Rep. Prog. Phys., Vol 47, pp 925–986, 1984. Printed in Great Britain

The inflationary universe

A D Linde

I E Tamm Department of Theoretical Physics, P N Lebedev Physical Institute, Academy of Sciences of the USSR, Moscow 117924, USSR

Abstract

According to the inflationary universe scenario the universe in the very early stages of its evolution was exponentially expanding in the unstable vacuum-like state. At the end of the exponential expansion (inflation) the energy of the unstable vacuum (of a classical scalar field) transforms into the energy of hot dense matter, and the subsequent evolution of the universe is described by the usual hot universe theory.

Recently it was realised that the exponential expansion during the very early stages of evolution of the universe naturally occurs in a wide class of realistic theories of elementary particles. The inflationary universe scenario makes it possible to obtain a simple solution to many longstanding cosmological problems and leads to a crucial modification of the standard point of view of the large-scale structure of the universe.

This review was received in February 1984.

0034-4885/84/080925+62$09.00

Contents

1. Introduction

The discovery of unified gauge theories of all fundamental interactions during the last 10–15 years has opened up a new era in the development of elementary particle theory. At the end of the 1960s the Glashow–Weinberg–Salam unified theory of weak and electromagnetic interactions was suggested (Glashow 1961, Weinberg 1967, Salam 1968). In 1974 the idea of grand unified theories was suggested, which unified strong, weak and electromagnetic interactions (Georgi and Glashow 1974). Finally, in 1976 it was suggested that all fundamental interactions could be uniquely described in the context of a supergravity theory (Freedman *et al* 1976, Deser and Zumino 1976). For a detailed discussion of the present status of elementary particle theory one can consult many excellent books and reviews (Abers and Lee 1973, Taylor 1976, Fayet and Ferrara 1977, Langacker 1981, van Nieuwenhuizen 1981, Okun' 1982).

Grand unified theories are renormalisable, just as quantum electrodynamics, and some of the most interesting versions of these theories prove to be asymptotically free, which means, roughly speaking, that the strength of interaction between different particles decreases with the increases in their energies (Gross and Wilczek 1973, Politzer 1973). This important property makes possible a quantitative description of the elementary particle interactions at energies up to $E \sim M_p \sim 10^{19}$ GeV in the centre-of-mass system, i.e. up to the Planck energy at which quantum gravity effects become important.

One of the consequences of the successful development of elementary particle theory is considerable progress in the theory of superdense matter. Ten years ago matter with a density greater than the nuclear density $\rho \sim 10^{14}$–10^{15} g cm^{-3} was usually called superdense, and there were almost no ideas of how to describe matter with $\rho \gtrsim 10^{15}$ g cm^{-3}. After the discovery of asymptotically free theories it became possible to describe matter at temperatures T up to $T \sim M_p \sim 10^{19}$ GeV, which corresponds to the Planck density $\rho_p \sim M_p^4 \sim 10^{94}$ g cm^{-3}. Thus, after the development of asymptotically free theories it became possible to investigate properties of matter at densities which are greater than the nuclear density by 80 orders of magnitude!

The results of this investigation were rather unexpected. It was shown that with an increase (decrease) of temperature a sequence of phase transitions should occur, each of which leads to a qualitative modification of the properties of superdense matter (Kirzhnits 1972, Kirzhnits and Linde 1972, 1974, 1976, Weinberg 1974, Dolan and Jackiw 1974, Linde 1979).

These phase transitions should have taken place during the process of cooling of the expanding universe soon after the Big Bang. The cosmological consequences of these phase transitions may be so important that by a comparison of the predictions of the theory of phase transitions in gauge theories with the present cosmological observational data one can obtain strong constraints on the parameters of elementary particle theory (Linde 1977, 1980a) and even on the possible classes of theories (Zeldovich *et al* 1974, Zeldovich and Khlopov 1978, Preskill 1979). This fact now becomes very important, since in the near future an investigation of particle properties in the energy range $E \sim 10^{15}$ GeV, which is necessary for a thorough study of grand unified theories, can be performed neither with the help of cosmic-ray experiments, nor by constructing new accelerators. The only 'laboratory' wherein elementary particles of such energies once existed was our universe in the very early stages of its evolution. Therefore, at present many experts in elementary particle physics consider the universe as a unique physical laboratory for testing the new elementary particle theories, and many new theories have already been rejected after this 'cosmological ability' test (Zeldovich and Khlopov 1978, Preskill 1979, Parke and Pi 1981, Lazarides *et al* 1982, Sikivie 1982, Weinberg 1982a, b).

During the investigation of the cosmological consequences of gauge theories a number of new, unexpected and very interesting possibilities have been revealed, some of which would have seemed absolutely crazy only a few years ago. One of the best examples is the idea that the baryon asymmetry of the universe could appear due to non-equilibrium CP-violating processes with baryon non-conservation in the very early universe (Sakharov 1967). This idea was successfully realised in the context of grand unified theories (Yoshimura 1978, Ignatiev *et al* 1978, Dimopoulos and Susskind 1978, Ellis *et al* 1979, Toussaint *et al* 1979, Weinberg 1979). A detailed discussion of the baryosynthesis scenario is contained in a number of review articles (see, for example, Dolgov and Zeldovich (1981), Langacker (1981) and Barrow (1983)). The main aim of the present review is to discuss some other ideas which have been suggested over the last few years and are related to the so-called inflationary universe scenario (Guth 1981, Linde 1982a, b, c, d, 1983d, e, Albrecht and Steinhardt 1982). This scenario, as well as the closely related Starobinsky model (Starobinsky 1979, 1980), opens up the possibility of getting a simple solution to many different cosmological problems, some of which for a long time seemed almost metaphysical.

An important feature of this scenario is the assumption that in the very early universe there was a stage of evolution in which the universe was in an unstable vacuum-like state with a large energy density. According to the Einstein equations, the universe in such a state expands exponentially, $a(t) \sim \exp(Ht)$, where $a(t)$ is the scale factor of the universe and H is the Hubble 'constant'. (Actually the parameter $H \equiv \dot{a}/a$ slowly decreases over time, and it is now many orders of magnitude smaller than it was at the stage of exponential expansion.) Later the vacuum-like state decays, its energy transforms into heat, the universe becomes hot and the scale factor $a(t)$ of the hot universe grows more slowly, $a(t) \sim \sqrt{t}$.

The possibility of exponential expansion in a vacuum-like state in the very early stages of evolution of the universe was first suggested by Gliner (1965, 1970). This possibility was also discussed by many other authors (Sakharov 1965, Altshuller 1972, Gurevich 1975, Gliner and Dyminikova 1975). However, the origin of the vacuum-like state investigated by these authors remained obscure.

Later it became clear that the constant homogeneous classical scalar field φ, which necessarily appears in all unified gauge theories, looks just like the vacuum state, which in some cases may have a very large energy density (Linde 1974, Veltman 1974, 1975, Dreitlein 1974). It was also shown that this energy is temperature-dependent (Linde 1974) and that during the first-order phase transition with large supercooling this vacuum energy may dominate the energy of the universe (Kirzhnits and Linde 1976, Linde 1979). After the phase transition the energy of the supercooled vacuum state transforms into heat, which may lead to a considerable growth of the total entropy of the universe (Linde 1979).

The actual significance of all these facts was not quite clear until the remarkable paper by Guth (1981), who suggested using the exponential expansion of the universe in the supercooled vacuum state in order to solve some longstanding cosmological problems, such as the flatness problem, the horizon problem and the primordial monopole problem. The word 'inflation', suggested by Guth, was originally related to the exponential expansion of the universe in the supercooled state before the phase transition. Later the same word was used to denote any intermediate stage of quasi-exponential expansion, in which the scale factor of the universe $a(t)$ was given by

$$a(t) \sim \exp\left(\int H(t)\,\mathrm{d}t\right) \tag{1.1}$$

where the Hubble 'constant' $H = \dot{a}/a$ varies slowly enough, $\dot{H} \ll H^2$.

The main idea of the inflationary universe scenario, suggested by Guth (1981), was very clear and attractive and many scientists worked enthusiastically in order to implement this scenario in the context of some realistic theory of elementary particles. However, as was pointed out by Guth himself, the universe after the phase transition in his scenario becomes extremely inhomogeneous (Guth 1981). This problem seemed unsolvable for some time (Hawking *et al* 1982, Guth and Weinberg 1983). Fortunately, the specific difficulties, which precluded the resolution of this problem in the original inflationary universe scenario, were removed in 1982, when the new inflationary universe scenario was suggested (Albrecht and Steinhardt 1982, Linde 1982a, b, c, d). With the help of this new scenario it became possible to solve not only the flatness, horizon and primordial monopole problems, but also many other difficult problems in the usual hot universe theory. However, the new inflationary universe scenario was still not quite perfect and a consistent realisation of the inflationary universe scenario was suggested only very recently, after the development of the chaotic inflation scenario which is not based on the idea of supercooling (Linde 1983d, e, g, Goncharov and Linde 1984a, b).

Thus, the inflationary universe scenario at present differs considerably from the original version of this scenario suggested by Guth, and we do not know what other modifications will be made to this scenario in the future. Nevertheless, it now seems possible to summarise some preliminary results of the development of the inflationary universe scenario, which could be of general physical interest.

This paper is based partially on my talks at the Nuffield Workshop in Cambridge (Linde 1983a) and at the Shelter Island Conference (Linde 1984a), and also on my previous review article (Linde 1984b). On the other hand, this paper may be considered as the second part of my review article on phase transitions in gauge theories and cosmology, which was published earlier in *Reports on Progress in Physics* (Linde 1979).

To make the presentation at least partially self-contained we will start with the discussion of some properties of gauge theories (§ 2) and of phase transitions in these theories (§ 3). In § 4 we will remind the reader of the main features of the standard hot universe scenario. In § 5 we will discuss some longstanding problems of this scenario, as well as some new problems, related to the cosmological consequences of gauge theories. Section 6 is devoted to the first version of the inflationary universe scenario as suggested by Guth (1981). In § 7 we discuss the first version of the new inflationary universe scenario (Linde 1982a, Albrecht and Steinhardt 1982). In § 8 an improved version of this scenario is discussed (Linde 1982b, c, d, Starobinsky 1982) and an important problem concerning the density perturbations generated after inflation is considered (Hawking 1982, Starobinsky 1982, Guth and Pi 1982, Bardeen *et al* 1983). In § 9 we discuss the Starobinsky model, which is closely related to the new inflationary universe scenario (Starobinsky 1979, 1980, 1983b). In § 10 we discuss a version of the new inflationary universe scenario based on a theory of phase transitions in $N = 1$

supergravity (Ellis *et al* 1983a, Nanopoulos *et al* 1983a, b, Linde 1983b). Section 11 contains a discussion of the chaotic inflation scenario (Linde 1983c, d). In § 12 we consider a possible realisation of the chaotic inflation scenario in supergravity (Linde 1983f, Goncharov and Linde 1984a, b). Our presentation would be somewhat incomplete if we did not discuss some less elaborate but extremely interesting questions related to quantum cosmology. These questions are considered in § 13 and in the appendices. In the conclusions (§ 14) we discuss the present status and outlook of further developments of the inflationary universe scenario.

2. Spontaneous symmetry breaking in gauge theories

One of the main features of unified gauge theories is the spontaneous symmetry breaking between different interactions due to the appearance of some constant classical scalar fields. To illustrate the main idea of the mechanism of spontaneous symmetry breaking let us consider a simple model of a real scalar field φ with the Lagrangian

$$L=\tfrac{1}{2}(\partial_\mu\varphi)^2-\frac{M^2}{2}\varphi^2-\frac{\lambda}{4}\varphi^4 \tag{2.1}$$

which describes the scalar field φ with the mass M and the coupling constant λ. The potential energy of the field φ (effective potential) in this theory at the classical level (i.e. without quantum corrections) is given by

$$V(\varphi)=\frac{M^2}{2}\varphi^2+\frac{\lambda}{4}\varphi^4. \tag{2.2}$$

From equation (2.2) it is clear that at $M^2>0$ the most energetically advantageous state is the state $\varphi=0$ (figure 1(a)). Let us consider the theory (2.1) with a 'wrong' sign of M^2, $M^2=-\mu^2<0$. In this case the state $\varphi=0$ is unstable and the energetically favourable state, corresponding to a minimum of $V(\varphi)$, is $\varphi=\varphi_0=\pm\mu/\sqrt{\lambda}$ (figure 1(b)). Therefore, in such a theory a constant classical field $\varphi=\mu/\sqrt{\lambda}$ should appear, which means that the symmetry $\varphi\leftrightarrow-\varphi$ of the original theory becomes spontaneously broken.

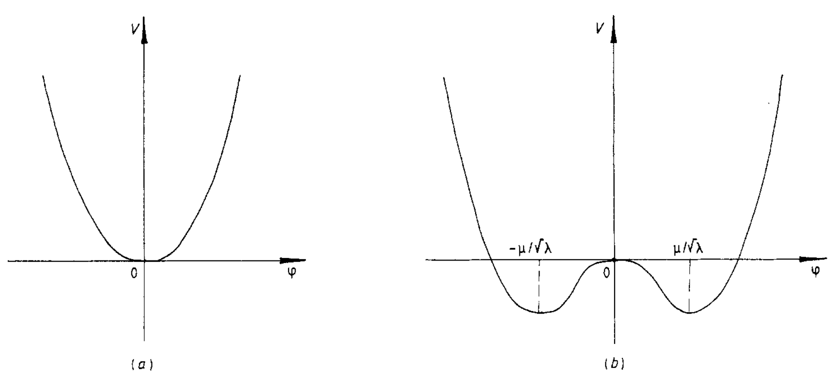

Figure 1. Effective potential $V(\varphi)$ in the theory (2.1). (a) $M^2>0$, (b) $M^2=-\mu^2<0$.

After the symmetry breaking the particle spectrum in the theory (2.1) changes. The effective mass squared of the scalar fluctuations near an extremum of $V(\varphi)$ is given simply by the value of the curvature of the effective potential near the extremum. For example, $m^2(\varphi=0)=\mathrm{d}^2V/\mathrm{d}\varphi^2|_{\varphi=0}=M^2=-\mu^2<0$. After symmetry breaking the sign of the mass squared becomes positive:

$$m^2(\varphi=\varphi_0)=\left.\frac{\mathrm{d}^2V}{\mathrm{d}\varphi^2}\right|_{\varphi=\varphi_0}=3\lambda\varphi_0^2-\mu^2=2\mu^2. \tag{2.3}$$

All other particles which interact with the field φ also change their masses after the symmetry breaking. Let us consider, for example, the Lagrangian of massless fermions ψ, interacting with the field φ:

$$L=\tfrac{1}{2}(\partial_\mu\varphi)^2+\frac{\mu^2}{2}\varphi^2-\tfrac{1}{4}\lambda\varphi^4+\bar{\psi}(\mathrm{i}\,\partial_\mu\gamma_\mu-h\varphi)\psi. \tag{2.4}$$

From equation (2.4) it follows that, after symmetry breaking, the fermions acquire a mass

$$m_\psi=|h\varphi|=h\mu\lambda^{-1/2}. \tag{2.5}$$

By a similar mechanism one can give a non-vanishing mass to the massless vector fields interacting with the field φ. As an important example we shall consider here the Higgs model (Higgs 1964a, b, 1966, Kibble 1967, Guralnik *et al* 1964, Englert and Brout 1964), which describes a massless vector field A_μ interacting with a complex scalar field $\chi=\frac{1}{2}(\chi_1+\mathrm{i}\chi_2)$ with a 'wrong' sign of the mass term $\mu^2\chi^*\chi$:

$$L=-\tfrac{1}{4}(\partial_\mu A_\nu-\partial_\nu A_\mu)^2+(\partial_\mu+\mathrm{i}eA_\mu)\chi^*(\partial_\mu-\mathrm{i}eA_\mu)\chi+\mu^2\chi^*\chi-\lambda(\chi^*\chi)^2. \tag{2.6}$$

The Lagrangian (2.6) is invariant under the U(1) group of gauge transformations:

$$\begin{aligned} A_\mu(x) &\to A_\mu(x)+\frac{1}{e}\,\partial_\mu\zeta(x) \\ \chi(x) &\to \chi(x)\exp[\mathrm{i}\zeta(x)]. \end{aligned} \tag{2.7}$$

This symmetry becomes spontaneously broken if the field χ acquires a classical part φ. This effect can most easily be described by the following change of variables:

$$\begin{aligned} \chi(x) &\to \frac{1}{\sqrt{2}}[\chi(x)+\varphi]\exp\left(\frac{\mathrm{i}\zeta(x)}{\varphi}\right) \\ A_\mu(x) &\to A_\mu(x)+\frac{1}{e\varphi}\,\partial_\mu\zeta(x) \end{aligned} \tag{2.8}$$

where φ is some constant classical field. After this change of variables the Lagrangian (2.6) is transformed into

$$\begin{aligned} L=&-\tfrac{1}{4}(\partial_\mu A_\nu-\partial_\nu A_\mu)^2+\frac{e^2}{2}(\chi+\varphi)^2A_\mu^2+\tfrac{1}{2}(\partial_\mu\chi)^2-\lambda\varphi^2\chi^2-\lambda\varphi\chi^3-\frac{\lambda}{4}\chi^4 \\ &-\frac{\lambda}{4}\varphi^4+\frac{\mu^2}{2}(\chi^2+\varphi^2)-(\lambda\varphi^2-\mu^2)\chi\varphi. \end{aligned} \tag{2.9}$$

Note that the auxiliary field $\zeta(x)$ is completely transformed away. The theory (2.9) describes the vector fields A_μ with the mass $m_A = e\varphi$, interacting with the real scalar field χ, which has a constant classical part φ with the effective potential (2.2) with $M^2 = -\mu^2$. At $\mu^2 > 0$, as before, there is a spontaneous symmetry breaking, the classical field φ becomes equal to $\varphi_0 = \mu/\sqrt{\lambda}$ and vector particles acquire the mass

$$m_A = e\varphi_0 = e\mu/\sqrt{\lambda}. \tag{2.10}$$

The scalar fields φ, which make vector mesons massive, are usually called Higgs bosons.

