cosmological model, making such a claim somewhat dubious. However, it has now been shown that all of the homogeneous models (with the exception of the highly-closed models) undergo inflation, isotropize and remain isotropic to the present epoch providing that the model would have inflated the requisite 60 or so e-folds in the absence of anisotropy [103].

The proof of this result involves three parts. First, Wald [104] demonstrated that all homogeneous models with a positive cosmological term asymptotically approach deSitter (less the aforementioned highly-closed models which recollapse before the cosmological term becomes relevant). Wald's result follows because all forms of 'anisotropy energy density' decrease with increasing proper volume element, whereas the cosmological term remains constant, and so eventually triumphs. Of course, inflationary models do not, in the strictest sense, have a cosmological term, rather they have a positive vacuum energy as long as the scalar field is displaced from the minimum of its potential. Thus the dynamics of the scalar field comes into play: does ϕ stay displaced from the minimum of its potential long enough so that the vacuum energy comes to dominate? Due to the presence of anisotropy the expansion rate is *greater* than if there were only vacuum energy density, and so the friction felt by ϕ as it tries to roll (the $3H\dot{\phi}$ term) is greater and it takes ϕ longer to evolve to its minimum than without anisotropy. For this reason the Universe does become vacuum dominated before the vacuum energy disappears, and in fact the Universe inflates slightly longer in the presence of anisotropy (one or two e-folds) [105]. Finally, is the anisotropy reduced sufficiently so that the Universe today is still nearly isotropic? As it turns out, the requisite 60 or so e-folds needed to solve the other conundrums reduces the growing modes of anisotropy sufficiently to render them small today.

Allowing for inhomogeneous initial spacetimes makes matters much more difficult. Jensen and Stein-Schabes [106] and Starobinskii [107] have proven the analogue of Wald's theorem for spacetimes which are negatively-curved. Jensen and Stein-Schabes [106] have gone on to conjecture that spacetimes which have sufficiently large regions of negative curvature will undergo inflation, resulting in a Universe today which although not globally homogeneous, at least contains smooth volumes as large as our current Hubble volume.

Does this improve the situation that Collins and Hawking [13] discussed in 1973? While the work of Jensen and Stein-Schabes [106] seems to indicate that many inhomogeneous spacetimes undergo inflation and even leads one to speculate that the measure of the set of initial spacetimes which eventually inflate is non-zero, it is not possible to draw a definite conclusion without first defining a measure on the space of initial data. In fact, as Penrose [25] pointed out there is at least one way of defining the measure such that this is not the case. Consider the set of all Cauchy data at the present epoch; intuitively it is clear that those spacetime slices which are highly irregular are the rule, and those which are smooth in regions much larger than our current Hubble volume are the exception. Defining the measure today, it seems very reasonable that the smooth spacetime slices are a set of measure zero. Now evolve the spacetimes back to some initial epoch (for example $t = 10^{-43}$ sec). Using the seemingly very reasonable measure defined today and the mapping back to 'initial' spacetimes, one could argue that the set of initial data which inflate is still of measure zero. While I believe that this argument is technically correct, I also believe that it is silly. First, upon close examination of all of those initial spacetimes which led to spacetimes today without smooth regions as large as our present Hubble volume, one would presumably find that the scalar field responsible for inflation would be very close to the minimum of its potential (in order that they not inflate) — not a very generic initial condition. Secondly, if one adopts the point-of-view of an evolving Universe which has an 'initial epoch' (and not everyone does), then there is a preferred epoch at which one would define a measure — the 'initial epoch', and at that epoch I believe any reasonably defined measure would lead to the set of initial spacetimes which inflate being of non-zero measure.

Although it is not possible *yet* to claim rigorously that inflation has resolved the problem of the seemingly special initial data required to reproduce the Universe we see today (at least within our Hubble volume), I think that any fairminded person would admit that it has improved the situation dramatically. Extrapolating from the solid results that exist, it seems to me that starting with a general inhomogeneous spacetime, there will exist regions which undergo inflation and which today are much larger than our present Hubble volume, thereby accounting for the smooth region we find ourselves in. From a more global perspective, one might expect that on scales $\gg H^{-1}$ the Universe would be highly irregular. (The evolution of a model universe which is isotropic and homogeneous except for one spherically-symmetric region of false vacuum (where $\phi \neq \sigma$) has been studied by the authors of Ref. [108]. The results are interesting in that they begin to address the problem of general initial conditions. The vacuum-dominated bubble becomes causally detached from the rest of the spacetime, becoming a 'child Universe' spawned by inflation.)

(*v*) The present vanishingly small value of the cosmological constant

Inflation has shed no light on this difficult and very fundamental puzzle (nor has anything else for that matter!). In fact, since inflation runs on vacuum energy so to speak, the fate of inflation hinges upon the resolution of this puzzle. For example, suppose there were a grand principle that dictated that the vacuum energy of the Universe is always zero, or that there were an axion-like mechanism which operated and ensured that any cosmological constant rapidly relaxed to zero; either would be a disaster to inflation, shorting out its source of power-vacuum energy. (Another possibility which has received a great deal of attention recently is the possibility that deSitter space might be quantum mechanically unstable [109] — of course, if its lifetime were at least 60 some e-folds that would not necessarily adversely affect inflation.)