The main idea of the unification of gauge theories is that, before symmetry breaking, all vector mesons, which are responsible for different types of interactions, are massless. After the symmetry breaking some of the vector mesons, interacting with the classical Higgs field φ, become massive and the corresponding interactions become short-range. For example, the Glashow–Weinberg–Salam theory before symmetry breaking is invariant under the $SU(2)\times U(1)$ group of gauge transformations and describes long-range electroweak interactions mediated by four different massless vector bosons. After the appearance of the classical Higgs field $H \sim 250$ GeV some of these vector bosons ($W^\pm_\mu$ and Z^0_μ) acquire mass $m_W \sim m_Z \sim eH \sim 10^2$ GeV and the corresponding interactions become short-range (weak interactions), whereas one more field (the electromagnetic field A_μ) remains massless (Weinberg 1967, Salam 1968).

In grand unified theories several different types of Higgs scalar fields are necessary. For example, in the minimal SU(5) theory one of these fields Φ is represented by a traceless matrix 5×5. The SU(5) effective potential with respect to the field Φ at the classical level looks as follows:

$$V(\Phi) = \frac{m^2}{2}\operatorname{Tr}\Phi^2 + \frac{a}{4}(\operatorname{Tr}\Phi^2)^2 + \frac{b}{2}\operatorname{Tr}\Phi^4. \tag{2.11}$$

There are two different symmetry breaking patterns in this theory. The first is the symmetry breaking $SU(5) \to SU(4)\times U(1)$ due to generation of the field

$$\Phi = \frac{1}{\sqrt{20}}\varphi\begin{pmatrix} 1 & & & & 0 \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ 0 & & & & -4 \end{pmatrix}. \tag{2.12}$$

This possibility is undesirable, and one should arrange things in such a way as to avoid such a symmetry breaking by an appropriate choice of effective potential $V(\Phi)$, or by some other means (see § 12).

Another (desirable) type of symmetry breaking is $SU(5) \to SU(3)\times SU(2)\times U(1)$, which occurs due to the appearance of the field

$$\Phi = \left(\tfrac{2}{15}\right)^{1/2}\varphi\begin{pmatrix} 1 & & & & 0 \\ & 1 & & & \\ & & 1 & & \\ & & & -\frac{3}{2} & \\ 0 & & & & -\frac{3}{2} \end{pmatrix}. \tag{2.13}$$

The value of the field φ in (2.13) is extremely large, $\varphi \sim 10^{15}$ GeV. Before symmetry breaking all vector particles in this theory are massless, there is no difference between weak, strong and electromagnetic interactions, and leptons can easily be transformed into quarks and vice versa. After the symmetry breaking X and Y vector mesons, responsible for the interactions transforming quarks into leptons, acquire a very large mass, $m_X = m_Y = \sqrt{\frac{5}{3}} g\varphi/2 \sim 10^{15}$ GeV, where $g^2 \sim 0.3$ is the SU(5) gauge coupling constant. The corresponding interactions become strongly suppressed, which is the reason why the proton at present is almost stable. After the appearance of the field (2.13) the original SU(5) symmetry breaks down to $SU(3) \times SU(2) \times U(1)$, which implies symmetry breaking between strong interactions (with the group of symmetry SU(3)) and electroweak interactions ($SU(2) \times U(1)$). Then there appears another classical field $H \sim 10^2$ GeV, which breaks the $SU(2) \times U(1)$ symmetry down to the $U(1)_{EM}$ symmetry of electromagnetic interactions, just as in the Glashow–Weinberg–Salam theory (Georgi and Glashow 1974).

A detailed discussion of different symmetry breaking patterns in grand unified theories is contained in the papers by Gell-Mann *et al* (1978), Langacker (1981) and Slansky (1981).

From equation (2.2) it could be inferred that symmetry breaking in gauge theories is possible only for $M^2 = -\mu^2 = d^2V/d\varphi^2|_{\varphi=0} < 0$. However, this is not quite correct. For example, it can be shown that in the one-loop approximation the effective potential in the Higgs model at $e^2 \gg \lambda$ looks as follows:

$$V(\varphi) = \frac{M^2\varphi^2}{2} + \frac{\lambda\varphi^4}{4} + \frac{3e^4}{32\pi^2}\varphi^4 \ln\frac{\varphi}{\mu_0} \tag{2.14}$$

where μ_0 is some normalisation mass. From (2.14) it follows that even at $M^2 > 0$ the effective potential may have its absolute minimum somewhere at $\varphi = \varphi_0 \neq 0$ (Coleman and Weinberg 1973, Linde 1976, Weinberg 1976). In particular, at $M^2 = 0$ the effective potential (2.14) can be represented in the following form:

$$V(\varphi) = \frac{3e^4\varphi^4}{32\pi^2}\left(\ln\frac{\varphi}{\varphi_0} - \frac{1}{4}\right). \tag{2.15}$$

This is the Coleman–Weinberg theory (Coleman and Weinberg 1973), which at $\varphi = 0$ looks just like a massless scalar electrodynamics. However, from (2.15) it follows that the state $\varphi = 0$ is unstable and the absolute minimum of $V(\varphi)$ is displaced at $\varphi = \varphi_0 \neq 0$. The SU(5) version of the Coleman–Weinberg effective potential with respect to the symmetry breaking $SU(5) \to SU(3) \times SU(2) \times U(1)$ (2.13) is

$$V(\varphi) = \frac{25g^4}{128\pi^2}\left(\varphi^4 \ln\frac{\varphi}{\varphi_0} - \frac{\varphi^4}{4} + \frac{\varphi_0^4}{4}\right). \tag{2.16}$$

The last term is added in order to make the vacuum energy zero at present, $V(\varphi_0) = 0$. This constraint follows from cosmological observational data, which implies that at present $|V(\varphi_0)| \lesssim 10^{-29}\,\mathrm{g\,cm^{-3}}$, whereas the typical value of $V(0)$ (2.16) is of the order of $10^{75}\,\mathrm{g\,cm^{-3}}$. By using the relation $m_X = \sqrt{\frac{5}{3}} g\varphi_0/2$ one may represent equation (2.16) in the following form:

$$V(\varphi) = \frac{25}{128\pi^2}\varphi^4\left(\ln\frac{\varphi}{\varphi_0} - \frac{1}{4}\right) + \frac{9}{32\pi^2}m_X^4. \tag{2.17}$$

The cosmological implications of this theory will be thoroughly discussed in §§ 7 and 8.

3. Phase transitions in gauge theories

The investigation of superdense matter described by gauge theories with spontaneous symmetry breaking was initiated by Kirzhnits (1972) (see also Kirzhnits and Linde 1972). He has predicted that at a sufficiently large temperature T the classical scalar field φ, which leads to the symmetry breaking, must disappear. A detailed theory of the corresponding phase transition has been developed by Weinberg (1974), Dolan and Jackiw (1974) and Kirzhnits and Linde (1974, 1976). A thorough discussion of the phase transitions in gauge theories is contained in Linde (1979), and therefore we will just remind readers here of some of the main features of the symmetry behaviour in gauge theories at a finite temperature.

The main point is that at a finite temperature an equilibrium value of $\varphi(T)$ corresponds not to the minimum of the potential energy $V(\varphi)$ but to the minimum of the free energy $F(\varphi, T) \equiv V(\varphi, T)$, which is equal to $V(\varphi)$ at $T=0$. It is known that the temperature-dependent contribution of ultrarelativistic scalar particles with mass $m \ll T$ to the value of $F(\varphi, T)$ is given by (Landau and Lifshitz 1976)

$$\Delta F = \Delta V(\varphi, T) = -\frac{\pi^2}{90}T^4 + \frac{m^2}{24}T^2\left[1 + \mathrm{O}\left(\frac{m}{T}\right)\right]. \tag{3.1}$$

Now let us take into account that in the theory (2.1)

$$m^2(\varphi) = \frac{\mathrm{d}^2 V}{\mathrm{d}\varphi^2} = 3\lambda\varphi^2 - \mu^2.$$

Therefore, at $T \gg m$

$$V(\varphi, T) = -\frac{\mu^2}{2}\varphi^2 - \frac{\lambda}{4}\varphi^4 + \frac{\lambda T^2}{8}\varphi^2 - \frac{\pi^2 T^4}{90} - \frac{\mu^2 T^2}{24} \tag{3.2}$$

(see figure 2).

From equation (3.2) it follows that with an increase of temperature the value of the field $\varphi(T)$ in the minimum of $V(\varphi, T)$ decreases, and at all temperatures exceeding the critical temperature

$$T_c = 2\varphi_0 = \frac{2\mu}{\sqrt{\lambda}} \tag{3.3}$$

the only minimum of $V(\varphi, T)$ is displaced at $\varphi = 0$, which means that at $T > T_c$ there is no symmetry breaking in the theory (2.1). From equation (3.2) it follows that the value of $\varphi(T)$ with the growth of temperature up to $T = T_c$ decreases continuously, which corresponds to the second-order phase transition.

Note that in the case $\lambda \ll 1$ the value of T_c (3.3) is much greater than m at $\varphi \lesssim \varphi_0$, which justifies our use of the high-temperature decomposition in (3.1). However, in many realistic theories (and in practically all grand unified theories (Linde 1980b, 1981a, Daniel 1981)) the phase transition occurs at $T \sim m$, or even at $T \ll m$. In such a case the effective potential $V(\varphi, T)$ may have more than one local minimum (Kirzhnits and Linde 1976, Linde 1979). A typical example is shown in figure 3. In some temperature intervals $T_c^1 < T_c < T_c^2$ the effective potential $V(\varphi, T)$ has two minima. The phase transition starts at the critical temperature T_c, at which the values of $V(\varphi, T)$ in these minima become equal to each other. Such a phase transition is the first-order one. It proceeds, like the boiling of water, by the formation and expansion of bubbles

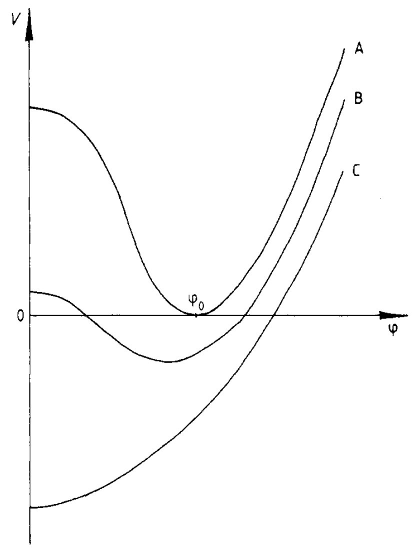

Figure 2. Effective potential $V(\varphi, T)$ in the theory (2.1) at $M^2 = -\mu^2 < 0$. A, $T = 0$; B, $0 < T < T_c$; C, $T > T_c$.

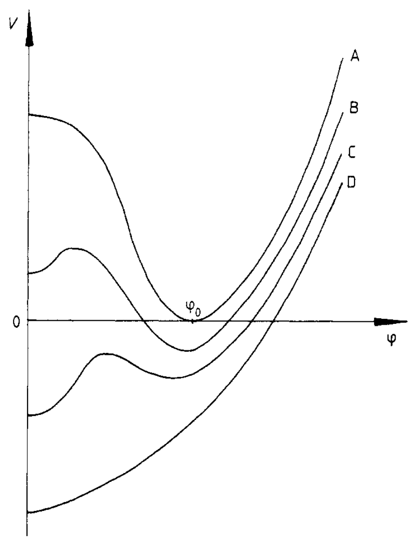

Figure 3. Behaviour of $V(\varphi, T)$ corresponding to the first-order phase transition. At $T_c^1 < T < T_c^2$ the effective potential has two different minima, which have the same depth at $T = T_c$. A, $T = 0$; B, $T_c^1 < T < T_c$; C, $T_c < T < T_c^2$; D, $T > T_c^2$.

filled with matter in a new, energetically favourable phase. An investigation of this process shows that in many theories the probability of bubble formation is very small (Voloshin *et al* 1974, Coleman 1977, Linde 1977, 1981b, 1983h). Therefore the phase transition with an increase of temperature in such theories proceeds from a strongly superheated asymmetric state $\varphi \neq 0$, whereas the phase transition with a decrease of temperature in these theories proceeds from a strongly supercooled symmetric state $\varphi = 0$. Such a phase transition in grand unified theories, which we have called the Grand Bang (Linde 1981a), in some cases looks like a large explosion all over the universe. As we shall see, such processes may lead to many interesting cosmological consequences.

4. The standard hot universe scenario

According to the hot universe theory, the universe has been expanding and gradually cooling from a state with infinite temperature and density (Weinberg 1972, Zeldovich and Novikov 1975). The universe is assumed to be homogeneous and isotropic (in accordance with the observational data) and is described by the Friedmann–Robertson–Walker metric

$$ds^2 = dt^2 - a^2(t)\left(\frac{d\tau^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta \, d\varphi^2)\right) \tag{4.1}$$

where $k = +1$, -1 or 0 for a closed, open or flat universe, respectively, $a(t)$ is the 'radius' of the universe or, to be more precise, its scale factor. The evolution of $a(t)$ is governed by the Einstein equations

$$\ddot{a} = -\frac{4\pi}{3} G(\rho + 3p)a \tag{4.2}$$

$$H^2 + \frac{k}{a^2} \equiv \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi}{3} G\rho. \tag{4.3}$$

Here ρ is the energy density, p denotes the pressure, G is the gravitational constant, $G = M_p^{-2}$ (where $M_p \sim 10^{19}$ GeV is the Planck mass) and $H = \dot{a}/a$ is the Hubble 'constant' (which, generally speaking, is time-dependent). Conservation of energy (which can be deduced from (4.2) and (4.3)) leads to the equation

$$\frac{\dot{a}}{a} = -\frac{\dot{\rho}}{3(\rho + p)}. \tag{4.4}$$

In the standard scenario it is usually assumed also that the expansion is adiabatic, in which case

$$\frac{d}{dt}(sa^3) = 0 \tag{4.5}$$

where s is the entropy density.

In the asymptotically free theories in the lowest approximation one may neglect particle interactions in the superdense matter, which means that superdense matter looks like an (almost) ideal gas of ultrarelativistic particles. Therefore the values of

ρ, p and s are given by (Landau and Lifshitz 1976)

$$\rho = 3p = \frac{\pi^2}{30} N(T) T^4 \tag{4.6}$$

$$s = \frac{2\pi^2}{45} N(T) T^3. \tag{4.7}$$

Here $N(T)$ is the effective number of particle species:

$$N(T) = N_B(T) + \tfrac{7}{8} N_F(T) \tag{4.8}$$

where $N_B(N_F)$ is the effective number of boson (fermion) degrees of freedom (for example, $N_B = 1$ for a real scalar field, $N_B = 3$ for a massive vector field, etc) with masses $m \ll T$.

It can be shown (see § 5) that, in the very early stages of evolution of the universe in the standard scenario, the universe was very flat and one may neglect the term k/a^2 in equation (4.3). In that case, from (4.3) and (4.4) it follows that

$$a(t) \sim \sqrt{t} \tag{4.9}$$

and the age of the universe is given by

$$t = \frac{1}{4\pi}\left(\frac{45}{\pi N}\right)^{1/2} \frac{M_p}{T^2}. \tag{4.10}$$

Rigorously speaking, these results are valid only at $T \leqslant T_p \sim M_p/\sqrt{N} \sim 10^{18}$ GeV, and at the density $\rho \leqslant \rho_p \sim M_p^4/N \sim 10^{92}$ g cm^{-3}, since at $T \geqslant M_p/\sqrt{N}$, $\rho \gtrsim M_p^4/N$ quantum corrections to the Einstein equations become considerable (Zeldovich and Novikov 1975) (see also Linde 1983b). Moreover, the thermodynamic equilibrium in the expanding universe is established only at $T \leqslant 10^{16}$ GeV (Dolgov and Zeldovich 1981), though some effectively equilibrium state can be formed due to the quantum gravity effects at $T \geqslant T_p$ (Weinberg 1979).

In this section we will consider phase transitions in grand unified theories, which occur at $T \leqslant 10^{15}$ GeV, when the state of thermodynamic equilibrium has already been established. At $T \gtrsim 10^{15}$ GeV symmetry in grand unified theories was unbroken. At $t_1 \sim 10^{-35}$ s after the Big Bang, when the temperature drops down to $T \sim T_{c_1} \sim 10^{14}$–$10^{15}$ GeV, the first phase transition with symmetry breaking occurs in grand unified theories. For example, in the SU(5) theory this may be a transition SU(5) $\to$ SU(3) $\times$ SU(2) $\times$ U(1). After this phase transition strong interactions become separated from electroweak interactions and leptons become separated from quarks. This phase transition is first order (Linde 1980b, 1981a, Daniel 1981). After the phase transition the superheavy X, Y bosons and the superheavy Higgs bosons decay, which leads to the baryon asymmetry generation (Sakharov 1967, Kuzmin 1970, Yoshimura 1978, Ignatiev *et al* 1978, Dimopoulos and Susskind 1978, Ellis *et al* 1979, Toussaint *et al* 1979, Weinberg 1979).

At $t_2 \sim 10^{-10}$ s, when the temperature drops down to $T_{c_2} \sim 10^2$ GeV, the phase transition separating weak and electromagnetic interactions occurs (SU(3) $\times$ SU(2) $\times$ U(1) $\to$ SU(3) $\times$ U(1)). This phase transition is a second-order or a weakly first-order one. At the temperature $T_{c_3} \sim 10^2$ MeV the phase transition (or two different phase transitions) occurs with the formation of baryons from quarks and with breaking of chiral invariance in the theory of strong interactions.