12. *Inflation confronts observation*

No matter how appealing a theory may be, it must meet and pass the test of experimental verification. Experiment and/or observation is the final arbiter. One of the few blemishes on early Universe physics is the lack, thus far, of experimental/observational tests of the many beautiful and exciting predictions. That situation is beginning to change as the field starts to mature. Inflation is one of the early Universe theories which is becoming amenable to verification or falsification. Inflation makes the following very definite predictions (postdictions?):

(*i*) $\Omega = 1.0$ (more precisely, $R_{\text{curv}} = R(t)\,|k|^{-1/2} = H^{-1}/|\Omega-1|^{1/2} \gg H^{-1}$,

(*ii*) Harrison-Zel'dovich spectrum of constant curvature perturbations (and possibly isocurvature perturbations as well) and tensor mode gravitational wave mode perturbations.

The prediction of $\Omega = 1.0$ together with the primordial nucleosynthesis constraint on the baryonic contribution, $0.014 \lesssim \Omega_B h^2 \lesssim 0.035 \lesssim 0.15$ (Ref. [6]), suggests that most of the matter in the Universe must be nonbaryonic. The simplest and most plausible possibility is that it exists in the form of relic WIMPs (Weakly-Interacting Massive Particles, e.g., axions, photinos, neutrinos; for a review, see Ref. [110]). Going a step further, these two original predictions then lead to testable consequences:

(*iii*) $H_0 t_0 = 2/3$ (providing that the bulk of the matter in the Universe today is in the form of NR particles)

The observational data on both H_0 and t_0 are far from being definitive: $H_0 \simeq 40$ –100 km sec^{-1} Mpc^{-1} and $t_0 \simeq 12$–20 Gyr, implying only that $H_0 t_0 \simeq 0.5$–2.0.

(*iv*) $\Omega = 1.0$.

All of the dynamical observations suggest that the fraction of critical density contributed by matter which is clumped on scales $\lesssim$ 10–30 Mpc is only about: $\Omega_{\lesssim 30} \simeq 0.2 \pm 0.1$

(± 0.1 is not meant to be a formal error estimate, but indicates the spread in the observations) (see Ref. [8]). If inflation is not to be falsified, that leaves but two options: (1) the observations are somehow misleading or wrong; or (2) there exists a component of energy density which is smoothly distributed on scales $\lesssim$10–30 Mpc (and therefore would not be reflected in the dynamical determinations). Candidates for the smooth component include: relic, light neutrinos, which by virtue of the large length scale ($\lambda_\nu \simeq 13\, h^{-2}$ Mpc) on which neutrino perturbations are damped by freestreaming, would likely still be smooth on these scales; relic relativistic particles produced by the recent decay of an unstable WIMP species [111]; a relic cosmological term [112]; 'failed galaxies', referring to a population of galaxies which have the same mix of dark matter to baryons, but are more smoothly distributed and are too faint to observe (at least thus far) [113]; relic population of light strings — either fast moving non-intercommuting strings or a tangled network of non-Abelian strings [114]. All of these smooth component scenarios have testable consequences [115] — their predictions for $H_0 t_0$ differ from 2/3; the growth of perturbations is different; the evolution of the cosmic scale factor $R(t)$ is different from the matter-dominated model and various kinematic tests (magnitude-redshift, angular size-redshift, lookback time-redshift, proper volume element-redshift, etc.) can in principle differentiate between them.

(*v*) Microwave fluctuations

Both the scalar and tensor metric perturbations lead to fluctuations in the CMBR on large angular scales ($\gg 1°$). On such large scales causal microphysical processes (such as reionization) cannot have erased the primordial fluctuations, and so if ever present, they must still be there. The scalar perturbations (if they have anything to do with structure formation) must be of amplitude $\gtrsim$ few $\times\ 10^{-6}$, which is within a factor of 10 or less of the current upper limits on these scales.

(*vi*) Two detailed stories of structure formation

The simplest possibility, namely that most of the mass density is in relic WIMPs ($\Omega_{\rm WIMP} = 1.0 - \Omega_{\rm B} \simeq 0.9$) leads to two very detailed scenarios of structure formation: hot dark matter (the case where the dark matter is neutrinos) and cold dark matter (essentially any other WIMP as dark matter). At present, the numerical simulations of these scenarios are sufficiently definite that it is possible to falsify them — and in fact, both of these simplest scenarios have difficulties (see the recent review by White [116]). In the hot dark matter case it is forming galaxies early enough. The large-scale structure which evolves in this case (voids, superclusters, froth) qualitatively agrees with what is observed; however, in order to get agreement with the galaxy-galaxy correlation function, galaxies must form very recently (redshifts $\lesssim 1$) in contradiction to all the galaxies (redshifts as large as 3.2) and QSO's (redshifts as large as 4.0) which are seen at redshifts $\gtrsim 1$.

With cold dark matter the simulations can nicely reproduce galaxy clustering, most of the observed properties of galaxies (masses and densities, rotation curves, etc.) [117]. However, the simulations do not seem to be able to produce sufficient large-scale structure. In particular, they fail to account for the amplitude of the cluster-cluster correlation function (by a factor of about 3), large amplitude, large-scale peculiar velocities, and voids. (In fairness I should mention that our knowledge of large-scale structure of the Universe is still very fragmentary, with the first moderate sized ($\sim 10^4$), 3-dimensional surveys having just recently been completed.) In order to account for $\Omega = 1.0$, galaxy formation must be biased (i.e., only density-averaged peaks greater than some threshold, typically 2–3σ, are assumed to evolve into galaxies which we see today, the more typical 1σ peaks resulting in 'failed galaxies' for some reason or another; see Ref. [113]).

(The situation with respect to large scale structure is becoming more interesting every moment. Several groups have now reported large-amplitude (600–1000 km sec^{-1}) peculiar velocities on large scales ($\sim 50h^{-1}$ Mpc) (Burştein et al. [118]; Collins et al. [119]). Such large peculiar velocities are very difficult, if not impossible, to reconcile with either hot or cold dark matter (or even smooth component models) and the Zel'dovich spectrum (see Ref. [120]). If these data hold up they may pose an almost insurmountable obstacle to any scenario with the Zel'dovich spectrum of density perturbations. The frothy structure observed in the galaxy distribution by de Lapparent et al. [121], galaxies distributed on the surfaces on large ($\sim 30h^{-1}$ Mpc), empty bubbles, although somewhat more qualitative, also seems difficult to reconcile with cold dark matter.)

There are a number of observations/experiments which can and will be done in the next few years and which should really put the inflationary scenario to the test. They include improved sensitivity measurements of the CMBR anisotropy. The microwave background anisotropies predicted in the hot dark matter scenario are very close to the observational upper limits on angular scales of both 5 or so arcminutes and $\gtrsim$ few degrees [12]. With cold dark matter, the predictions are a factor of 3–10 away from the observational limits (for the isocurvature spectrum, the quadrupole upper limit may actually rule out this possibility; see, Efstathiou and Bond [122]). An improvement in sensitivity to microwave anisotropies of the order of 3–10 could either begin to confirm one of the scenarios or rule them both out, and is definitely within the realm of experimental reality (Wilkinson in Ref. [5]).

The relic WIMP hypothesis for the dark matter can also be tested. While it was once almost universally believed that all WIMP dark matter candidates were, in spite of their large abundance, essentially impossible to detect because of the feebleness of their interactions, a number of clever ideas have recently been suggested (and are being experimentally implemented) for detecting axions [123], photinos, sneutrinos, heavy neutrinos, etc. [124]. Results and/or limits will be forth coming soon. With the coming online of the Tevatron at Fermilab, the SLC at SLAC, and hopefully the SSC it is possible that one of the candidates may be directly produced in the lab. Experiments to detect neutrino masses in the eV mass range also continue.

A geometric measurement of the curvature of the Universe (which uses the dependence of the comoving volume element as a function of redshift) has recently been made by Loh and Spillar [125]. Their preliminary results indicate $\Omega = 0.9^{+0.7}_{-0.5}$ (95% confidence) (for a matter-dominated model). This technique appears to have great cosmological leverage and looks very promising (especially the value!) — far more promising than the traditional approach of determining the density of the Universe through the deceleration parameter q_0.

Another area with great potential for improvement is 3d surveys of the distribution of galaxies. The largest redshift surveys at present contain only a few 1000 galaxies, yet have been very tantalizing, indicating evidence of voids and froth-like structure to the galaxy distribution [121]. The large, automated surveys which are likely to be done in the next decade could very well lead to a quantum leap in our understanding of the large scale features of the Universe and help to provide hints as to how they evolved.

The peculiar velocity field of the Universe is potentially a very valuable and direct probe of the density field of the Universe:

$$|\delta v_k| = |\dot{\delta}_k/k| \quad (= (\lambda H/2\pi)\delta_k \quad \text{for} \quad \Omega = 1), \tag{77}$$

$$(\delta v/c)_\lambda \simeq (\lambda/10^4 h^{-1}\ \mathrm{Mpc})\,(\delta\varrho/\varrho)_\lambda, \tag{78}$$

where δ_k and δv_k are the k-th Fourier components of $\delta\varrho/\varrho$ and $\delta v/c$, respectively. The very recent measurements which indicate large amplitude peculiar velocities on scales of $\sim 50h^{-1}$ Mpc are surprising in that they indicate substantial power on these scales, and are problematic to almost every scenario of structure formation. Should they be confirmed they will provide a very acute test of structure formation in inflationary models.