The subsequent evolution of the universe is described in many textbooks on cosmology (see, for example, Zeldovich and Novikov 1975). The most important stages of evolution of the universe are shown in figure 4 (see also Zee 1980, Kirzhnits and Linde 1982). The main part of this review will be devoted to the discussion of events which took place about 10^{10} yr ago, in the period $t \lesssim 10^{-30}$ s after the creation of the universe.

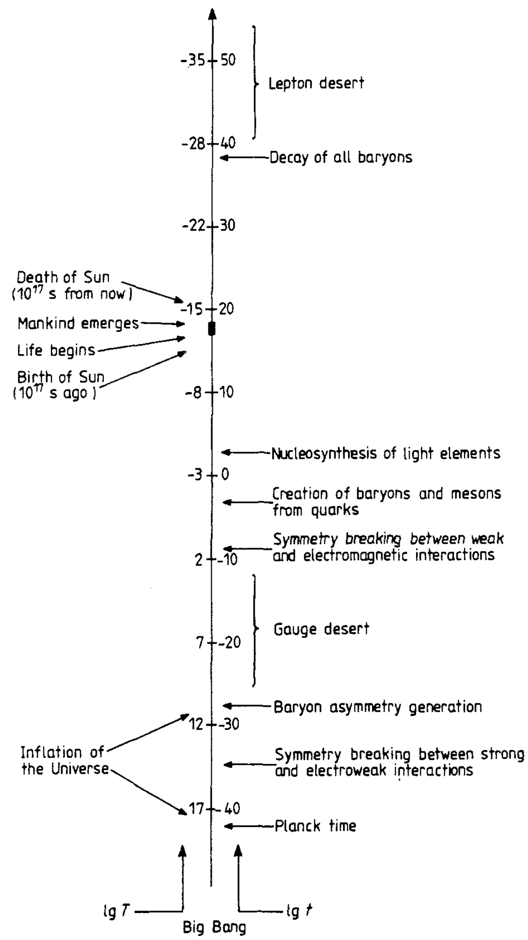

Figure 4. Some important stages of evolution of the universe (with an account taken of the inflationary universe scenario). The age of the universe is given in lg t(s) and the temperature of the universe is given in lg T(GeV). 1 GeV corresponds approximately to 10^{13} K. The classical description of evolution of the universe becomes possible at $t \gg t_p \sim 10^{-43}$ s. *In typical grand unified theories nothing interesting occurs at* 10^2 GeV $\ll T \ll 10^{14}$ GeV (the gauge desert). However, some oases in the gauge desert may exist, e.g. at $T \sim 10^{10}$–10^{11} GeV (see § 12). At $t \sim 10^{38}$–10^{40} s all baryons decay and the universe becomes filled with a dilute gas of leptons and photons (the lepton desert). In the middle of the way from the gauge desert to the lepton desert there exists a small oasis of life, in which we now live. Inflation of the universe presumably occurred at $t \lesssim 10^{-25}$ s, and only after inflation was the baryon asymmetry of the universe generated.

5. Problems of the standard scenario

Despite the great phenomenological success of the standard hot universe scenario, this scenario was still somewhat incomplete. Here we will list some of the longstanding problems of this scenario.

5.1. *The singularity problem*

From equations (4.6) and (4.10) it follows that the scale factor of the universe $a(t)$ vanishes at $t \to 0$, whereas the energy density at $t \to 0$ becomes infinitely large (Hawking and Ellis 1973, Zeldovich and Novikov 1975). One may wonder, therefore, what was *before* the singularity? If our universe did not exist at $t < 0$, then how could it originate from 'nothing'? The singularity problem is certainly one of the most puzzling problems of contemporary science.

5.2. *The flatness problem*

From equation (4.3) it follows that at $t \to 0$

$$\frac{|\rho - \rho_c|}{\rho_c} = \dot{a}^{-2} \tag{5.1}$$

where ρ_c is the energy density corresponding to $k = 0$ (flat universe) and ρ is the energy density in the closed or open universe ($k = \pm 1$ in (4.3)) with the same H.

The present energy density ρ is not known exactly, $0.03 \leqslant \rho/\rho_c \leqslant 2$. Therefore the value of $|\rho - \rho_c|/\rho_c$ may now be rather large. However, according to (4.9), $\dot{a}^{-2} \sim t$ in the very early stages of evolution of the universe, which means that in the very early universe the value of $|\rho - \rho_c|/\rho_c$ was extremely small. One can show that, from the fact that the present 'radius' of the universe $a(t)$ exceeds 10^{28} cm, it follows that the value of $|\rho - \rho_c|/\rho_c$ in the standard scenario should be smaller than $10^{-59} M_p^2/T^2$ (Guth 1981, Linde 1984b). Therefore near the Planck time $t_p \sim M_p^{-1}$, when $T \sim T_p \sim M_p$,

$$|\rho - \rho_c|/\rho_c \leqslant 10^{-59}. \tag{5.2}$$

This means, for example, that if the density of the universe at the Planck time was slightly greater than ρ_c, say $\rho \geqslant \rho_c(1 + 10^{-55})$, then the universe would be closed and it would have collapsed millions of years ago. If, on the other hand, $\rho \leqslant \rho_c(1 - 10^{-55})$ near the Planck time, then the universe would be open and the present energy density of the universe would be negligibly small. In the standard hot universe scenario it is absolutely unclear why our universe was created flat, or almost flat, with such fantastic accuracy. It can be shown that the question of why our universe is so flat is equivalent to the question of why the total entropy S of the observable part of the universe is so large, $S \geqslant (aT_\gamma)^3 \geqslant 10^{87}$. Here $a \sim 10^{28}$ cm is the radius of the observable part of the universe and $T_\gamma \sim 2.7$ K (the temperature of the microwave background radiation). The equivalence of the flatness and total entropy problems can most easily be understood in the case of a closed universe, since the maximal 'radius' of a closed universe grows with an increase in S: $a_{max} \sim M_p^{-1} S^{2/3}$ (Landau and Lifshitz 1973).

5.3. *The homogeneity and isotropy problems*

In § 4 it was assumed that the universe was initially absolutely homogeneous and isotropic. Meanwhile, even at present the universe is not totally homogeneous and isotropic, at least at a sufficiently small length scale. This means that there are no reasons for believing that the universe was homogeneous and isotropic *ab initio.* A more natural assumption is that the universe initially was in some chaotic state, and that the initial conditions in parts of the universe sufficiently far removed from each other were practically uncorrelated (see, for example, Misner (1969a, b) and Rees (1972)). However, as was shown by Collins and Hawking (1973), the class of all initial conditions for which the universe at large time behaves as a homogeneous and isotropic Friedmann universe (4.1) is a class of measure zero. Therefore, it is very difficult to understand why our universe is so homogeneous and isotropic. For a detailed discussion of all the subtleties of this problem see also Zeldovich and Novikov (1975).

5.4. *The horizon problem*

The isotropy problem was somewhat moderated after it was understood that the presence of matter and the effects connected with elementary particle creation in the expanding universe can make the universe *locally* isotropic (Zeldovich and Novikov 1975). However, such effects presumably cannot make the universe *globally* isotropic, since in the standard scenario the properties of space–time in the causally unconnected regions of the universe, displaced a distance exceeding the size of the particle horizon $l \sim ct$ from each other, cannot be correlated and cannot influence each other in any way. However, from the investigation of the isotropy of the microwave background radiation it follows that, at $t \sim 10^5$ yr, the universe was homogeneous and isotropic at a scale much greater than ct. The corresponding problem is called the horizon problem or the causality problem (Rindler 1956, Misner 1969a, b).

5.5. *The galaxy formation problem*

It is well-known that the universe is not exactly homogeneous. Such inhomogeneities as stars, galaxies and clusters of galaxies are too important to be overlooked in our discussion of the structure of the universe. To explain their formation one should assume that in the very early universe there existed small density perturbations $\delta\rho$ with an almost scale-independent spectrum $\delta\rho/\rho \sim 10^{-4}$ (Zeldovich 1972). However, it was not quite clear what was the source of these inhomogeneities with such a specific spectrum.

5.6. *The baryon asymmetry problem*

The essence of this problem is to understand why in the observable part of the universe the density of baryons is many orders greater than the density of antibaryons and why, on the other hand, the density of baryons is much less than the density of photons, $n_B/n_\gamma \sim 10^{-9}$.

These problems, given above, have for a long time seemed to be almost metaphysical. For example, the essential part of the cosmological singularity problem can be formulated as the following question: has anything existed at $t < 0$ when our universe did not exist? As for all other questions, it was argued that the universe is unique, and there is no point in thinking about any different initial conditions in the universe. Another possible answer was based on the so-called anthropic principle (Dicke 1961, Collins and Hawking 1973, Carr and Rees 1979, Rozental 1980, 1984). It was argued that in a matter–antimatter symmetric, anisotropic and inhomogeneous universe there would be no observers who could ask such questions. This argument is very witty but is not quite convincing enough for two reasons. First of all, such an answer implies that there may exist many universes and we live in just one of them, which is sufficiently suitable for the existence of intelligent life. However, it was not quite clear in what sense one could speak about many universes if our universe is unique. One of the possible answers to this objection is connected with the quantum gravity effects and with the many-world interpretation of quantum mechanics (Everett 1957, DeWitt 1967) (see also § 13). Another possibility is connected with the inflationary universe scenario and will be discussed in § 12. One more difficulty of the anthropic principle is that it cannot explain why the universe is almost exactly homogeneous and isotropic, why the spectrum of inhomogeneities in our universe is almost scale-independent, why the numerical value of n_B/n_γ is $O(10^{-9})$, etc. Therefore the anthropic principle *by itself* (without the help of the inflationary universe scenario) cannot answer all the questions discussed above.

Besides the problems outlined in §§ 5.1–5.6 there also exist some other problems, which are related to the cosmological consequences of the new theories of elementary particles.

5.7. The domain wall problem

As was shown in § 3, at a sufficiently high temperature $T > T_c$ the symmetry in the theory (2.1) was restored, $\varphi(T) = 0$. With a decrease of temperature in the expanding universe the symmetry breaking phase transition should occur. However, in sufficiently far removed (causally unconnected) domains of the universe this phase transition may proceed into two different states: into the state $\varphi = +\mu/\sqrt{\lambda}$ or into the state $\varphi = -\mu/\sqrt{\lambda}$. The domains of the field $\varphi = +\mu/\sqrt{\lambda}$ are separated from the domains of the field $\varphi = -\mu/\sqrt{\lambda}$ by thin walls, inside which the field φ varies from $\mu/\sqrt{\lambda}$ to $-\mu/\sqrt{\lambda}$. The surface energy density of such domain walls is so large that, if at least one such wall existed at present in the observable part of the universe (which seems to be unavoidable in the standard scenario), the observable part of the universe would be largely anisotropic.

This result implies that most of the theories with spontaneous breaking of a discrete symmetry of the type given in (2.1) (which is symmetric with respect to the change $\varphi \leftrightarrow -\varphi$) contradict cosmological data (Zeldovich *et al* 1974). Among such theories is the simplest version of the SU(5) theory with the effective potential (2.11) (Parke and Pi 1981, Lazarides *et al* 1982), many theories with spontaneously broken *CP* invariance, including the Weinberg model of *CP* violation (Weinberg 1976), most theories of axions (Sikivie 1982), etc. Many of these theories are very attractive in all other respects and it would be very desirable to save at least some of them.

5.8. The primordial monopole problem

In those theories with other types of symmetry breaking some other structures can be formed. For example, in the Higgs model with broken U(1) symmetry, and in some other theories, strings of the Abrikosov vortex tube type can be formed (Kibble 1976, 1980). Formation of such strings in grand unified theories may be very important for the theory of galaxy formation (Zeldovich 1980, Vilenkin 1981). However, the most important effect here is formation of the 't Hooft–Polyakov monopoles ('t Hooft 1974, Polyakov 1974), which should be copiously produced during the phase transitions in grand unified theories at $T = T_{c_1} \sim 10^{14}$–10^{15} GeV (Kibble 1976, 1980). As was shown by Zeldovich and Khlopov (1978) and Preskill (1979), annihilation of such monopoles is rather ineffective, and at present the density of monopoles in the universe would be of the same order as the density of protons. This would lead to catastrophic cosmological consequences, since the mass of each monopole is approximately 10^{16} times greater than the proton mass, and therefore the density of matter in our universe would be approximately 10^{15} times greater than the critical density $\rho_c \sim 10^{-29}$ g cm^{-3}. With such a density the universe would have collapsed a long time ago. The primordial monopole problem is one of the most difficult problems for new theories of elementary particles, since this problem arises in practically all grand unified theories.

5.9. The primordial gravitino problem

One of the most interesting directions in the development of unified theories is connected with supersymmetry, which is a symmetry between bosons and fermions (Gol'fand and Likhtman 1971, Volkov and Akulov 1972, Wess and Zumino 1974). Supersymmetric theories have many beautiful properties (Fayet and Ferrara 1977, van Nieuwenhuizen 1981). In particular, in the context of $N = 1$ supergravity coupled to matter it is possible to solve one of the most difficult problems of the unified theories, namely the gauge hierarchy problem (Ibanez 1982, Barbieri *et al* 1982, Nath *et al* 1982). The essence of the problem is to explain why several different mass scales exist, $M_p \gg m_X \sim 10^{15}$ GeV, $m_X \gg m_W \sim 10^2$ GeV (Gildener 1976). One of the important features of the proposed solution of the gauge hierarchy problem is the existence of the gravitino (spin-$\frac{3}{2}$ superpartner of the graviton) with mass $m_{3/2} \sim m_W \sim 10^2$ GeV. However, according to Weinberg (1982a), such particles, being present in the very early stages of evolution of the universe, decay after the process of nucleosynthesis, which would lead to some undesirable cosmological consequences. This would not lead to a contradiction with the cosmological data only if the relative abundance of gravitinos in the very early universe was extremely small, $n_{3/2}/n_\gamma \leqslant 10^{-10}$–$10^{-12}$ (Khlopov and Linde 1984), whereas in the standard scenario one would expect $n_{3/2}/n_\gamma \sim 10^{-3}$. The question then arises whether it is possible to avoid the undesirable consequences of gravitino decay in the very early universe, or if one should abandon the above-mentioned possibility to solve the gauge hierarchy problem.

5.10. The problem of proper symmetry breaking

In the unified theories the effective potential $V(\varphi)$ often has more than one local minimum. For example, in the theory (2.1) with $M^2 = -\mu^2 < 0$ there are two minima, at $\varphi = +\mu/\sqrt{\lambda}$ and at $\varphi = -\mu/\sqrt{\lambda}$. In the minimal supersymmetric SU(5) theory (Fradkin 1980, Dimopoulos and Georgi 1981, Sakai 1982) there are three different minima of $V(\Phi)$ with $V(\Phi_i) = 0$. The first minimum corresponds to the SU(5)-symmetric vacuum state, the second one corresponds to the symmetry breaking SU(4) × U(1) and the third one corresponds to the desirable symmetry breaking SU(3) × SU(2) × U(1) (see figure 5). The number of different degenerate minima of the effective potential

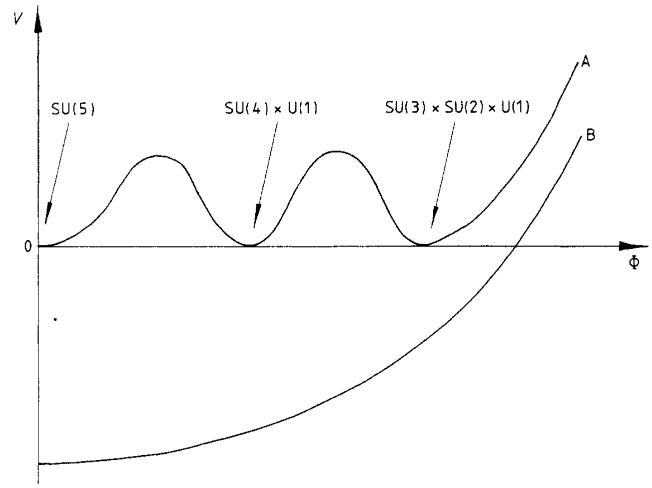

Figure 5. Effective potential $V(\Phi, T)$ in the minimal supersymmetric SU(5) model. A, $T=0$; B, $T>T_c$.

becomes even greater if one also considers other Higgs fields, such as the light Higgs bosons H (Dragon 1982, Frampton and Kephart 1982, Buccella *et al* 1982).

The question now arises of why we are now in the vacuum state with the symmetry breaking $SU(3)\times U(1)$ (which originated from $SU(3)\times SU(2)\times U(1)$) and not in some other undesirable state, corresponding to some other minimum of $V(\Phi, H)$. This problem becomes even more complicated if one takes into account that, in the early universe, the SU(5) symmetry was restored due to high-temperature effects (see figure 5) and in the standard scenario there are no reasons for the universe to jump from the SU(5) minimum to the $SU(3)\times SU(2)\times U(1)$ minimum of $V(\Phi)$ (Nanopoulos and Tamvakis 1982, Srednicki 1982a). There were some attempts to solve this problem based on the idea of SU(5) confinement (Srednicki 1982b, Nanopoulos *et al* 1982). In our opinion, however, the problem remained unsolved (Linde 1983f).

5.11. The problem of space–time dimensionality

The question of why our space–time is four-dimensional would have seemed rather meaningless and scholastic only a few years ago. However, at present the theories of the type given by Kaluza (1921) and Klein (1926) have become more and more popular. In such theories it is assumed that our space has dimension $d>4$, but $d-4$ dimensions are spontaneously compactified, i.e. the curvature radius in $d-4$ dimensions is extremely small, of the order of M_p^{-1}. That is why we cannot move in $d-4$ directions and our space–time is apparently four-dimensional (Cremmer and Scherk 1976, Witten 1981). Such theories became especially interesting in relation to the extended supergravities $N=4$ and $N=8$, which can most easily be formulated in $d=10$ and $d=11$ spaces (van Nieuwenhuizen 1981). This leads us to the question of why just $d-4$ dimensions have been compactified, and not $d-3$ or $d-5$.

5.12. The vacuum energy (cosmological constant) problem

From the cosmological observations it follows that the vacuum energy density $V(\varphi_0)$ in the universe at present cannot be much greater than the critical density $\rho_c \sim 10^{-29}$ g cm^{-3}:

$$|\rho_{\text{vac}}| = |V(\varphi_0)| \lesssim 10^{-29} \text{ g cm}^{-3}. \tag{5.3}$$

This value of $V(\varphi)$ was achieved after a sequence of symmetry breaking phase transitions (Linde 1974). In the minimal SU(5) theory after the phase transition $\text{SU}(5) \to \text{SU}(3) \times \text{SU}(2) \times \text{U}(1)$ the vacuum energy $V(\varphi)$ decreases approximately by 10^{80} g cm^{-3}. After the phase transition $\text{SU}(3) \times \text{SU}(2) \times \text{U}(1) \to \text{SU}(3) \times \text{U}(1)$ the vacuum energy decreases by $\sim 10^{25}$ g cm^{-3}. Finally, after the phase transition with the baryon formation from quarks the vacuum density decreases by $\sim 10^{14}$ g cm^{-3} and, surprisingly enough, becomes zero with an accuracy of $\pm 10^{-29}$ g cm^{-3}! Such a precise cancellation of the vacuum energy is unbelievable (Veltman 1974, 1975) unless there exist some unknown reasons for this cancellation. The vacuum energy problem at present is regarded as one of the most difficult problems of unified theories with spontaneous symmetry breaking. Since the vacuum energy $V(\varphi)$ multiplied by $8\pi G$ enters as a cosmological term into the Einstein equations, the vacuum energy problem is sometimes also called the cosmological constant problem (Linde 1974, Veltman 1974, 1975, Dreitlein 1974).

The first and the last of the problems mentioned above have not yet been finally solved. However, some very interesting attempts to solve them are directly related to the inflationary universe scenario and/or to the Starobinsky model (see § 13 and the appendices). A solution to the baryon asymmetry problem was suggested by Sakharov (1967) many years before the proposal of the inflationary universe scenario, but the inflationary universe scenario makes this solution more effective (Dolgov and Linde 1982). What makes the inflationary universe scenario so attractive is that the remaining problems can be either completely or partially solved in the context of this scenario, which we are now going to discuss.

6. The first version of the inflationary universe scenario

The inflationary universe scenario, in its present form, differs considerably from the first version of this scenario suggested by Guth (1981). Nevertheless, we consider it expedient to begin our presentation with the discussion of the main ideas of the original scenario.

As we have mentioned in § 3, the phase transitions from the state $\varphi = 0$ to the symmetry breaking state $\varphi = \varphi_0$ in some theories proceed with a large supercooling. The energy density of the ultrarelativistic particles $\sim T^4$ in a strongly supercooled state $\varphi = 0$ becomes smaller than the vacuum energy density $V(0)$. This means that in the limit of an extremely strong supercooling the energy density of matter ρ in the expanding and cooling universe becomes constant, $\rho = V(0)$. In such a case, according to (4.3), the universe at large t becomes exponentially expanding (Kolb and Wolfram 1980)

$$a(t) \sim a_0 \exp(Ht) \tag{6.1}$$

where the Hubble constant at that time is given by

$$H = \left(\frac{8\pi V(0)}{3M_p^2}\right)^{1/2}. \tag{6.2}$$

After the phase transition the vacuum energy $V(0)$ transforms into heat and the universe becomes very hot again (Linde 1979).

If the transformation of the vacuum energy into heat occurs rapidly enough (during the time $\Delta t \lesssim H^{-1}$), the universe after the phase transition reheats up to the temperature $T_R \sim V_{(0)}^{1/4}$. Note that the reheating temperature does not depend on the duration of the exponential expansion before the phase transition. This observation is the starting point of the inflationary universe scenario suggested by Guth (1981), which makes it possible to solve simultaneously the horizon and flatness problems discussed in the previous section. Indeed, the only quantity which depends on the duration of the stage of exponential expansion (inflation) is the scale factor $a(t)$, which becomes exponentially large after the expansion. Therefore, after the expansion one can neglect the term k/a^2 in equation (4.3), which means that the universe becomes very flat. An alternative way of understanding this effect is to consider the total entropy of the universe. After the phase transition the total entropy $S \gtrsim a^3 T_R^3 \sim a^3 V_{(0)}^{3/4}$ becomes exponentially large. Let us assume, for example, that the exponential expansion starts in a closed universe with a radius $a_0 = c_1 M_P^{-1}$ and with the vacuum energy $V(0) = c_2 M_P^4$, where c_1 and c_2 are some constants. In realistic theories $1 \lesssim c_1 \lesssim 10^{10}$ and $10^{-20} \lesssim c_2 \lesssim 1$. Actual values of c_i will be not very important for us (see below).

The total entropy of a closed universe after exponential expansion during the time Δt becomes of the order of

$$S \sim a_0^3 \exp(3H\Delta t) T_R^3 \sim c_1^3 c_2^{3/4} \exp(3H\Delta t) \tag{6.3}$$

which means that the entropy S becomes greater than 10^{87} (see § 5) if

$$\Delta t \gtrsim H^{-1}(67 - \ln c_1 c_2^{1/4}). \tag{6.4}$$

Typically $|\ln c_1 c_2^{1/4}| \lesssim 10$. This means that for the solution of the flatness (entropy) problem in this scenario the universe should be exponentially expanding during the time

$$\Delta t \gtrsim 70 H^{-1} = 70\left(\frac{3M_P^2}{8\pi V(0)}\right)^{1/2}. \tag{6.5}$$

If the value of Δt considerably exceeds 70 H^{-1} (which is the case in all realistic versions of the inflationary universe scenario, see §§ 10–12), then the universe after expansion becomes extremely flat, $\Omega = \rho/\rho_c \approx 1$. This is one of the most important predictions of the inflationary universe scenario, which can be experimentally tested (Guth 1981, 1983).

One should note, however, that this solution of the flatness problem is complete only if the universe is open. If the universe is closed, a typical lifetime of the universe would be $\Delta t \sim M_P^{-1}$, and the energy density of such a universe $\sim T^4 \sim M_P^4$ never becomes as small as $V(0)$. Such a universe recollapses before the exponential expansion starts. One could argue that this just means that our universe should be open (with $\Omega \approx 1$) (see Linde 1983a). Fortunately, however, the flatness problem can be completely solved in the chaotic inflation scenario (Linde 1983d, e) even if the universe is closed, since in the chaotic inflation scenario the exponential expansion may start even at $V(\varphi) \gtrsim M_P^4$ (see § 11).

The condition $S \sim (aT)^3 \gtrsim 10^{87}$ implies that, at present, when the temperature of photons $T_\gamma \sim 3$ K, the 'radius' of the universe a is greater than the size of the observable part of the universe $l \sim 10^{28}$ cm. This means that after the exponential expansion during the time $\Delta t'$, which slightly exceeds Δt (6.5) (by $\sim H^{-1} \ln c_1$), and after the subsequent expansion according to the hot universe theory any domain of a size $\Delta l \sim M_p^{-1}$ acquires a size exceeding the size of the observable part of the universe.

Since we consider here the processes in the post-Planckian epoch ($\rho \lesssim M_p^4$, $T \lesssim M_p$, $t \gtrsim M_p^{-1}$), the size of the domains $\Delta l \sim M_p^{-1}$ at the beginning of the exponential expansion is smaller than the size of the causally connected domains of the universe $\Delta l_{\text{horizon}} \sim t \gtrsim M_p^{-1}$. Thus, in this scenario the observable part of the universe arises due to the exponential expansion of one extremely small causally connected domain. This solves the horizon problem.

The same scenario could also help us to solve the primordial monopole problem. Indeed, as we have mentioned in § 5, the primordial monopoles are effectively created only during the phase transition with symmetry breaking, and only in those points at which several different bubbles of the field φ collide (Kibble 1976, 1980). If supercooling is very large, then the typical size of the bubbles produced during the phase transition at the moment when different bubbles collide with each other is also very large and the resulting density of the monopolies is very small (Guth 1981).

Unfortunately, however, as was noted by Guth himself, this scenario leads to some unacceptable cosmological consequences. According to this scenario, the field φ inside the bubbles of the new phase, formed during the phase transition, immediately grows up to its equilibrium value φ_0, which corresponds to the minimum of $V(\varphi)$, and the reheating of the universe occurs only after the bubble wall collisions. This would lead to an extremely large inhomogeneity and anisotropy of the universe after the phase transition. Nevertheless, the main idea of the inflationary universe scenario was so attractive that many scientists extensively studied this scenario during the year after the appearance of the paper by Guth (1981). The main results of this investigation have been summarised in the papers by Hawking *et al* (1982) and Guth and Weinberg (1983), in which it was concluded that the difficulties in the original scenario were insurmountable. Fortunately, at the same time, a new version of the inflationary universe scenario was suggested (Linde 1982a) (see also Albrecht and Steinhardt 1982), which was free of the main difficulties of the original scenario, and which provided the possibility of solving not only the horizon, flatness and primordial monopole problems but also many other problems mentioned in the previous section.

7. The new inflationary universe scenario: a simplified version

The first version of the new inflationary universe scenario was based on the theory of high-temperature phase transitions in the SU(5) Coleman–Weinberg theory (2.16) and (2.17). This theory is very complicated. Therefore we would first like to give a somewhat simplified description of the phase transition in SU(5) Coleman–Weinberg theory (Linde 1982a), which will make it possible to outline the main ideas of the new inflationary universe scenario. A more detailed discussion of this scenario is contained in the next section.

First of all, let us study the high-temperature behaviour of the effective potential in the SU(5) Coleman–Weinberg theory with respect to the symmetry breaking SU(5) $\to$ SU(3) $\times$ SU(2) $\times$ U(1) (2.13).

As was shown in § 3, at very high temperatures there should be no symmetry breaking in gauge theories. In particular, the effective potential $V(\varphi, T)$ in the SU(5)

Coleman–Weinberg theory at $T \gg m_X$, m_φ looks as follows:

$$V(\varphi, T) = \tfrac{5}{8}g^2T^2\varphi^2 + \frac{25g^4\varphi^4}{128\pi^2}\left(\ln\frac{\varphi}{\varphi_0} - \frac{1}{4}\right) + \frac{9m_X^4}{32\pi^2} + cT^4 \tag{7.1}$$

where c is some constant, $c = O(10)$. At $T \gg \varphi_0$ the only minimum of $V(\varphi, T)$ is the minimum at $\varphi = 0$, which means that at $T \gg \varphi_0$ the symmetry is restored. At $T \ll \varphi_0 \sim 10^{15}$ GeV all high-temperature corrections to $V(\varphi)$ at $\varphi \sim \varphi_0$ vanish, since in this case the masses m_X and m_φ are much greater than T. However, the masses of all particles in the Coleman–Weinberg theory vanish near $\varphi = 0$. Therefore in the vicinity of the point $\varphi = 0$ equation (7.1) holds even at $T \ll 10^{15}$ GeV. This means that at any $T \neq 0$ the point $\varphi = 0$ remains a local minimum of $V(\varphi, T)$:

$$m^2(T) = \left.\frac{d^2V(\varphi, T)}{d\varphi^2}\right|_{\varphi=0} = \tfrac{5}{4}g^2T^2. \tag{7.2}$$

To be more precise, one should also take into account the scalar field contribution to $V(\varphi, T)$, which changes the curvature of $V(\varphi, T)$ near $\varphi = 0$:

$$\left.\frac{d^2V}{d\varphi^2}\right|_{\varphi=0} = \left(\frac{5}{4}g^2 + \frac{65a + 47b}{30}\right)T^2. \tag{7.3}$$

At $T \sim \varphi_0$ the last term in (7.3) can be omitted, since in the Coleman–Weinberg theory $a, b \ll g^2$ at $T, \varphi \approx \varphi_0$. However, at $T, \varphi \ll \varphi_0$ the effective constants $a(T, \varphi)$ and $b(T, \varphi)$ become large and negative, which may change the sign of $d^2V/d\varphi^2|_{\varphi=0}$ at a sufficiently small temperature (Linde 1982c). We will forget about this effect for a while in order to simplify our presentation but will return to the discussion of this point in the next section.

The phase transition from the local minimum of the effective potential at $\varphi = 0$ to the global minimum at $\varphi \sim \varphi_0$ proceeds by the formation and subsequent expansion of the bubbles of the field φ. According to Sher (1982) and Billoire and Tamvakis (1982), the probability of this process becomes significant only at a very small temperature, $T \lesssim 10^6$ GeV (this statement is not quite correct—see the next section—but we shall assume for a moment that it is true). In that case, from our investigation of the bubble formation at a finite temperature (Linde 1981a, b, 1983h) it follows that the bubble of the field φ at the moment of its formation has a size $O(T_c^{-1})$, and the field inside the bubble is much smaller than φ_0 (see figure 6):

$$\varphi \lesssim 3\varphi_1 \approx \frac{12\pi T_c}{g(5\ln m_X/T_c)^{1/2}} \ll \varphi_0 \tag{7.4}$$

where the point φ_1 is defined by the condition $V(0, T_c) = V(\varphi_1, T_c)$. This means that the mass squared of the field φ inside the bubble is

$$|m^2| = \left|\frac{d^2V}{d\varphi^2}\right| \lesssim 75g^2T_c^2 \sim 25T_c^2. \tag{7.5}$$

It is clear that the field φ inside the bubble will increase to its equilibrium value $\varphi \sim \varphi_0$ during the time interval Δt, which is of the order of or greater than $|m^{-1}| \sim 0.2T_c^{-1}$. (Actually Δt is even greater than $|m^{-1}|$ (Linde 1982b).) During the main part of this interval the field φ remains much smaller than φ_0. This means that during some time $\Delta t \gtrsim 0.2T_c^{-1}$ the vacuum energy $V(\varphi)$ remains almost equal to $V(0)$, and the part of the universe inside the bubble expands exponentially just as it expanded before the

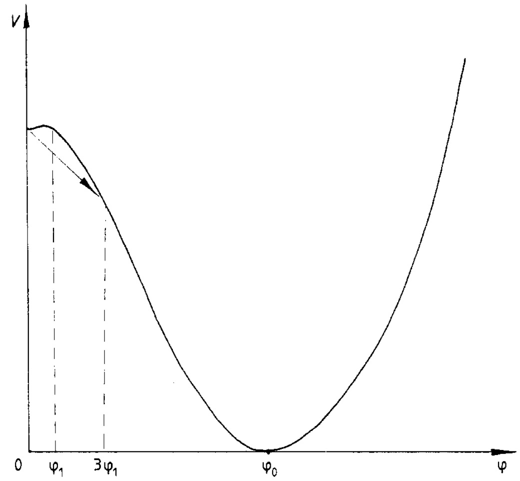

Figure 6. Effective potential in the Coleman–Weinberg theory at $T \ll \varphi_0$. The arrow indicates the direction of tunnelling with the bubble formation. The field inside the bubble $\varphi \leqslant 3\varphi_1$, where $V(\varphi_1, T) = V(0, T)$.

bubble creation. This is the main difference between this scenario and the scenario suggested by Guth (1981), in which it was assumed that the exponential expansion finished immediately after bubble formation.

The value of the Hubble 'constant' in this theory at $\varphi \ll \varphi_0$, $m_X \sim 5 \times 10^{14}$ GeV is given by

$$H = \left(\frac{8\pi}{3M_p^2} V(0)\right)^{1/2} = \frac{m_X^2}{2M_p}\left(\frac{3}{\pi}\right)^{1/2} \approx 10^{10}\,\mathrm{GeV}. \tag{7.6}$$

During the period of time $\Delta t \gtrsim 0.2 T_c^{-1}$ the universe grows $\exp(H\Delta t)$ times, where

$$\exp(H\Delta t) \gtrsim \exp(0.2H/T_c) \sim \exp(2000) \sim 10^{800}. \tag{7.7}$$

A typical size of the bubble at the moment of its creation is $O(T_c^{-1}) \sim 10^{-20}$ cm. After the period of exponential expansion this bubble acquires a size of the order of $10^{-20} \exp(H\tau)$ cm $\sim 10^{800}$ cm, which is very much greater than the size of the observable part of the universe at present, $l \sim 10^{28}$ cm. Therefore the whole observable part of the universe is contained *inside one bubble*, so we see no inhomogeneities caused by bubble wall collisions.

When the field φ grows sufficiently large, the rate of its growth increases, and finally the field φ becomes convergently oscillating near its equilibrium state $\varphi = \varphi_0$ with frequency equal to the Higgs meson mass m at $\varphi \sim \varphi_0$, $m \approx \frac{1}{4}m_X \approx 10^{14}$ GeV. Note that this frequency is much larger than $H \sim 10^{10}$ GeV. Therefore the whole process of symmetry breaking can be approximately divided into two parts.

(a) The field φ grows very slowly, whereas the size of the bubble increases exponentially, and all the observable part of the universe becomes filled with the almost homogeneous field $\varphi \ll \varphi_0$.

(b) This almost homogeneous field rapidly grows. During this time the universe expansion rate can be neglected compared with the rate of growth and subsequent oscillations of the field φ. The oscillating field φ creates Higgs bosons and X, Y bosons, which rapidly decay and reheat the universe up to the temperature $T_R \sim V_{(0)}^{1/4} \sim \frac{1}{6} m_X \sim 10^{14}$ GeV (Linde 1982a).

Thus the mechanism of reheating the universe in the new inflationary universe scenario also differs considerably from the mechanism of reheating in the first version of the inflationary universe scenario.