Of course, theorists are very accommodating and have already started suggesting alternatives to the simplest scenarios for structure formation. As I mentioned earlier, scenarios with a smooth component to the energy density have been put forward to solve the Ω problem. Cosmic strings present a radically different approach to structure formation with their non-gaussian spectrum of density fluctuations (for further discussion see Ref. [131]). (It is interesting to note that cosmic strings of the right 'weight' ($G\mu \simeq 10^{-6}$ or so, where μ is the string tension) seem to be somewhat incompatible with inflation, as they must necessarily be produced after inflation and require reheating to a temperature $\gtrsim \mu^{1/2} \simeq 10^{16}$ GeV which seems difficult.) Somewhat immodestly I mention a proposal Silk and I recently made: 'double inflation' [126]. While the Harrison-Zel'dovich spectrum is a beautiful prediction both because of its geometric simplicity and its definiteness, it may well be in conflict with observation because it does not seem to allow enough power on large scales to account for the recent observations of froth and large amplitude peculiar velocities. In the variant we have proposed there are two (or more) episodes of inflation, with the final episode lasting only about 40 e-folds or so, so that the amplitudes of perturbations on large scales are set by the first episode and those on small scales by the second episode. That there might be multiple episodes of inflation seems quite plausible given the number of different microphysical scenarios which result in inflation. Arranging the most recent episode to last for only 40 or so e-folds so that some of the scales within our present Hubble volume crossed outside the horizon during an earlier episode of inflation is a more formidable task — but not an impossible or implausible one! If this can be arranged then it is possible to have very large amplitude perturbations on small scales (of order 10^{-1}) and larger than usual amplitude perturbations on large scales (nearly saturating the large scale microwave limits), thereby providing enough power for the large scale structure which the recent redshift surveys and peculiar velocity measurements indicate. The large amplitude perturbations on small scales allow for very early galaxy formation (and reionization of the Universe, thereby erasing the CMBR fluctuations on small angular scales). If the second episode of inflation proceeds via the nucleation of bubbles, they might directly explain the froth-like structure recently reported by de Lapparent et al. [121].

13. *Epilogue*

Despite the absence of a compelling model which successfully implements the inflationary paradigm, inflation remains a very attractive means of accounting for a number of very fundamental cosmological facts by microphysics that we have some understanding of: namely, scalar field dynamics at sub-Planck energies. The lack of a compelling model at present must be viewed in the light of the fact that at present we have no compelling, detailed model for the '*Theory of Everything*' and the fact that despite vigorous scrutiny there has yet to be a *No-Go Theorem* for inflation unearthed. It is my belief that the undoing of inflation (if it should come) will involve observations and not theory. At the very least. The Inflationary Paradigm is still worthy of further consideration — and I hope that I have convinced you of that fact!

Due to space/time limitations my review of inflation has necessarily been incomplete, for which I apologize. I refer the interested reader to the more complete reviews by Linde [127]; by Abbott and Pi [128]; by Steinhardt [129]; by Brandenberger [130]; by Bonometto and Masiero [138]; and by Blau and Guth [132]. My prescription for successfully implementing inflation borrows heavily from the paper written by Steinhardt and myself [65]. This work was supported in part by the DoE (at Chicago) and by my Alfred P. Sloan Fellowship.

Editorial note. This article was proofread by the editors only, not by the author.

REFERENCES

[1] S. Djorgovski, H. Spinrad, P. McCarthy, M. Strauss, *Astrophys. J.* **299**, L1 (1985).

[2] B. A. Peterson *et al.*, *Astrophys. J.* **260**, L27 (1982); D. Koo, *Astron. J.* **90**, 418 (1985); C. Hazard, R-G. McMahon, W. L. W. Sargent, *Nature* **322**, 38 (1986); *Nature* **322**, 40 (1986); S. J. Warren et al., *Nature* **325**, 131 (1987).

[3] P. J. E. Peebles, *Physical Cosmology*, Princeton Univ. Press, 1971; S. Weinberg, *Gravitation and Cosmology*, Wiley, NY 1972, chapter 15; *Physical Cosmology*, eds. J. Audouze, R. Balian, D. N. Schramm, North-Holland, Amsterdam 1980.

[4] J. Peterson, P. Richards, T. Timusk, *Phys. Rev. Lett.* **55**, 332 (1985); G. F. Smoot et al., *Astrophys. J.* **291**, L23 (1985); D. Meyer, M. Jura, *Astrophys. J.* **276**, L1 (1984); D. Woody, P. Richards, *Astrophys. J.* **248**, 18 (1981).

[5] D. Wilkinson, in *Inner Space/Outer Space*, eds. E. Kolb et al., Univ. of Chicago Press, Chicago 1986.

[6] J. Yang, M. S. Turner, G. Steigman, D. N. Schramm, K. Olive, *Astrophys. J.* **281**, 493 (1984); A. Boesgaard, G. Steigman, *Ann. Rev. Astron. Astrophys.* **23**, 319 (1985).

[7] J. Huchra, A. Sandage, G. Tammann, in *Inner Space/Outer Space*, eds. E. Kolb et al., Univ. of Chicago Press, Chicago 1986; J. Gunn, B. Oke, *Astrophys. J.* **195**, 255 (1975); J. Kristian, A. Sandage, J. Westphal, *Astrophys.* **221**, 383 (1978); R. Buta, G. de Vaucouleurs, *Astrophys. J.* **266**, 1 (1983); W. D. Arnett, D. Branch, J. C. Wheeler, *Nature* **314**, 337 (1985).

[8] S. Faber, J. Gallagher, *Ann. Rev. Astron. Astrophys.* **17**, 135 (1979); V. Trimble, *Ann. Rev. Astron. Astrophys.* **25**, in press (1987).

[9] I. Iben, Jr., *Ann. Rev. Astron. Astrophys.* **12**, 215 (1974); A. Sandage, *Astrophys. J.* **252**, 553 (1982); D. N. Schramm, in *Highlights of Astronomy*, ed. R. West, Reidel, Dordrecht 1983, vol. 6; I. Iben, A. Renzini, *Phys. Rep.* **105**, 329 (1984).

[9a] C. B. Collins, M. J. Perry, *Phys. Rev. Lett.* **34**, 1353 (1975).

[10] G. F. R. Ellis, *Ann. NY. Acad. Sci.* **336**, 130 (1980).

[11] R. Sachs, A. Wolfe, *Astrophys. J.* **147**, 73 (1967).

[12] N. Vittorio, J. Silk, *Astrophys. J.* **285**, L39 (1984); J. R. Bond, G. Efstathiou, *Astrophys. J.*

285, L44 (1984); J. Silk, in *Inner Space/Outer Space*, eds. E. W. Kolb et al., University of Chicago Press, Chicago 1986.

[13] C. B. Collins, S. Hawking, *Astrophys. J.* **180**, 317 (1973).

[14] See, e.g., P. J. E. Peebles, *Large-Scale Structure of the Universe*, Princeton Univ. Press, Princeton 1980; G. Efstathiou, J. Silk, *Fund. Cosmic Phys.* **9**, 1 (1983); S. D. M. White, in *Inner Space/Outer Space*, eds. E. Kolb et al., Univ. of Chicago Press, Chicago 1986.

[14a] R. H. Dicke, P. J. E. Peebles, in *General Relativity: An Einstein Centenary Survey*, eds. S. W. Hawking, W. Israel, Cambridge University Press, Cambridge 1979.

[15] G. Steigman, *Ann. Rev. Astron. Astrophys.* **14**, 339 (1976).

[16] G. G. Ross, *Grand Unified Theories*, Benjamin Cummings, Menlo Park 1984.

[17] E. W. Kolb, M. S. Turner, *Ann. Rev. Nucl. Part. Sci.* **33**, 645 (1983).

[18] G. 't Hooft, *Nucl. Phys.* **B79**, 276 (1974); A. M. Polyakov, *JETP Lett.* **20**, 194 (1974).

[19] T. W. B. Kibble, *J. Phys.* **A9**, 1387 (1976); J. Preskill, *Phys. Rev. Lett.* **43**, 1365 (1979); Ya. B. Zel'dovich, M. Yu. Khlopov, *Phys. Lett.* **79B**, 239 (1978).

[20] J. Preskill, *Ann. Rev. Nucl. Part. Sci.* **34**, 461 (1984); M. S. Turner, in *Monopole '83*, ed. J. Stone, Plenum Press, NY 1984.

[21] A. Guth, S-H. H. Tye, *Phys. Rev. Lett.* **44**, 631 (1980).

[22] A. Guth, *Phys. Rev.* **D23**, 347 (1981).

[23] S. Bludman, M. Ruderman, *Phys. Rev. Lett.* **38**, 255 (1977); Ya. B. Zel'dovich, *Sov. Phys. Usp.* **11**, 381 (1968).

[24] C. W. Misner, *Astrophys. J.* **151**, 431 (1968); in *Magic Without Magic*, ed. J. Klauder, Freeman, San Francisco 1972; R. Matzner, C. W. Misner, *Astrophys. J.* **171**, 415 (1972).

[25] R. Penrose, in *General Relativity: An Einstein Centenary Survey*, eds. S. W. Hawking, W. Israel, Cambridge University Press, Cambridge 1979.

[26] B. de Witt, *Phys. Rev.* **90**, 357 (1953).

[27] L. Parker, *Nature* **261**, 20 (1976).

[28] Ya. B. Zel'dovich, *JETP Lett.* **12**, 307 (1970).

[29] A. A. Starobinskii, *Phys. Lett.* **91B**, 99 (1980).

[30] P. Anderson, *Phys. Rev.* **D28**, 271 (1983).

[31] J. M. Hartle, B.-L Hu, *Phys. Rev.* **D20**, 1772 (1979).

[32] M. V. Fischetti, J. Hartle, B.-L. Hu, *Phys. Rev.* **D20**, 1757 (1979).

[33] J. M. Hartle, S. W. Hawking, *Phys. Rev.* **D28**, 2960 (1983).

[34] B. J. Carr, M. J. Rees, *Nature* **278**, 605 (1979); J. D. Barrow, F. Tipler, *The Anthropic Cosmological Principle*, Oxford University Press, Oxford 1986.

[35] D. Lindley, Fermilab preprint (1985).

[36] A. Guth, E. Weinberg, *Nucl. Phys.* **B212**, 321 (1983); S. Hawking, I. Moss, J. Stewart, *Phys. Rev.* **D26**, 2681 (1982).

[37] A. Linde, *Phys. Lett.* **108B**, 389 (1982).

[38] A. Albrecht, P. J. Steinhardt, *Phys. Rev. Lett.* **48**, 1220 (1982).

[39] S. Coleman, F. de Luccia, *Phys. Rev.* **D21**, 3305 (1980); S. Coleman, *Phys. Rev.* **D15**, 2929 (1977); C. G. Callan, S. Coleman, *Phys. Rev.* **D16**, 1762 (1977).

[40] S. W. Hawking, I. G. Moss, *Phys. Lett.* **110B**, 35 (1982); L. Jensen, P. J. Steinhardt, *Nucl. Phys.* **239B**, 176 (1984); K. Lee, E. J. Weinberg, *Nucl. Phys.* **B267**, 181 (1986).