The process of baryon asymmetry generation in the inflationary universe is also different from that in the standard baryosynthesis scenario (Dolgov and Linde 1982). In the first papers discussing this question it was assumed that the baryon asymmetry in the inflationary universe was generated either directly by the oscillating field φ (Linde 1982a, Hawking and Moss 1982, Abbott *et al* 1982) or by the decay of the Higgs bosons *after* the reheating of the universe (Albrecht *et al* 1982). However, the process of baryon production actually proves to be much more complicated and, in the most interesting cases, proceeds due to the decay of Higgs bosons and/or X, Y bosons *before* the final reheating of the universe (Dolgov and Linde 1982). It appears that the baryon asymmetry in the inflationary universe, under certain constraints on the masses of superheavy bosons, can be one or two orders of magnitude greater than that in the standard scenario. It is important also that any initial baryon asymmetry of the universe vanishes after the inflation. Therefore the final baryon asymmetry in the inflationary universe does not depend on the initial baryon asymmetry of the universe in contrast to what occurs in the standard baryon synthesis scenario based on BL conserving grand unified theories.

Just as in the Guth scenario, the exponential expansion of more than exp (70) times (7.7) provides us with a solution of the horizon and flatness problems (see the previous section). However, the new scenario opens up a possibility of solving the homogeneity and isotropy problems as well. As we have noted in § 5, the processes connected with particle creation in the very early universe and with the existence of dense hot matter can make our universe locally isotropic at a scale exceeding the Planck scale $l_p \sim M_p^{-1} \sim 10^{-33}$ cm. Then the exponential expansion extends this isotropy to all the observable part of the universe (Linde 1982a). Moreover, the remaining small anisotropy inside the bubble decreases rapidly during the exponential expansion (see, for example, Gibbons and Hawking 1977, Boucher and Gibbons 1983, Starobinsky 1983a, Steigman and Turner 1983, Wald 1983, Boucher 1983). The solution of the homogeneity problem is very similar to that of the isotropy problem. Density inhomogeneities inside the bubble immediately after its formation are negligibly small compared with $V(0)$, i.e. the space inside the bubble is almost homogeneous. Then the exponential expansion extends this homogeneity to the whole observable part of the universe (Linde 1982a).

As we have noted in § 5, monopoles and domain walls are created only after the phase transition in the regions in which bubbles with different types of the Higgs field φ collide (Kibble 1976, 1980). In the new inflationary universe scenario the typical size of a bubble is much greater than the size of the observable part of the universe. Therefore no monopoles or domain walls are created in the observable part of the universe in this scenario. Inflation of the universe dilutes the density of all objects which existed before the exponential expansion stage. In particular, it dilutes the density of gravitinos, which were created near the time of the Big Bang (Ellis *et al* 1982). However, gravitinos can be created again if the reheating temperature T_R is large enough (S Weinberg 1983 private communication, Nanopoulos *et al* 1983a, Khlopov and Linde 1984). An example of a theory in which the gravitino problem can actually be solved is discussed in § 12.

Thus we see that the new inflationary universe scenario may actually provide us with a simple solution of many problems considered in § 5 (see also subsequent sections). The main idea of the new scenario is very simple. This scenario can be realised if the field φ in the first stages of the symmetry breaking process changes very slowly (this is necessary in order to have a large inflation), and if the frequency of oscillations of the field φ near $\varphi = \varphi_0$ is sufficiently large (this is necessary for efficient reheating of the universe after inflation). The same idea was also used in the improved version of this scenario to be discussed now, as well as in all other versions of the inflationary universe scenario considered so far.

8. Improvement of the new inflationary universe scenario

The version of the new inflationary universe scenario discussed in the previous section (as well as a similar version suggested later by Albrecht and Steinhardt (1982)) was oversimplified. As was noted in the first paper in which this scenario was suggested (Linde 1982a), to have a more precise theory of the Coleman–Weinberg phase transition one should take into account the effects connected with the non-vanishing curvature and rapid expansion of the universe, which become important at $T \lesssim H$. One should also carefully study the behaviour of the effective coupling constants $a(T)$ and $b(T)$ in (7.3) at small temperatures.

The last question was studied in our paper (Linde 1982c), in which it was shown that the effective coupling constants $a(T)$ and $b(T)$ at small T become large and negative, and therefore the phase transition from $\varphi = 0$ to $\varphi = \varphi_0$ may occur due to the change of sign of $\mathrm{d}^2 V/\mathrm{d}\varphi^2|_{\varphi=0}$ without tunnelling. The critical temperature T_c, at which the curvature of $V(\varphi, T)$ near $\varphi = 0$ changes its sign, is model-dependent. In the theories with $T_c \gg H$ inflation does not occur. Therefore, in what follows we will consider only those theories in which $T_c \ll H$. However, in such theories all high-temperature effects are irrelevant for the theory of the phase transition (Linde 1982b, c). Indeed, the time interval which is necessary for the phase transition to start due to high-temperature effects exceeds $T_c^{-1} \gg H^{-1}$, but during this interval the temperature falls to $T_c \exp(-H/T_c)$, i.e. it practically vanishes. This means that the theory of the inflationary phase transition in the Coleman–Weinberg theory is determined not by the high-temperature effects but by the effects connected with the non-vanishing curvature and the exponential expansion of the universe.

Before discussing these effects we would like to make a comment, which is necessary in order to avoid some terminological misunderstandings which often appear in the literature. The metric of the exponentially expanding Friedmann universe (4.1) at large t can be written as

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \exp(2Ht)(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2). \tag{8.1}$$

This is the metric of a flat de Sitter space. As was shown by Gibbons and Hawking (1977), from the point of view of a comoving observer the temperature in a de Sitter space is equal to the Hawking temperature $T_H = H/2\pi$ and does not decrease in the course of the exponential expansion. This effect is connected with the fact that any comoving observer in the exponentially expanding universe cannot have any information about anything removed at a distance greater than the event horizon radius H^{-1} from the observer. However, if one tries to describe the whole universe (bearing in mind that after the phase transition the event horizon disappears) one can use a formalism in which there is no lower bound on the temperature, $T \sim T_0 \exp(-Ht)$ (Birrell and Davies 1982, Linde 1982d, 1983a). Just this formalism has been used in most of the papers in which gravitational effects in the inflationary universe have been studied (Linde 1982b, d, Vilenkin and Ford 1982, Starobinsky 1982, Vilenkin 1983b), though an alternative approach to these effects is also possible (Hawking and Moss 1982, 1983). All the physical results, of course, do not depend on the formalism used for their derivation.

Investigation of gravitational effects in the inflationary universe scenario is rather complicated, which caused many different misunderstandings. For example, the theory of tunnelling with bubble formation in de Sitter space has been developed by many authors. However, about half of the papers discussing this question contain errors, whereas the correct investigation of bubble formation in de Sitter space (Coleman and De Luccia 1980, Hawking and Moss 1982, Parke 1983, Guth and Weinberg 1983) is directly applicable only to the case of bubble formation in the eternally existing de Sitter space but not to the hot universe with the de Sitter stage of exponential expansion (Linde 1983a, Goncharov and Linde 1984c). In the first paper discussing tunnelling in the new inflationary universe scenario with an account taken of gravitational effects (Hawking and Moss 1982) it was claimed that the tunnelling occurs simultaneously in the whole universe and proceeds from the minimum of $V(\varphi)$ at $\varphi = 0$ to the nearby maximum of $V(\varphi)$ at $\varphi = \tilde{\varphi}$ with the probability

$$P \sim \left(\frac{\mathrm{d}^2 V}{\mathrm{d}\varphi^2}\bigg|_{\varphi=0}\right)^2 \exp\left[-\frac{3M_{\mathrm{P}}^4}{8}\left(\frac{1}{V(0)} - \frac{1}{V(\tilde{\varphi})}\right)\right] \tag{8.2}$$

per unit four-volume. The simultaneity of the phase transition in the whole universe was considered as a cause of the homogeneity of the universe after inflation. The same statement was made later by Mottola and Lapedes (1983) and Abbott and Burges (1983). In Linde (1983a) it was noted, however, that the probability of a globally homogeneous tunnelling in the inflationary universe is strongly suppressed. Later Hawking and Moss noted that their result (8.2) should be valid for the probability of tunnelling which is not absolutely homogeneous but looks homogeneous in the domains of the universe of size $l \gtrsim H^{-1}$ (Hawking and Moss 1983). Very recently this statement has been proven by Starobinsky (1984a, c) for the case $\mathrm{d}^2 V/\mathrm{d}\varphi^2|_{\varphi=0} \ll H^2$ by a method totally different from those used by the previous authors.

We will not discuss this complicated question any more (for a detailed discussion of tunnelling in the inflationary universe see, for example, Goncharov and Linde (1984c)), since in the course of the investigation of this problem it was understood that the tunnelling with bubble formation is not a necessary ingredient of the new inflationary universe scenario, and a simpler version of this scenario can be suggested which is not based on the theory of bubble formation in the exponentially expanding universe (Linde 1982b, d, Starobinsky 1982).

To explain this modification of the new inflationary universe scenario let us consider the scalar field theory with an effective potential which, at small φ, looks as follows:

$$V(\varphi) = \frac{m^2\varphi^2}{2} - \frac{\lambda}{4}\varphi^4 + V(0) \tag{8.3}$$

without the terms $\sim \xi R \varphi^2$ where R is the curvature scalar. This potential at small $|m^2|$ and $\lambda > 0$ imitates the effective potential in the Coleman–Weinberg theory.

If the field φ initially was zero then, according to the classical equations of motion, it should remain zero forever even if $m^2 < 0$. It is clear, however, that spontaneous symmetry breaking should occur in this theory due to quantum fluctuations. To describe this process let us consider the behaviour of quantum fluctuations of the field φ, i.e. the vacuum average $\langle\varphi^2\rangle$ in the exponentially expanding universe. We will assume here that $|m^2| \ll H^2$, which is necessary for the realisation of the new inflationary universe scenario. In this case, from the results of Bunch and Davies (1978) it follows that

$$\langle\varphi^2\rangle = \frac{3H^4}{8\pi^2 m^2}. \tag{8.4}$$

From equation (8.4) it follows that $\langle\varphi^2\rangle \to \infty$ for $m^2 \to 0$. The physical reason for such a pathological behaviour of $\langle\varphi^2\rangle$ at $m^2 \to 0$ in the exponentially expanding universe is connected with the anomalously large density of long-wave fluctuations in de Sitter space. The leading contribution to $\langle\varphi^2\rangle$ comes from fluctuations with momenta $|k| \ll H$ and is given by

$$\langle\varphi^2\rangle \approx \frac{H^{2-4m^2/3H^2}}{2\pi^2} \int_0^H \frac{k^2\,\mathrm{d}k}{[\Delta m^2(T) + k^2]^{\frac{3}{2}-2m^2/3H^2}} \tag{8.5}$$

(Vilenkin and Ford 1982, Starobinsky 1982, Linde 1982d), where $\Delta m^2(T)$ is the temperature-dependent contribution to m^2, $\Delta m^2(T) \sim \mathrm{O}(g^2T^2)$. At $T = 0$ this yields the Bunch–Davies result (8.4). At $T \gg H$ this anomalous contribution disappears and $\langle\varphi^2\rangle = T^2/12$ just as in Minkowski space (Linde 1979). In the intermediate region $0 < T \ll H$ (in which $\Delta m^2(T)$ exponentially decreases due to the expansion of the universe) the value of $\langle\varphi^2\rangle$ in (8.5) is given by

$$\langle\varphi^2\rangle = \frac{3H^4}{8\pi^2 m^2}\left[1 - \exp\left(-\frac{2m^2}{3H}(t-t_0)\right)\right] \tag{8.6}$$

where t_0 is the time at which the value of $m^2(T) = m^2 + \Delta m^2(T)$ becomes smaller than $2H^2$. Note that at $t - t_0 \ll 3H/2m^2$ fluctuations of the field φ (8.6) grow linearly in time:

$$\langle\varphi^2\rangle = \frac{H^3}{4\pi^2}(t-t_0). \tag{8.7}$$

The value of $\langle\varphi^2\rangle$ in the theory with $m^2 < 0$ at small $t - t_0$ also grows linearly, but at $t - t_0 \gtrsim 3H/2m^2$ it grows exponentially:

$$\langle\varphi^2\rangle \sim \frac{3H^2}{8\pi^2|m^2|}\left[\exp\left(\frac{2|m^2|}{3H}(t-t_0)\right) - 1\right]. \tag{8.8}$$

It is very unusual and very important that the leading contribution to $\langle\varphi^2\rangle$ (8.4)–(8.8) comes from the fluctuations of the field φ with an extremely large wavelength. For example, from equation (8.5) it follows that at $T = 0$ the leading contribution to $\langle\varphi^2\rangle$ (8.4) comes from the fluctuations with the wavelength $l \sim H^{-1}\exp(3H^2/2m^2)$. The leading contribution to $\langle\varphi^2\rangle$ (8.7) comes from the fluctuations with the wavelength $l \sim H^{-1}\exp[H(t-t_0)]$. This means, for example, that in the regime (8.7) the fluctuations of the field φ at a length scale $l \lesssim H^{-1}\exp[H(t-t_0)]$ are practically indistinguishable from the homogeneous classical field φ with magnitude

$$\varphi = (\langle\varphi^2\rangle)^{1/2} = \frac{H}{2\pi}[H(t-t_0)]^{1/2} \tag{8.9}$$

(see figure 7). The domains filled with an almost homogeneous field φ play, in this scenario, the same role as the bubbles of the classical field φ in the first version of the inflationary universe scenario (Linde 1982b, d, Starobinsky 1982, Hawking and Moss 1983).

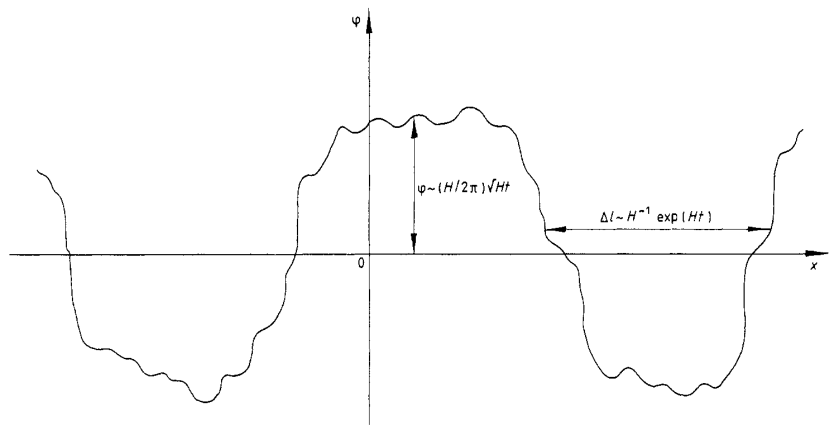

Figure 7. Spatial distribution of fluctuations of the scalar field φ with mass $m^2 \ll H^2$ in the inflationary universe at the stage of linear growth of $\langle\varphi^2\rangle$ (8.7).

Equations (8.4)–(8.8) are valid only at the first stages of the process of the growth of the fluctuations $\langle\varphi^2\rangle$, at which the back-reaction of these fluctuations on the effective mass of the field φ is unimportant. With a growth of $\langle\varphi^2\rangle$ the effective mass squared of the field φ acquires a negative time-dependent contribution:

$$m^2(t) = m^2 - 3\lambda\langle\varphi^2\rangle \tag{8.10}$$

which speeds up the growth of $\langle\varphi^2\rangle$. Let us assume for simplicity that $m^2 = 0$. In this case it can be shown (Linde 1982d, Starobinsky 1982) that the linear growth of $\langle\varphi^2\rangle$ (8.7) persists only during the time

$$\Delta t = t - t_0 \sim \frac{\sqrt{2}\pi}{\sqrt{\lambda} H} \tag{8.11}$$

after which the growth of $\langle\varphi^2\rangle$ becomes very rapid and stops only when the value of $\varphi = (\langle\varphi^2\rangle)^{1/2}$ becomes of the order of φ_0, where φ_0 corresponds to the minimum of $V(\varphi)$. The duration of the process of symmetry breaking in such a theory is of the same order as Δt (8.11). During this time the universe expands $\exp(H\Delta t)$ times, where

$$\exp(H\Delta t) \gtrsim \exp(\sqrt{2}\pi/\lambda). \tag{8.12}$$

This means that the universe expands more than exp(70) times only if the constant λ is sufficiently small:

$$\lambda \lesssim 4 \times 10^{-3}. \tag{8.13}$$

A similar investigation for the SU(5) Coleman–Weinberg theory shows that with the realistic values of the coupling constants the universe during the stage of exponential expansion grows less than exp(70) times (Linde 1982d, Starobinsky 1982, 1984a).

The situation becomes even more complicated if one reconsiders the homogeneity problem in this scenario with an account taken of the quantum fluctuations of the field φ. The point is that the field φ in this scenario is not exactly homogeneous (see figure 7). During the growth of the field φ these inhomogeneities are amplified, which gives rise to perturbations of density after inflation.