[41] A. Albrecht, P. J. Steinhardt, M. S. Turner, F. Wilczek, *Phys. Rev. Lett.* **48**, 1437 (1982).

[42] L. Abbott, E. Farhi, M. Wise, *Phys. Lett.* **117B**, 29 (1982).

[43] A. Dolgov, A. Linde, *Phys. Lett.* **116B**, 329 (1982).

[44] P. J. Steinhardt, M. S. Turner, *Phys. Rev.* **D29**, 2162 (1984).

[45] J. R. Gott, *Nature* **295**, 304 (1982); J. R. Gott, T. S. Statler, *Phys. Lett.* **136B**, 157 (1984).

[46a] M. S. Turner, *Phys. Lett.* **115B**, 95 (1982); J. Preskill, in *The Very Early Universe*, eds. G. Gibbons, S. Hawking, S. Siklos, Cambridge Univ. Press, Cambridge 1983; G. Lazarides, Q. Shafi, W. P. Trower, *Phys. Rev. Lett.* **49**, 1756 (1982).

[46b] G. Lazarides, Q. Shafi, in *The Very Early Universe, ibid*; K. Olive, D. Seckel, in *Monopole '83*, ed. J. Stone, Plenum Press, NY 1984.
[47] M. S. Turner, *Phys. Rev.* **D28**, 1243 (1983).
[48] J. D. Barrow, M. S. Turner, *Nature* **292**, 35 (1981); J. D. Barrow, D. H. Sonoda, *Phys. Rep.* **139**, 1 (1986).
[49] D. Nanopoulos, K. Olive, M. Srednicki, *Phys. Lett.* **127B**, 30 (1983).
[50] R. J. Scherrer, M. S. Turner, *Phys. Rev.* **D31**, 681 (1985).
[51] T. Bunch, P. C. W. Davies, *Proc. R. Soc. London* **A360**, 117 (1978); also see, G. Gibbons, S. Hawking, *Phys. Rev.* **D15**, 2738 (1977).
[52] L. Abbott, M. Wise, *Nucl. Phys.* **B244**, 541 (1984).
[53] S. Hawking, *Phys. Lett.* **115B**, 295 (1982).
[54] A. A. Starobinskii, *Phys. Lett.* **117B**, 175 (1982).
[55] A. Guth, S.-Y. Pi, *Phys. Rev. Lett.* **49**, 1110 (1982).
[56] J. Bardeen, P. Steinhardt, M. S. Turner, *Phys. Rev.* **D28**, 679 (1983).
[57] R. Brandenberger, R. Kahn, W. Press, *Phys. Rev.* **D28**, 1809 (1983); W. Fischler, B. Ratra, L. Susskind, *Nucl. Phys.* **B259**, 730 (1985); R. Brandenberger, *Rev. Mod. Phys.* **57**, 1 (1985).
[58] E. M. Lifshitz, I. M. Khalatnikov, *Adv. Phys.* **12**, 185 (1963); E. M. Lifshitz, *Zh. Eksp. Teor. Fiz.* **16**, 587 (1946).
[59] W. Press, E. T. Vishniac, *Astrophys. J.* **239**, 1 (1980).
[60] J. M. Bardeen, *Phys. Rev.* **D22**, 1882 (1980).
[61] W. Press, in *Cosmology and Particles*, eds. J. Audouze et al., Editions Frontieres, Gif-sur-Yvette 1981, p. 137.
[62] J. A. Frieman, M. S. Turner, *Phys. Rev.* **D30**, 265 (1984); R. Brandenberger, R. Kahn, *Phys. Rev.* **D29**, 2172 (1984).
[63] E. Harrison, *Phys. Rev.* **D1**, 2726 (1970).
[64] Ya. B. Zel'dovich, *Mon. Not. R. Astron. Soc.* **160**, 1p (1972).
[65] P. J. Steinhardt, M. S. Turner, *Phys. Rev.* **D29**, 2162 (1984).
[66] M. S. Turner, E. T. Vishniac, in preparation (1985).
[67] J. Bond, A. Szalay, *Astrophys. J.* **276**, 443 (1983); P. J. E. Peebles, *Astrophys. J.* **263**, L1 (1982); J. Bond, A. Szalay, M. S. Turner, *Phys. Rev. Lett.* **48**, 1036 (1982).
[68] M. Axenides, R. Brandenberger, M. S. Turner, *Phys. Lett.* **120B**, 178 (1983); P. J. Steinhardt, M. S. Turner, *Phys. Lett.* **129B**, 51 (1983).
[69] D. Seckel, M. S. Turner, *Phys. Rev.* **D32**, 3178 (1985).
[70] A. D. Linde, *JETP Lett.* **40**, 1333 (1984); *Phys. Lett.* **158B**, 375 (1985).
[71] J. Preskill, M. Wise, F. Wilczek, *Phys. Lett.* **120B**, 127 (1983); L. Abbott, P. Sikivie, *Phys. Lett.* **120B**, 133 (1983); M. Dine, W. Fischler, *Phys. Lett.* **120B**, 137 (1983).
[72] M. S. Turner, F. Wilczek, A. Zee, *Phys. Lett.* **125B**, 35, 519(E) (1983); J. Ipser, P. Sikivie *Phys. Rev. Lett.* **50**, 925 (1983).
[73] M. S. Turner, *Phys. Rev.* **D33**, 889 (1986).
[74] S. Coleman, E. Weinberg, *Phys. Rev.* **D7**, 1888 (1973).
[75] A. D. Linde, *Phys. Lett.* **116B**, 335 (1982).
[76] A. Vilenkin, L. Ford, *Phys. Rev.* **D26**, 1231 (1982).
[77] S. Dimopoulos, S. Raby, *Nucl. Phys.* **B219**, 479 (1983).
[78] A. Albrecht, S. Dimopoulos, W. Fischler, E. Kolb, S. Raby, P. Steinhardt, *Nucl. Phys.* **B229**, 528 (1983).
[79] J. Ellis, D. Nanopoulos, K. Olive, K. Tamvakis, *Phys. Lett.* **118B**, 335 (1982); D. V. Nanopoulos, K. Olive, M. Srednicki, K. Tamvakis, *Phys. Lett.* **123B**, 41 (1983); J. Ellis, K. Enqvist, D. Nanopoulos, K. Olive, M. Srednicki, *Phys. Lett.* **152B**, 175 (1985); K. Enqvist, D. Nanopoulos, *Phys. Lett.* **142B**, 349 (1984); C. Kounnas, M. Quiros, *Phys. Lett.* **151B**, 189 (1985); for a review of the CERN SUSY/SUGR models see D. V. Nanopoulos, *Comments on Astrophysics* **X**, 219 (1986).

[80] B. Ovrut, P. J. Steinhardt, *Phys. Lett.* **133B**, 161 (1983); L. Jensen, K. Olive, *Nucl. Phys.* **B263**, 731 (1986).
[81] B. Ovrut, P. J. Steinhardt, *Phys. Rev. Lett.* **53**, 732 (1984); *Phys. Lett.* **147B**, 263 (1984).
[82] R. Holman, P. Ramond, G. G. Ross, *Phys. Lett.* **137B**, 343 (1984); C. D. Coughlan, R. Holman, P. Ramond, G. G. Ross, *Phys. Lett.* **140B**, 44 (1984); **158B**, 47 (1985).
[83] J. Breit, S. Gupta, A. Zaks, *Phys. Rev. Lett.* **51**, 1007 (1983).
[84] A. A. Starobinsky, *JETP Lett.* **30**, 682 (1979); V. Rubakov, M. Sazhin, A. Veryaskin, *Phys. Lett.* **115B**, 189 (1982); R. Fabbri, M. Pollock, *Phys. Lett.* **125B**, 445 (1983); L. Abbott, M. Wise, *Nucl. Phys.* **B244**, 541 (1984).
[85] J. Ellis, J. Kim, D. Nanopoulos, *Phys. Lett.* **145B**, 181 (1984); L. L. Krauss, *Nucl. Phys.* **B227**, 556 (1983); M. Yu. Khlopov, A. D. Linde, *Phys. Lett.* **138B**, 265 (1984).
[86] J. Polonyi, Budapest preprint KFKI 1977-93, unpublished (1977).
[87] C. Coughlan, W. Fischler, E. Kolb, S. Raby, G. G. Ross, *Phys. Lett.* **131B**, 54 (1983).
[88] S. Gupta, H. R. Quinn, *Phys. Rev.* **D29**, 2791 (1984).
[89] Q. Shafi, A. Vilenkin, *Phys. Rev. Lett.* **52**, 691 (1984).
[90] S.-Y. Pi, *Phys. Rev. Lett.* **52**, 1725 (1984).
[91] A. Guth, S.-Y. Pi, *Phys. Rev.* **D32**, 1899 (1985).
[92] A. D. Linde, *Phys. Lett.* **129B**, 177 (1983).
[93] V. A. Belinsky, L. P. Grischuk, I. M. Khalatnikov, Ya. B. Zel'dovich, *Phys. Lett.* **155B**, 232 (1985); T. Piran, R. M. Williams, *Phys. Lett.* **163B**, 331 (1985).
[94] S. Adler, *Rev. Mod. Phys.* **54**, 729 (1982); L. Smolin, *Nucl. Phys.* **B160**, 253 (1979); A. Zee, *Phys. Rev. Lett.* **42**, 417 (1979).
[95] F. Accetta, D. Zoller, M. S. Turner, *Phys. Rev.* **D31**, 3046 (1985).
[96] B. L. Spokoiny, *Phys. Lett.* **147B**, 39 (1984).
[97] Q. Shafi, C. Wetterich, *Phys. Lett.* **129B**, 387 (1983); **152B**, 51 (1985).
[98] A. A. Starobinskii, *Phys. Lett.* **91B**, 99 (1986); M. B. Mijic, M. S. Morris, W.-M. Suen, *Phys. Rev.* **D34**, 2934 (1986).
[99] A. Albrecht, R. Brandenberger, *Phys. Rev.* **D31**, 1225 (1985); **D32**, 1280 (1985).
[100] W. Fischler, B. Ratra, L. Susskind, *Nucl. Phys.* **B259**, 730 (1985).
[101] G. Semenoff, N. Weiss, *Phys. Rev.* **D31**, 689 (1985).
[102] M. Evans, J. McCarthy, *Phys. Rev.* **D31**, 1799 (1985).
[103] M. S. Turner, L. Widrow, *Phys. Rev. Lett.* **57**, 2237 (1986); L. Jensen, J. Stein-Schabes, *Phys. Rev.* **D34**, 931 (1986).
[104] R. M. Wald, *Phys. Rev.* **D28**, 2118 (1983).
[105] G. Steigman, M. S. Turner, *Phys. Lett.* **128B**, 295 (1983).
[106] L. Jensen, J. Stein-Schabes, *Phys. Rev.*, in press (1987).
[107] A. A. Starobinskii, *JETP Lett.* **37**, 66 (1983).
[108] S. K. Blau, E. I. Guendelman, A. H. Guth, *Phys. Rev.* **D**, in press (1987); K. Sato, M. Sasaki, H. Kodama, K. Maeda, *Phys. Lett.* **108B**, 103 (1982); *Prog. Theor. Phys.* **65**, 1443 (1981); **68**, 1979 (1982); **66**, 2287 (1981).
[109] N. Myhrvad, *Phys. Lett.* **132B**, 308 (1983); E. Mottola, *Phys. Rev.* **D31**, 754 (1985); **D33**, 2136; L. Parker, *Phys. Rev. Lett.* **50**, 1009 (1983); L. Ford, *Phys. Rev.* **D31**, 710 (1985); P. Anderson, *Phys. Rev.* **D32**, 1302 (1985); J. Traschen, C. T. Hill, *Phys. Rev.* **D33**, 3519 (1986).
[110] M. S. Turner, in *Dark Matter in the Universe*, eds. J. Kormendy, G. Knapp, Reidel, Dordrecht 1987, p. 445.
[111] M. S. Turner, G. Steigman, L. L. Krauss, *Phys. Rev. Lett.* **52**, 2090 (1984); D. A. Dicus, E. W. Kolb, V. Teplitz, *Phys. Rev. Lett.* **39**, 168 (1977); K. A. Olive, D. Seckel, E. T. Vishniac, *Astrophys. J.* **292**, 1 (1985).
[112] P. J. E. Peebles, *Astrophys. J.* **284**, 439 (1984); M. S. Turner et al., in Ref. [111].
[113] N. Kaiser, *Astrophys. J.* **273**, L17 (1983); J. M. Bardeen, J. Bond, N. Kaiser, A. Szalay, *Astrophys. J.* **304**, 15 (1986).