The theory of generation of density perturbations after inflation has been developed by many authors (Mukhanov and Chibisov 1981, 1982, Hawking 1982, Starobinsky 1982, Guth and Pi 1982, Bardeen *et al* 1983, Hawking and Moss 1983, Brandenberger and Kahn 1984). The main idea of these papers is that the field φ in different regions of the inflationary universe reaches its equilibrium value φ_0 at different times due to the inhomogeneity of the initial distribution of the field φ (figure 7). The universe will be reheated up to the same temperature T_R in different regions of the universe, but at different times, which just implies the appearance of density perturbations after inflation. We will not discuss here the details of the theory of generation of density perturbations, which is rather involved, but will just present the final result for the spectrum of density perturbations $\delta\rho/\rho$ (Hawking 1982, Starobinsky 1982, Guth and Pi 1982, Bardeen *et al* 1983):

$$\frac{\delta\rho(k)}{\rho}=\frac{1}{\sqrt{2\pi^3}}\frac{H^2}{|\dot{\varphi}|}\bigg|_{\varphi=\varphi_*} \tag{8.14}$$

where k is the momentum, corresponding to the density perturbation $\delta\rho(k)$ at the end of inflation and φ_* is the value of the field $\varphi=(\langle\varphi^2\rangle)^{1/2}$ at the moment at which the momentum k of this perturbation was equal to $k_*=H$. (Some authors take $k_*^2\sim \mathrm{d}^2V/\mathrm{d}\varphi^2|_{\varphi=\varphi_*}$, but for most theories this yields the same result for $\delta\rho(k)/\rho$ with logarithmic accuracy.) The behaviour of the field φ at the stage of linear growth of $\langle\varphi^2\rangle$ is governed by equation (8.7), whereas at later stages the field φ obeys the equation of motion of a classical homogeneous field φ in the exponentially expanding universe (8.1):

$$\ddot{\varphi}+3H\dot{\varphi}=-\mathrm{d}V/\mathrm{d}\varphi. \tag{8.15}$$

(The second term in equation (8.15) appears due to the expansion of the universe.) By means of equations (8.14) and (8.15) it can be shown that in the theory (8.2) after inflation:

$$\frac{\delta\rho(k)}{\rho}=\left(\frac{4\lambda}{3\pi^3}\right)^{1/2}\ln^{3/2}\frac{H}{k}. \tag{8.16}$$

This spectrum is almost scale-independent. Just such a flat spectrum is necessary as a spectrum of initial density perturbations in the theory of galaxy formation (Zeldovich 1972). However, the amplitude of density perturbations at a galactic scale should be very small, $\delta\rho/\rho\sim10^{-4}$. This condition is satisfied in the theory (8.3) only for

$$\lambda\sim10^{-12}. \tag{8.17}$$

For the SU(5) Coleman–Weinberg theory the value of $\delta\rho/\rho$ (8.16) at the galactic scale proves to be O(50), which is absolutely unacceptable.

After this discrepancy was revealed there were many attempts to suggest either another realistic theory of elementary particles, in which a condition of the type (8.17) could be satisfied, or another version of the inflationary universe scenario. The first of these two directions has led to the realisation of the new inflationary universe scenario in the context of supergravity (Ellis *et al* 1983a, b, Nanopoulos *et al* 1983a, b, Linde 1983c) and the second direction has led to the chaotic inflation scenario (Linde 1983d, e). Both these directions are now unified in the context of the chaotic inflation scenario in supergravity (Linde 1983g, Goncharov and Linde 1984a, b, d).

However, before describing these versions of the inflationary universe scenario we would first like to discuss one more scenario, closely related to the inflationary universe scenario—the Starobinsky model (Starobinsky 1979, 1980, 1983b).

9. The Starobinsky model

Two years before the formulation of the inflationary universe scenario a very similar scenario was suggested by Starobinsky (1979, 1980). It is known that de Sitter space (8.1) with the scalar curvature $R = 12H^2 \sim G^{-1} = M_P^2$ is a self-consistent solution of the Einstein equations with quantum corrections (Dowker and Critchley 1976). Starobinsky has pointed out that such a solution is unstable and de Sitter space eventually evolves into the hot Friedmann universe. The theory of this transition (Starobinsky 1980, 1983c) is very similar to the theory of the phase transition in the new inflationary universe scenario (Linde 1982d, Starobinsky 1982).

The main aim of the Starobinsky model was to solve the singularity problem. This has not been completely realised so far. However, in the course of the investigation of the Starobinsky model it became clear that, in the context of this model, one can also solve many of the problems which can be solved in the context of the inflationary universe scenario. At the same time it was realised that the first version of the Starobinsky model (Starobinsky 1979, 1980), just as the first version of the new inflationary universe scenario (Linde 1982a, b, c, d, Albrecht and Steinhardt 1982), was far from being perfect. For example, the universe in the Starobinsky model was assumed to be absolutely homogeneous and isotropic *ab initio*. On the other hand, the density perturbations, which arise after the reheating of the universe in this model, prove to be too large, just as in the model discussed in the previous section (Mukhanov and Chibisov 1981, 1982, Starobinsky 1983b). The temperature of the universe after the reheating in the first version of the Starobinsky model (1979, 1980) was much greater than the critical temperature T_{c_1} of the SU(5) phase transition, $T_R \gg T_{c_1} \sim 10^{15}$ GeV, and therefore it was impossible to solve the primordial monopole problem in this scenario. The main difficulty of this scenario was the following. De Sitter space is non-singular and has an infinite lifetime, $-\infty < t < +\infty$, whereas the de Sitter stage in the Starobinsky model, just as in the inflationary universe scenario, can exist only during some finite time interval. This means that before the de Sitter stage the universe should have been in some other state. In the first versions of the inflationary universe scenario the de Sitter stage appears as a result of the supercooling of the phase $\varphi = 0$. However, in the Starobinsky model the universe was assumed to be in a pure vacuum state before the transition to the hot Friedmann universe. Therefore it was rather difficult to understand what was before the de Sitter stage in this scenario. One possible way of answering this question is that the de Sitter stage in the Starobinsky model appears as a result of an anisotropic collapse of the universe (Gurovich and Starobinsky 1979). Another possibility, suggested by Zeldovich (1981), is that the de Sitter universe as a whole is created from 'nothing' due to quantum gravity effects. Such a realisation is possible not only for the Starobinsky model but for the inflationary universe scenario as well. (See a discussion of this question in § 13.) However, the theory of quantum creation of the universe is far from being completely elaborated. Therefore the inflationary universe scenario, which can be realised in the context of a more usual approach, based on the hot universe theory and described in the previous sections, was somewhat simpler than the Starobinsky model.

Recently the status of the Starobinsky model has changed. It was understood that the perturbations of density in this scenario can be of the necessary magnitude $\delta\rho/\rho \sim 10^{-4}$ if one adds to the Einstein Lagrangian of the gravitational field $-R/16\pi G$ a term of the type R^2/M^2, where $M \sim 10^{14}$ GeV (Starobinsky 1983b). At $M \sim 10^{14}$ GeV the reheating temperature in this model becomes much smaller than the temperature of the SU(5) phase transition, and therefore no monopoles and domain walls appear after reheating. As for the problem of the origin of the de Sitter stage in this model, this problem can be solved not only with the help of the idea of quantum creation of the universe, but also in the context of a scenario similar to the chaotic inflation scenario (Kofman *et al* 1984) (see § 11). Therefore at present the scenario based on the Starobinsky model (Starobinsky 1979, 1980, 1983b) can be considered as a viable alternative to the inflationary universe scenario, and it is not excluded that the future theory of the early stages of evolution of the universe will be based on some synthesis of the inflationary universe scenario and the Starobinsky model (Kofman *et al* 1984).

10. Supergravity and inflation

Now let us return to the inflationary universe scenario and try to understand how one can reduce the density perturbations in this scenario. The simplest way to do this (Linde 1982d) is connected with supersymmetric versions of the Coleman–Weinberg model. Indeed, in the usual Coleman–Weinberg model the effective coupling constant $\lambda = \frac{1}{6}\mathrm{d}^4V/\mathrm{d}\varphi^4$ is of the order of e^4 due to the vector boson contribution to $V(\varphi)$ (Coleman and Weinberg 1973). However, the fermion contribution to $V(\varphi)$ has the opposite sign. In theories with (broken) supersymmetry the contributions from bosons and fermions (partially) cancel each other, which may lead to a very small effective coupling constant $\lambda(\varphi)$. A first realisation of this possibility was suggested by Albrecht *et al* (1983) in the context of the Witten–Dimopoulos–Raby inverted hierarchy model (Witten 1981, Dimopoulos and Raby 1983). It proved possible to have enough inflation and small density perturbations in this model (Albrecht *et al* 1983). However, the effective potential in this model is too flat and therefore no baryon asymmetry can be produced in this model after inflation (Ellis *et al* 1982). For the same reason large inflation in this model proves to be incompatible with efficient reheating after inflation (Albrecht *et al* 1983, Ovrut and Steinhardt 1983).

The problems of reheating and baryon generation can be solved more easily if the inflationary phase transition occurs at a large mass scale of the order of $M_{\mathrm{p}} \sim 10^{19}$ GeV (primordial inflation). This possibility was first suggested in the context of a globally supersymmetric theory by Ellis *et al* (1983a). Later it was suggested that the idea of primordial inflation in the context of $N = 1$ supergravity coupled to matter be implemented (Nanopoulos *et al* 1983b).

The idea that supergravity rather than grand unified theories may be responsible for the most fundamental features of the structure of the universe seems very natural and attractive. However, there were some difficulties, which hampered a direct realisation of the original version of the primordial inflation scenario. These difficulties and their possible resolution in the context of the chaotic inflation scenario in supergravity

(Linde 1983c, g, Goncharov and Linde 1984a, b) will be discussed in this section and also in § 12.

Let us start with the investigation of the original version of the primordial inflation scenario based on the theory of a chiral superfield Σ coupled to supergravity $N=1$ (Nanopoulos *et al* 1983b). According to Cremmer *et al* (1979), the effective potential $V(z, z^*)$ of the first (scalar) component of this superfield is given by

$$V(z, z^*) = \exp\left(\frac{zz^*}{2}\right)\left(2\left|g_z - \frac{z^*}{2}g\right|^2 - 3|g|^2\right) \tag{10.1}$$

where $g_z \equiv \mathrm{d}g/\mathrm{d}z$, $g(z)$ is some arbitrary function called the superpotential which can be represented in the following form:

$$g(z) = \mu^3 f(z) \tag{10.2}$$

where μ is some mass parameter and $f(z)$ is an arbitrary dimensionless function of the field z. To simplify the notation in this section we use the system of units in which $M_{\mathrm{p}}/\sqrt{8\pi} = 1$ (Cremmer *et al* 1979).

Nanopoulos *et al* (1983a, b) have assumed that, due to high-temperature effects in the very early universe, the field z initially was zero (or almost zero), and then the inflationary phase transition with the generation of the classical field $\varphi = \varphi_0$ occurred, where φ is the real part of the field z. It is clear, however, that such a scenario could be realised only for some particular choice of superpotential $g(z)$. In the first papers on primordial inflation the function $f(z)$ was written as follows:

$$f(z) = \sum_{n=1}^{m} \frac{\lambda_n}{n} z^n \tag{10.3}$$

and it was assumed that $\lambda_0 \geqslant 0$, $\lambda_1 > 0$. The effective potential $V(\varphi)$ in this theory is given by

$$V(\varphi) = \mu^6(\alpha + \beta\varphi + \gamma\varphi^2 + \delta\varphi^3 + \ldots) \tag{10.4}$$

where α, β, γ, δ are some functions of λ_n. It was assumed that by a proper choice of the constants λ_n one may choose an effective potential $V(z, z^*)$ which has an absolute minimum at $\varphi = \varphi_0 = 1$. It was also assumed that $V(\varphi_0) = 0$ (in order to have a vanishing cosmological constant at present) and that $g(\varphi_0) = 0$ (to avoid large breaking of supersymmetry at $\varphi = \varphi_0$, see § 12). One could expect that, due to the freedom of choice of different constants λ_n, all these conditions can be satisfied simultaneously. Unfortunately, however, these conditions prove to be incompatible with each other (Ovrut and Steinhardt 1983, Goncharov and Linde 1984a, b).

To make our investigation as simple as possible we introduce the function

$$\psi(z) = \exp(z^2/4) f(z) \tag{10.5}$$

and rewrite the effective potential (10.1) as follows:

$$V(z, z^*) = \mu^6 \exp[-\tfrac{1}{4}(z - z^*)^2](2|\psi_z + \tfrac{1}{2}(z - z^*)\psi|^2 - 3|\psi|^2) \tag{10.6}$$

where $\psi_z = \mathrm{d}\psi/\mathrm{d}z$. At the real axis $z = \varphi$ the effective potential acquires a very simple form:

$$V(\varphi) = \mu^6(2\psi_\varphi^2 - 3\psi^2). \tag{10.7}$$

Now let us note that in the theories discussed above $\psi(0)=f(0)=\lambda_0\geqslant 0$, $\psi_\varphi(0)=f_\varphi(0)=\lambda_1>0$, $\psi(\varphi_0)=0$. However, the conditions $\psi(0)\geqslant 0$, $\psi_\varphi(0)>0$ imply that the condition $\psi(\varphi_0)=0$ can be satisfied only if ψ_φ vanishes at some point $\tilde{\varphi}$ between $\varphi=0$ and $\varphi=\varphi_0$, in which $\psi(\tilde{\varphi})\neq 0$. From (10.7) it follows that $V(\tilde{\varphi})<0$, which means that the point $\varphi=\varphi_0$ is not the absolute minimum of $V(\varphi)$, and that a deeper minimum with $V(\tilde{\varphi})<0$ lies somewhere between $\varphi=0$ and $\varphi=\varphi_0$. Therefore it seems impossible to come from the point $\varphi=0$ to the local minimum of $V(\varphi)$ at $\varphi=\varphi_0$ after inflation in such theories.

Another realisation of the inflationary universe scenario in the context of supergravity was suggested in Linde (1983c, g). In these papers we have not assumed that $g(\varphi_0)=0$ (see, however, § 12) and have considered the simplest effective potential which satisfies the condition $V(\varphi_0)=0$:

$$V(\varphi)=3\mu^6\left(1-\alpha^2\varphi^2+\frac{\alpha^4}{4}\varphi^4\right). \tag{10.8}$$

The minimum of $V(\varphi)$ occurs at $\varphi=\varphi_0=\sqrt{2}/\alpha$. Such a potential with respect to φ can be obtained from (10.1) by a proper choice of $g(z)$.

Let us assume now, as usual, that the universe initially was in the state $\varphi=0$ (for a discussion of this point see, however, the next section). With the expansion of the universe the temperature decreases, the effective potential acquires the form (10.8) and the phase transition with the growth of the classical field φ starts. A typical time, which is necessary for the field φ to grow from $\varphi=0$ to $\varphi=\varphi_0$, proves to be $\mathrm{O}(\mu^{-3}\alpha^{-2})$ (see below). During this time the vacuum energy is approximately equal to $V(0)=3\mu^6$ and the universe expands exponentially, $a(t)\sim\exp(Ht)$, where the Hubble constant $H=((8\pi/3M_\mathrm{p}^2)V(\varphi))^{1/2}=\mu^3$ (in units of $M_\mathrm{p}/\sqrt{8\pi}=1$). To be more precise, one should note that in the case under consideration the classical field $\varphi=\langle\varphi\rangle$ is not generated, but the field $\langle\varphi^2\rangle$ is. However, as was explained in § 8, the leading contribution to $\langle\varphi^2\rangle$ is given by fluctuations with exponentially large wavelengths and therefore the fluctuations of the field φ are almost indistinguishable from the homogeneous classical field $\varphi=(\langle\varphi^2\rangle)^{1/2}$. According to (8.8), the amplitude of the field $\varphi=(\langle\varphi^2\rangle)^{1/2}$ grows as follows:

$$\varphi=\frac{\mu^3}{4\pi\alpha}\{\exp[4\mu^3\alpha^2(t-t_0)]-1\}^{1/2}. \tag{10.9}$$

Fluctuations of the field φ give rise to the density perturbations after inflation, which at the galactic scale for $\alpha=\mathrm{O}\,(10^{-1})$ have the magnitude (Linde 1983c, g):

$$\frac{\delta\rho}{\rho}\sim\frac{\mu^3\exp(10^2\alpha^2)}{20\alpha}. \tag{10.10}$$

Therefore the desirable value $\delta\rho/\rho\sim 10^{-4}$ can be obtained, e.g. at $\alpha\sim 10^{-1}$, $\mu^3\sim 10^{-4}$, which seems quite reasonable.

Now let us estimate the duration of inflation Δt. The value of Δt approximately coincides with the time necessary for the field φ to increase to $\varphi\sim\varphi_0=\sqrt{2}/\alpha$. From (10.9) one obtains

$$\Delta t\sim 6\mu^{-3}\alpha^{-2}\sim 600H^{-1}\sim(3\times 10^{11}\ \mathrm{GeV})^{-1}. \tag{10.11}$$

From (10.10) and (10.11) it follows that in the theory (10.8) one may obtain both a sufficiently large inflation (since $\Delta t \gg 70H^{-1}$) and the desirable value of $\delta\rho/\rho \sim 10^{-4}$, which was the most difficult problem of the new inflationary universe scenario.

Now let us consider the primordial monopole problem in this scenario. In the first papers on primordial inflation it was assumed that the phase transition with breaking of SU(5) and with the monopole production occurs *after* primordial inflation, and therefore one should find some other method to solve the primordial monopole problem in this scenario (Ellis *et al* 1983a, Nanopoulos *et al* 1983b). Note, however, that the temperature of the universe during inflation almost vanishes, and the typical time, which is necessary for the SU(5) symmetry breaking to occur (if supercooling in this phase transition is not anomalously large), is $\tau \sim (10^{15}\,\text{GeV})^{-1}$, which is much smaller than the duration of the inflation (10.11). Therefore the SU(5) phase transition with monopole production occurs not *after* inflation but long *before* the end of inflation, which solves the primordial monopole problem in this scenario.

To be more precise, one should note that in the *minimal* supersymmetric model (Fradkin 1980, Dimopoulos and Georgi 1981, Sakai 1982) the supercooling is very large indeed. The process of symmetry breaking and the solution of the primordial monopole problem in the minimal SU(5) theory coupled to supergravity will be discussed in § 12.

One could argue, however, that if the reheating of the universe occurs during the time $\tau_R \lesssim H^{-1}$ after inflation, as in the SU(5) Coleman–Weinberg theory (Dolgov and Linde 1982), then all the vacuum energy $V(0)$ transforms into heat and the temperature of the universe increases to the reheating temperature $T_R \sim V_{(0)}^{1/4} \sim 10^{17}$ GeV. At such a temperature the SU(5) symmetry would be restored, and then with the cooling of the universe the SU(5) symmetry breaking phase transition with monopole production would occur again.