[114] A. Vilenkin, *Phys. Rev. Lett.* **53**, 1016 (1984).
[115] J. C. Charlton, M. S. Turner, *Astrophys. J.* **313**, 495 (1987).
[116] S. D. M. White, in *Inner Space/Outer Space*, eds. E. W. Kolb et al., University of Chicago Press, Chicago 1986.
[117] G. Blumenthal, S. Faber, J. Primack, M. Rees, *Nature* **311**, 517 (1984); M. Davis, G. Efstathiou, C. Frank, S. D. M. White, *Astrophys. J.* **292**, 371 (1985).
[118] D. Burstein et al., in *Galaxy Distances and Deviations from Universal Expansion*, eds. B. Madore, R. Tully, Reidel, Dordrecht 1987, p. 123.
[119] C. A. Collins et al., *Nature* **320**, 506 (1986).
[120] N. Vittorio, M. S. Turner, *Astrophys. J.* **316**, in press (1987).
[121] V. deLapparent, M. Geller, J. Huchra, *Astrophys. J.* **302**, L1 (1986).
[122] G. Efstathiou, J. R. Bond, *Mon. Not. R. Astron. Soc.* **218**, 103 (1986).
[123] P. Sikivie, *Phys. Rev. Lett.* **51**, 1415 (1983).
[124] M. Goodman, E. Witten, *Phys. Rev.* **D31**, 3059 (1985).
[125] E. Loh, E. Spillar, *Astrophys. J.* **307**, L1 (1986); *Phys. Rev. Lett.* **57**, 2865 (1986).
[126] H. Silk, M. S. Turner, *Phys. Rev.* **D35**, 419 (1986); M. S. Turner et al., *Astrophys. J.*, in press (1987).
[127] A. Linde, *Rep. Prog. Phys.* **47**, 925 (1984); *Comments on Astrophysics* **10**, 229 (1985); *Prog. Theor. Phys. (Suppl.)* **85**, 279 (1985).
[128] *The Inflationary Universe*, eds. L. Abbott and S.-Y. Pi, World Publishing, Singapore 1986.
[129] P. J. Steinhardt, *Comments Nucl. Part. Phys.* **12**, 273 (1984).
[130] R. Brandenberger, *Rev. Mod. Phys.* **57**, 1 (1985).
[131] A. Vilenkin, *Phys. Rep.* **121**, 263 (1985); A. Albrecht, N. Turok, *Phys. Rev. Lett.* **54**, 1868 (1985); N. Turok, *Phys. Rev. Lett.* **55**, 1801 (1985); R. Scherrer, *Astrophys. J.*, in press (1987); A. Melott, R. Scherrer, *Nature*, in press (1987).
[132] S. K. Blau, A. H. Guth, in *300 Years of Gravitation*, eds. S. W. Hawking and W. Israel, Cambridge University Press, Cambridge 1987.
[133] S. Bonometto, A. Masiero, *Riv. Nuovo Cimento* **9**, 3 (1986).