Fortunately, this process does not actually occur in the theories under consideration because of the very slow reheating in these theories. The particles of the field φ are very weakly coupled to each other and to all other matter fields, and therefore thermalisation occurs at a very late time, and the reheating temperature T_R in this scenario is typically of the order of 10^{10} GeV, which is much smaller than the critical temperature of the SU(5) phase transition (Nanopoulos *et al* 1983a, Linde 1983c, g). The baryon asymmetry of the universe in this scenario can be generated by decay of the Higgs bosons and fermions with masses $m \sim 10^{10}$ GeV (Nanopoulos *et al* 1983a, Krauss 1983). One should note that the existence of the Higgs bosons H_5 with $m_H \sim 10^{10}$ GeV may be incompatible with the large lifetime of the proton, $\tau_p \gtrsim 10^{31}$ yr (A Yu Smirnov 1984 private communication), whereas the existence of fermions with $m_\psi \sim 10^{10}$ GeV is not forbidden but is not very natural. However, in the inflationary universe scenario the baryon asymmetry can be generated even at $T_R \ll m$, since the heavy Higgs bosons and fermions in this scenario can be copiously produced *before* the reheating (Dolgov and Linde 1982). In such a case the baryon asymmetry generation is suppressed by a factor of T_R/m, but one should remember that in the inflationary universe scenario the baryon production may be one or two orders of magnitude more efficient than in the standard baryosynthesis scenario (Dolgov and Linde 1982). Moreover, by a slight variation of the shape of $V(\varphi)$ near $\varphi = \varphi_0$ one can make T_R as large as 10^{16} GeV, which makes it possible to generate the baryon asymmetry by decays of superheavy Higgs bosons (Linde 1983g). Unfortunately, the version of the inflationary universe scenario discussed above is still incomplete for some other reasons (see the next section). An improved version of this scenario will be discussed in § 12.

11. Chaotic inflation scenario

Our previous discussion of phase transitions in SU(5) theory, supergravity, etc, was based on the implicit assumption that the universe initially was in the state corresponding to a minimum of the effective potential $V(\varphi, T)$. Such an assumption at first sight seems absolutely natural, since any non-equilibrium field φ eventually rolls down to a minimum of $V(\varphi, T)$. Let us, however, investigate this question in a more detailed way.

A typical curvature of the effective potential, which arises due to the high-temperature effects, in the high-temperature limit is given by

$$m^2(T) = \frac{d^2 V}{d\varphi^2} = cT^2 \tag{11.1}$$

see equation (3.2). Here c is some combination of coupling constants: $c = \lambda/4$ for the theory $(\lambda/4)\varphi^4$ (2.1), whereas in the theory (10.8) $c \sim \mu^6\alpha^2$. The time necessary for the field φ to drop to the minimum of $V(\varphi)$ exceeds $\tau \sim m^{-1}(T) \sim c^{-1/2}T^{-1}$. On the other hand, the age of the hot universe t is given by equation (4.10), where the typical number of particle species is $N \geqslant 2\times10^2$:

$$t \leqslant \frac{1}{50}\frac{M_p}{T^2}. \tag{11.2}$$

By comparison of τ and t one concludes that the field φ can be influenced by high-temperature effects only at

$$T \leqslant T_* \sim 10^{-2}c^{1/2}M_p \tag{11.3}$$

or, equivalently, at the moment at which the energy density of the hot matter ρ (4.6) becomes sufficiently small:

$$\rho \leqslant \rho_* \sim \frac{10^{-2}}{3N}c^2M_p^4 \sim 10^{-5}c^2M_p^4. \tag{11.4}$$

However, if the effective potential $V(\varphi, T)$ is sufficiently flat, and if the universe initially was in the state with $V(\varphi) \geqslant \rho_*$, then at $T \leqslant T_*$ the universe becomes exponentially expanding and the temperature T vanishes before it could have any effect on the value of the field φ. For example, let us consider the theory (10.8) with any field $\varphi \leqslant \varphi_0/2$. The value of $V(\varphi)$ at $\varphi \leqslant \varphi_0/6$ is given approximately by $3\mu^6(M_p/\sqrt{8\pi})^4$, which is much greater than $\rho_* \sim 10^{-5}\mu^6\alpha^2M_p^4 \sim 10^{-7}\mu^6M_p^4$. Therefore the high-temperature effects cannot lead to the symmetry restoration in this theory. (Another way to show it is to note that the 'critical temperature' in this theory $T_c \sim \alpha^{-1}M_p$ is much greater than $T_* \sim 10^{-1}\mu^3\alpha M_p$(11.3).) Similarly, it can be shown that the high-temperature effects cannot influence the behaviour of the field φ in the theory $(\lambda/4)\varphi^4$ if the field φ was initially greater than $O(10^{-1}\lambda^{1/4}M_p)$ (Linde 1983e).

More generally, this means that in the theories with large values of $V(\varphi)$ and with sufficiently small values of coupling constants (and just such theories are of the most interest for us, see, for example (8.17)) the inflationary universe scenario cannot be realised in a standard way, based on the theory of high-temperature phase transitions (Guth 1981, Linde 1982a, Albrecht and Steinhardt 1982). However, the same reason which leads to the failure of the standard version of the inflationary universe scenario simultaneously makes it possible to suggest a much better scenario, which can be naturally implemented in a large class of realistic theories (Linde 1983d, e).

To explain the main idea of the new scenario, which we have called the chaotic inflation scenario for reasons soon to become clear, let us try to understand how the classical field $\varphi(x)$ could be distributed in the very early universe.

Of course, one could just *assume* that the field φ initially was exactly at the minimum of $V(\varphi, T)$ near $\varphi=0$. However, such an assumption would be even less reasonable than the assumption that the universe was initially in an absolutely symmetric, homogeneous and isotropic state. Indeed, one can easily verify that the value of the effective potential $V(\varphi, T)$ at the Planck time $t_p \sim M_p^{-1}$, at which the classical description of evolution of the universe becomes possible, is defined only with an accuracy of $O(M_p^4)$ due to the uncertainty principle. Therefore one may expect that in the hot universe at $t \sim t_p$ any field $\varphi(x)$ such that $V(\varphi) \lesssim M_p^4$, $(\partial_\mu \varphi)^2 \lesssim M_p^4$ can appear in any point x with an almost φ-independent probability (or at least that there is no strong suppression of the fields $\varphi(x)$ such that $V(\varphi) \lesssim M_p^4$). Moreover, even the constraint $V(\varphi) \lesssim M_p^4$ does not seem obligatory, since according to the classical theory of evolution of the universe the total energy of matter, including $V(\varphi)$, near the singularity was infinitely large (Hawking and Ellis 1973, Zeldovich and Novikov 1975). Therefore we will assume that the initial distribution of the field φ in the universe was more or less chaotic and will not impose any constraints on the field φ, except for a possible constraint $V(\varphi) \lesssim M_p^4$.

Now let us study the evolution of such an initial distribution of the field $\varphi(x)$ in the simplest model with $V(\varphi)=(\lambda/4)\varphi^4$ (without the term $\xi R\varphi^2$, where R is the curvature scalar). We will be especially interested in the evolution of the domains of the universe in which the field φ was initially sufficiently homogeneous (at a scale $l \gtrsim H^{-1}$) and sufficiently large ($\varphi \gtrsim M_p$). As will be shown below, the field φ in such domains decreases very slowly. Consequently, the space inside such a domain behaves as the interior of the quasiexponentially expanding universe with a scale factor

$$a = a_0 \exp\left(\int_0^t H(t)\,\mathrm{d}t\right) \tag{11.5}$$

where $\dot{H} \ll H^2$, and the Hubble 'constant' is given by

$$H = \left(\frac{8\pi}{3M_p^2} V(\varphi)\right)^{1/2} = \left(\frac{2\pi\lambda}{3}\right)^{1/2} \frac{\varphi^2}{M_p}. \tag{11.6}$$

It is important that if the size of such a domain exceeds $2H^{-1}$, where H^{-1} is the radius of the event horizon in de Sitter space, the evolution of the homogeneous field φ inside the domain does not depend on the evolution of this field in the nearby domains (Gibbons and Hawking 1977). The equation of motion of the homogeneous field φ (8.15) in the theory $(\lambda/4)\varphi^4$ looks as follows:

$$\ddot{\varphi} + 3H\dot{\varphi} = \ddot{\varphi} + (6\pi\lambda)^{1/2} \frac{\varphi^2}{M_p} \dot{\varphi} = -\lambda\varphi^3. \tag{11.7}$$

An investigation of this equation shows that at $\varphi \gtrsim M_p/3$ the behaviour of φ practically does not depend on the initial value of $\dot{\varphi}$. If $\dot{\varphi}$ initially was not too large one can neglect the term $\ddot{\varphi}$ in (11.7), which yields

$$\varphi = \varphi_1 \exp\left(-\frac{\sqrt{\lambda}}{\sqrt{6\pi}} M_p t\right) \tag{11.8}$$

where $\varphi = \varphi_1$ at $t=0$. Note that at $\varphi \gg M_p/3$ the kinetic energy $\frac{1}{2}\dot{\varphi}^2$ of the field (11.8) is much smaller than the potential energy $V(\varphi)$, and therefore equation (11.6) is approximately valid. From (11.5) and (11.6) it follows that at large a the scale factor of the universe is given by

$$a = a_0 \exp\left\{\frac{\pi\varphi_1^2}{M_p^2}\left[1-\exp\left(-\frac{2\sqrt{\lambda}M_p}{\sqrt{6\pi}}t\right)\right]\right\}$$
$$= a_0 \exp\left(\frac{\pi}{M_p^2}(\varphi_1^2-\varphi^2)\right). \tag{11.9}$$

Thus, at small t

$$a(t) \sim a_0 \exp\left[\left(\frac{2\pi\lambda}{3}\right)^{1/2}\frac{\varphi_1^2}{M_p}t\right] \tag{11.10}$$

(compare with (11.5)). During the time of quasiexponential expansion (11.9) and (11.10) the universe expands approximately $\exp(\pi\varphi_1^2/M_p^2)$ times. This means that the universe expands more than $\exp(70)$ times if $\varphi_1 \gtrsim 5M_p$.

As we have mentioned above, the only possible constraint on the initial value of φ is the condition $V(\varphi_1) = (\lambda/4)\varphi_1^4 \lesssim M_p^4$. The field $\varphi = \varphi_1 \sim 5M_p$ is quite possible if $\lambda \lesssim 10^{-2}$. This constraint can be satisfied in many reasonable theories including the Glashow–Weinberg–Salam theory.

Analogous results are valid for a rather wide class of effective potentials $V(\varphi)$. Indeed, let us assume again that at large $\varphi(\varphi \gtrsim M_p)$ one can neglect the term $\ddot{\varphi}$ in (11.7) (which can be verified *a posteriori*). In such a case, from equation (11.7) it follows that

$$\frac{d\varphi}{dt} = -\frac{1}{3H}\frac{dV}{d\varphi}. \tag{11.11}$$

The duration of the rolling of the field $\varphi = \varphi_1$ down to $\varphi = 0$ can be roughly estimated as

$$t \sim \varphi_1\left(\left.\frac{d\varphi}{dt}\right|_{\varphi=\varphi_1}\right)^{-1}. \tag{11.12}$$

For a wide class of potentials $V(\varphi)$ (in particular for all polynomial potentials) at large φ one can use an estimate

$$\left.\frac{dV}{d\varphi}\right|_{\varphi=\varphi_1} \sim \frac{V(\varphi_1)}{\varphi_1}. \tag{11.13}$$

From (11.11)–(11.13) it follows that

$$tH \sim \frac{3H^2\varphi_1^2}{V} \sim 8\pi\frac{\varphi_1^2}{M_p^2} \tag{11.14}$$

which, together with equation (11.5), is in qualitative agreement with equation (11.9). (This agreement for the theory $V(\varphi) = (\lambda/4)\varphi^4$ becomes almost complete if one takes into account that in this theory $dV/d\varphi = 4V/\varphi$.) This means that, in a wide class of theories in which the estimates (11.12) and (11.13) are reasonable, the domains of the universe filled with the field $\varphi \gtrsim M_p$ expand quasiexponentially ($\dot{H} \ll H^2$), and the chaotic inflation scenario can be realised. From this point of view inflation is not a peculiar phenomenon which is desirable for a number of reasons discussed in § 5. In a wide class of theories discussed above, inflation proves to be a natural, and maybe even an inevitable, consequence of the chaotic initial conditions in the very early universe.

Thus we see that the original problem of obtaining a sufficiently large inflation in the context of some natural theory of elementary particles is not a problem any more. There still remains the problem, though, of obtaining sufficiently small density perturbations after inflation, $\delta\rho/\rho \sim 10^{-4}$. However, in order to obtain $\delta\rho/\rho \sim 10^{-4}$ in the chaotic inflation scenario it is sufficient, for example, to introduce any field φ weakly interacting with all other fields, which at $\varphi \sim 4M_{\mathrm{P}}$ has an effective coupling constant $\lambda \sim 10^{-12}$ (8.17). The existence of such a weakly interacting field does not contradict any experimental data. In our opinion the introduction of one new field is a very low price to pay for solving about ten different cosmological problems. A possible candidate for the role of such a field is the field φ coupled to supergravity and discussed in the previous section. We will continue the discussion of this possibility in the next section. However, this possibility is certainly not unique and we will consider it just as an example of a more or less realistic model which can be completely investigated. We would also like to note that the field φ is not necessarily an elementary scalar field. It can be a composite field such as a condensate of fermions (or vector bosons) of the type $\langle\bar{\psi}\psi\rangle$ (or $\langle G^a_{\mu\nu}G^a_{\mu\nu}\rangle$). It can also be the curvature scalar R. An assumption of a chaotic initial distribution of R can lead to a local realisation of the Starobinsky model similar to the chaotic inflation scenario (Kofman *et al* 1984).

Before proceeding further let us derive some approximate equations for $\delta\rho/\rho$ in the chaotic inflation scenario, which may be useful, since they are valid for a rather wide class of theories. Thus, from equations (8.14) and (11.6) it follows that

$$\frac{\delta\rho}{\rho} \sim \frac{16}{\sqrt{3}} \left.\frac{V^{3/2}}{M_{\mathrm{P}}^3 |\mathrm{d}V/\mathrm{d}\varphi|}\right|_{\varphi=\varphi_*}. \tag{11.15}$$

From the estimate (11.13) it follows that

$$\frac{\delta\rho}{\rho} \sim \frac{16}{\sqrt{3}} \frac{\sqrt{V(\varphi_*)}}{M_{\mathrm{P}}^3} \varphi_*. \tag{11.16}$$

A typical value of φ_* corresponding to fluctuations on a galactic scale in the chaotic inflation scenario in the theories with $V(\varphi) \sim \varphi^n$ is $\mathrm{O}(5M_{\mathrm{P}})$. For example, in the theory $V(\varphi) = (\lambda/4)\varphi^4$, $\varphi_* \sim 4M_{\mathrm{P}}$. Therefore, from equation (11.16) it follows that the density perturbations on the galactic scale in this scenario can be roughly estimated by the following expression:

$$\frac{\delta\rho}{\rho} \sim 40 \frac{\sqrt{V(\varphi \sim 4M_{\mathrm{P}})}}{M_{\mathrm{P}}^2}. \tag{11.17}$$

This means that $\delta\rho/\rho \sim 10^{-4}$ in the theories in which

$$V(\varphi_*) \sim V(\varphi \sim 4M_{\mathrm{P}}) \sim 10^{-11} M_{\mathrm{P}}^4. \tag{11.18}$$

For the theory $(\lambda/4)\varphi^4$ this yields the same value of λ as the value obtained in § 8:

$$\lambda \sim 10^{-12}. \tag{11.19}$$

In conclusion we would like to remind the reader that the very idea of chaotic initial conditions in the very early universe is rather old (Wheeler 1964, Misner 1969a, b, Rees 1972) (see also the discussion of this question in § 5). However, the main aim of the previous works was to obtain a *globally* homogeneous and isotropic universe, which is hardly possible. In our approach the universe as a whole remains inhomogeneous and anisotropic after expansion, but it contains many exponentially large domains, in each of which space is almost homogeneous and isotropic. Equation (11.9) suggests that the main part of the physical volume of our universe appears as a result of expansion of the domains filled initially with the largest possible field $\varphi = \varphi_1$. If the maximal value of φ_1 is determined by the condition $V(\varphi_1) = \lambda\varphi_1^4/4 \sim M_p^4$, then from equation (11.9) it follows that the typical size of the above-mentioned homogeneous and isotropic domains of the universe after expansion exceeds

$$l \gtrsim M_p^{-1} \exp\left(\frac{6\pi^3}{\lambda}\right)^{1/2} \tag{11.20}$$

which for $\lambda \sim 10^{-12}$ yields

$$l \gtrsim \exp(10^7)\ \text{cm}. \tag{11.21}$$

Thus, though we cannot guarantee equally good conditions everywhere in the universe, according to the chaotic inflation scenario many rather comfortable domains (mini-universes) should exist of a size $l \gtrsim \exp(10^7)$ cm, which is much greater than the size of the observable part of the universe. We cannot guarantee also that we live in the best of the domains. However, as will become clear from the discussion of the large-scale structure of the universe in the next section, one should think twice before crossing the boundary of his own domain.

12. Chaotic inflation in supergravity

Now let us return to the discussion of the inflationary universe scenario in supergravity. As was shown in the previous section, the inflationary universe scenario cannot be implemented in the context of supergravity in a standard way, based on the theory of high-temperature phase transitions. However, this does not invalidate the results obtained in § 10. According to the chaotic inflation scenario, all these results remain valid for domains of the universe in which the field φ was initially sufficiently small, $\varphi \ll \varphi_0$ (Linde 1983g). Moreover, the main results of § 10 (the value of $\delta\rho/\rho$, the reheating temperature T_R, etc) remain practically unchanged for the domains of the universe in which the field φ was initially sufficiently large, $\varphi \gtrsim 5M_p + 2\varphi_0$ (Linde 1983g).

However, the model (10.8) is still not perfect. As was mentioned in § 10, the superpotential $g(z)$ in this model does not vanish at $z = \varphi_0$. Though such a behaviour of $g(z)$ is quite possible, it may not be very good, since in the theories under consideration the mass of the gravitino is proportional to $g(\varphi_0)$ and is usually assumed to be very small, $m_{3/2} \sim 10^2$ GeV (Ellis and Nanopoulos 1982), or $m_{3/2} \sim 10^{-16}$ in the units used in § 10 ($M_p/\sqrt{8\pi} = 1$). This is necessary in order to obtain a comparatively simple solution to the gauge hierarchy problem in the context of supergravity (Ibanez 1982, Barbieri *et al* 1982, Nath *et al* 1982) (see § 5). Therefore, one must either solve the gauge hierarchy problem in some other way or keep $g(z) \approx 0$ in the minimum of $V(z, z^*)$ which corresponds to the vacuum state at present, $z = \varphi_0$.

In the paper by Ovrut and Steinhardt (1983) it was claimed that it is impossible to implement the inflationary universe scenario in supergravity with $V(\varphi_0) = 0$, $g(\varphi_0) = 0$. However, their statement was based on the investigation of high-temperature corrections to $V(z, z^*)$ in this theory (Gelmini *et al* 1983, Ovrut and Steinhardt 1983), which are irrelevant for the realisation of the inflationary universe scenario in supergravity (Linde 1983g) (see § 11). As will be shown below, in the context of the chaotic inflation scenario one can have both $V(\varphi)=0$ and $g(\varphi)=0$ in the absolute minimum of the effective potential (Goncharov and Linde 1984a, b). We will not try to propose here a realisation of the chaotic inflation scenario in a theory with the most natural superpotential, since at present we have no criterion for making a choice between different possible functions $g(z)$. We would just like to show that superpotentials of the necessary type do actually exist and to reveal some basic features of the chaotic inflation scenario in supergravity.

As an example of an appropriate superpotential let us consider the function $g(z)=\mu^3 \exp(-z^2/4)\psi(z)$ (10.5), where

$$\psi(z)=\tanh\zeta \sinh\zeta \qquad \zeta=\xi+\mathrm{i}\eta. \tag{12.1}$$

Here

$$\zeta=\sqrt{\tfrac{3}{2}}(z-\varphi_0) \tag{12.2}$$

and φ_0 is some real number (see below). (In this section, just as in § 10, we use the system of units, in which $M_\mathrm{p}/\sqrt{8\pi}=1$.)

It is seen that $\psi(z)=0$ at $z=\varphi_0$. The effective potential (10.1) in the new variables looks as follows:

$$V(\zeta,\zeta^*)=3\mu^6 \exp\left(\frac{2(\mathrm{Im}\,\zeta)^2}{3}\right)[|\psi_\zeta-\tfrac{1}{3}(\zeta-\zeta^*)\psi|^2-|\psi|^2]. \tag{12.3}$$

With the superpotential (12.1) the effective potential (12.3) at the real axis is given by

$$V(\xi)=3\mu^6(\psi_\xi^2-\psi^2)=9\mu^6\left(1-\frac{2}{3\cosh^2\xi}-\frac{1}{3\cosh^4\xi}\right). \tag{12.4}$$

This potential has a minimum at $\xi=0$ ($\varphi=\varphi_0$) with $V(\varphi_0)=0$. At large $|\xi|$ the effective potential grows and asymptotically approaches $9\mu^6$. In the complex plane $V(\zeta,\zeta^*)$ is positive semi-definite, $V(\zeta,\zeta^*)=0$ only at $\zeta=\mathrm{i}\pi n$ ($n=0,\pm1,\pm2,\ldots$). Due to the presence of the factor $\exp[2(\mathrm{Im}\,\zeta)^2/3]$ in (12.3), the potential $V(\zeta,\zeta^*)$ is exponentially large everywhere outside the narrow region near the real axis. The only exceptions are the exponentially narrow holes (of width of the order of $\exp(-2\pi^2n^2/3)$) near the above-mentioned points $n=\pm1,\pm2,\ldots$.

The chaotic inflation scenario can be implemented in the theory with the effective potential (12.1)–(12.4) as follows. Almost independently of the initial value of the field z in the universe, this field exponentially rapidly drops to the real axis, $\zeta=\xi=\sqrt{\tfrac{3}{2}}(\varphi-\varphi_0)$. After this process the main part of the universe becomes filled with some real field φ, which typically is very large, $|\varphi-\varphi_0|\gg1$. (The probability of dropping to the holes near $\zeta=\mathrm{i}\pi n$, $n=\pm1,\pm2,\ldots$, or *directly* to the minimum at $\zeta=0$ is negligibly small.) Then the field φ exponentially slowly rolls along the real axis down to the minimum of $V(\varphi)$ at $\varphi=\varphi_0$. If the field $\xi=\sqrt{\tfrac{3}{2}}(\varphi-\varphi_0)$ is initially sufficiently large ($|\xi|\gtrsim1$), then the value of $V(\varphi)$ at that time is given by $9\mu^6$ (12.4) and the domains of the universe filled with a field $|\xi|\gtrsim1$ expand exponentially:

$$a(t)\sim\exp(Ht)\sim\exp(\sqrt{3}\mu^3t). \tag{12.5}$$

The effective potential at large $|\xi|$ can be approximated by the asymptotic formula

$$V(\varphi)=9\mu^6(1-\tfrac{8}{3}\exp(-\sqrt{6}|\varphi-\varphi_0|)). \qquad (12.6)$$

From equations (8.15), (12.5) and (12.6) it can be shown that, during the rolling from some initial field $\varphi=\varphi_1$ down to $\varphi=\varphi_0$ the universe expands as follows:

$$a(\varphi=\varphi_0)\sim a(\varphi=\varphi_1)\exp(\tfrac{9}{16}\exp(\sqrt{6}|\varphi_1-\varphi_0|). \qquad (12.7)$$

This means that the universe expands more than exp (70) times (which is necessary for the realisation of the inflationary universe scenario) if the value of $|\varphi-\varphi_0|$ was initially sufficiently large, $|\varphi-\varphi_0|\gtrsim 2$.

In the chaotic inflation scenario in the theory with effective potential (12.7) any initial value of φ can exist in a given domain of the universe with the probability which is almost φ-independent, at least for sufficiently large φ. Therefore, a considerable part of the universe (presumably the main part of the universe) was initially in a state with $|\varphi-\varphi_0|\gtrsim 2$. The regions of the universe with $|\varphi-\varphi_0|\gtrsim 2$ exponentially expand and acquire the size exceeding the size of the observable part of the universe, $l\sim 10^{28}$ cm.

The spectrum of density perturbations in this model is given by

$$\frac{\delta\rho(k)}{\rho}\sim 0.3\mu^3\ln\frac{\mu^6}{k^2} \qquad (12.8)$$

which yields for the fluctuations at the galactic scale ($k\sim\exp(-50)\mu^3$)

$$\frac{\delta\rho}{\rho}\sim 30\mu^3. \qquad (12.9)$$

This means that $\delta\rho/\rho\sim 10^{-4}$ if $\mu^3\sim 3\times 10^{-6}$, i.e. $\mu\sim 10^{-2}$, which seems quite reasonable. In this theory, unlike all other theories considered up to now, we do not need any other small parameters to obtain a sufficiently large inflation and the small value of density perturbations $\delta\rho/\rho\sim 10^{-4}$. As for the parameter φ_0, it determines the rate of decay of particles φ to other particles, $\Gamma_\varphi\sim\mu^3\varphi_0^2$ (Nanopoulos *et al* 1983a, Goncharov and Linde 1984a), and with an appropriate choice of this parameter one can get an adequate temperature T_R of the universe after reheating, which is important for the solution of the gravitino problem in this scenario. In particular, at $\varphi_0=1$ in our scenario $T_R\sim 10^{10}$–10^{11} GeV and the decrease (increase) of φ_0 leads to the decrease (increase) of T_R proportional to φ_0. At small T_R the process of gravitino production after inflation is inefficient (S Weinberg 1983 private communication, Nanopoulos *et al* 1983a), which may help us to solve the primordial gravitino problem (Ellis *et al* 1982). However, for a complete solution of this problem it might be necessary to have T_R as small as 10^9 GeV (Khlopov and Linde 1984). It is still possible to generate the baryon asymmetry of the universe at $T_R\sim 10^9$ GeV (see § 10), but one must confess that it is not very easy (Khlopov and Linde 1984) and it would be much better to make the gravitino heavy and harmless, $m_{3/2}\gtrsim 10^4$ GeV (Weinberg 1982a).

As was shown above, the observable part of the universe in our scenario was formed when the field φ was in the region $|\varphi-\varphi_0|\lesssim 2$. Therefore this scenario can be realised not only in the theory with superpotential (12.1), but in many theories in which $\psi(z)$ is approximately given by (12.1) at $|z-\varphi_0|\lesssim 2$. This means, in particular, that the effective potential $V(\varphi)$ should not be absolutely flat at $|\varphi-\varphi_0|\to\infty$ (12.6), it may grow (though not too rapidly) at $|\varphi-\varphi_0|\gtrsim 2$. Thus, we have found a rather wide class of superpotentials, which have all the properties necessary for the realisation of the chaotic inflation scenario in supergravity (Goncharov and Linde 1984a, b).

Now let us study the problem of symmetry breaking in the supersymmetric SU(5) theory coupled to supergravity in this scenario. As was noted in § 5, the effective potential $V(\Phi)$ in the minimal version of this theory has several different minima with $V(\Phi)=0$: SU(5) invariant minimum ($\Phi=0$), SU(4)×U(1) minimum and SU(3)×SU(2)×U(1) minimum (see figure 5). Supergravity makes the values of $V(\Phi)$ in these minima slightly different from each other, but the SU(5) minimum usually remains the most energetically favourable one (Weinberg 1982b) and, which is even more important, the SU(5) minimum is the only minimum of $V(\Phi, T)$ at high temperatures. Therefore one may wonder how we have succeeded to jump into the SU(3)×SU(2)×U(1) minimum.

To answer this question let us first assume that, due to the high-temperature effects, the field Φ was initially equal to zero. Let us assume that the curvature of the effective potential $V(\Phi)$, i.e. the effective mass squared of the field Φ, is much smaller than H^2, where H is the Hubble constant during inflation in this scenario: $m_\Phi^2 \ll H^2 = (8\pi/3M_p^2)V(\varphi)$. (Note that inflation is governed not by the SU(5) fields Φ, but by some other fields φ with a very large value of $V(\varphi)$ which, in the chaotic inflation scenario, can be as large as M_p^4.) In this case from equations (8.6)–(8.9) it follows that during the inflation the fluctuations of the field Φ grow, which in the exponentially large domains of the universe cannot be distinguished from the homogeneous classical field Φ with magnitude $\Phi \sim (\mathrm{Tr}\langle\Phi^2\rangle)^{1/2}$. With a proper choice of the parameters of the model, the magnitude of the field Φ after inflation can become much greater than $m_x \sim 10^{15}$–10^{16} GeV. (An opposite statement by Olive and Seckel (1983) was based on the investigation of models with small H, whereas in the chaotic inflation scenario the inflation typically starts at $H \gtrsim M_p$.) Note that the growth of the field Φ occurs in all possible directions in the isotopic space. After the end of inflation the field Φ stops its growth and becomes convergently oscillating in the vicinity of a nearest minimum of $V(\Phi)$. Therefore, after the end of inflation the universe becomes divided into many domains with all possible types of symmetry breaking, the typical size of each domain being many orders of magnitude greater than the size of the observable part of the universe, $l \sim 10^{28}$ cm. In particular, there will be many (in an open universe, infinitely many) domains of the phase SU(3)×SU(2)×U(1), in one of which we now live (Linde 1983f).

We would like to emphasise the difference between the above-mentioned mechanism of symmetry breaking (Linde 1983f) and the standard one. Usually it is assumed that it is impossible for the phase transition to occur from the global minimum of the effective potential $V(\varphi)$ to any local minimum of $V(\varphi)$. From our results it follows, however, that such a phase transition becomes effectively possible in the inflationary universe scenario due to the anomalous growth of the long-range scalar field fluctuations in the exponentially expanding universe. Usually the phase transition proceeds to one preferred phase in the whole universe. In our case the phase transition proceeds with a comparable probability to many different phases. Each of these phases fills a domain (mini-universe) of a size exceeding the size of the observable part of our universe. All these phases with a non-negative vacuum energy density are practically stable. The only phase in which life of our type may exist is the phase SU(3)×U(1) with vanishing vacuum energy. This phase appears after symmetry breaking in the SU(3)×SU(2)×U(1) domains and proves to be absolutely stable in the theories under consideration (Weinberg 1982b).

Note that in the versions of the inflationary universe scenario in which the value of $V(\varphi)$ in some domains was initially sufficiently large, high-temperature effects could not lead to the symmetry restoration in the SU(5) theory due to the rapid decrease of temperature in such domains (see the previous section). Therefore the field Φ may take different values in different domains of the universe not only due to the anomalous growth of the fluctuations of the field Φ in the inflationary universe scenario, but also due to the chaotic initial distribution of the field Φ in different parts of the universe. In both cases the universe after inflation becomes divided into many mini-universes, in which all possible symmetry breaking patterns of the SU(5) theory can be realised.

The large size of domains, just as the large size of bubbles in the first version of the new inflationary universe scenario (Linde 1982a), implies that no monopoles appear in the observable part of the universe after inflation in this scenario. This solves the primordial monopole problem in the minimal supersymmetric SU(5) theory. As was shown in § 10, in the non-minimal SU(5) theory the primordial monopole problem can be solved in an even more simple way.

The results obtained above show that the chaotic inflation scenario can be completely realised in the context of $N=1$ supergravity coupled to matter (Linde 1983f, g, Goncharov and Linde 1984a, b) under several conditions which are necessary to solve the gravitino problem and to make possible the baryon asymmetry generation (Khlopov and Linde 1984).

As we have noted in § 11, such a realisation certainly is not unique. One can suggest a realisation of the inflationary universe scenario even without any help from supersymmetry. For example, in a recent paper by Shafi and Vilenkin (1984) (see also Pi (1984)) it was suggested to implement this scenario in a theory of a singlet scalar field φ weakly coupled to the SU(5) scalar fields Φ and H_5. It can be shown that the chaotic inflation scenario (though not the new inflationary scenario based on the theory of high-temperature phase transitions) can be realised in the models of this type and the value $\delta\rho/\rho \sim 10^{-4}$ can be obtained. However, as we have argued in § 11, it is now not very difficult to suggest a new realisation of the inflationary universe scenario; the real problem here is to implement this scenario in the context of a natural and realistic theory of elementary particles. It is now widely believed that the theories based on supergravity are the best candidates for the role of a realistic theory of all fundamental interactions. This is the main reason why we have discussed inflation in supergravity in this section and in § 10.

The models considered above were based on the simplest version of spontaneously broken $N=1$ supergravity. For a discussion of inflation in other theories such as SU(1, 1) supergravity see Gelmini *et al* (1984) and Goncharov and Linde (1984d).

To conclude this section we would like to note that the appearance of the domain structure of the universe after inflation, which was discussed above, may help us to answer some other questions, including the question of why our universe is four-dimensional. As we have noted in § 5, in the Kaluza–Klein theories it is assumed that our space originally had dimension $d>4$, but extra $d-4$ dimensions are spontaneously compactified (see, for example, Cremmer and Sherk 1976, Witten 1981). One may wonder, therefore, why just $d-4$ dimensions are compactified, and not $d-3$ or $d-5$?

A possible answer is that, due to some properties of the Kaluza–Klein theories, the only possible compactification is just $d \to 4$ (Freund and Rubin 1980, Freund 1982). The inflationary universe scenario suggests another possible solution to this problem (Linde 1983a). Let us assume that space, after compactification, may have any number of dimensions including $d=4$. It is clear that the process of compactification (or decompactification?) proceeds independently in different causally unconnected domains of the universe. Therefore, the universe after compactification becomes divided into many different domains, in which the number of uncompactified dimensions may be different. (There is nothing paradoxical in such a possibility, since locally (i.e. at a scale $\Delta l \lesssim M_p^{-1}$) the number of dimensions everywhere remains equal to the initial number d.) After inflation the universe becomes divided into many mini-universes, each of which may have a different number of uncompactified dimensions. Now let us note that the conditions necessary for the existence of life (or, at least, of our kind of life, based essentially on electromagnetic and gravitational interactions) can be realised just in four-dimensional space–time. Indeed, as was noted by Ehrenfest many years ago (Ehrenfest 1917), in a space with $d > 4$ electrostatic and gravitational forces decrease too rapidly with the increase in distance between interacting objects, and therefore any bounded states such as atoms or planetary systems at $d > 4$ are impossible. On the other hand, according to general relativity theory, the gravitational attraction between far-removed objects at $d < 4$ vanishes altogether. Therefore, according to our scenario, the universe may consist of many mini-universes with a different number of dimensions, and we live in the four-dimensional universe since life of our type is impossible in domains with any other number of uncompactified dimensions.

One could notice that in this section we have used the anthropic principle several times, which was criticised in § 5. However, after the formulation of the inflationary universe scenario the status of the anthropic principle was changed. First of all, the inflationary universe scenario makes it possible to answer some questions which could not be answered by means of the anthropic principle alone (why the universe is almost isotropic and homogeneous, why the spectrum of inhomogeneities in the universe is almost scale-independent, etc (see § 5)). On the other hand, the inflationary universe scenario provides the conditions which are necessary for the implementation of the anthropic principle. Our universe after inflation becomes divided into many domains (mini-universes) with different properties of elementary particles inside each of them, with different values of the vacuum energy (of the cosmological term) and may even have a different number of dimensions, and life can exist only in some of these mini-universes which are sufficiently suitable for it. The anthropic principle can be developed even further in the context of quantum cosmology, which will be discussed in the next section.

13. The inflationary universe scenario and quantum cosmology

The standard assumption which is usually made in the inflationary universe scenario is that the exponential expansion starts after some earlier stage of expansion and cooling of the hot singular universe. However, one may wonder what was *before* the beginning of the universe expansion (if such a question makes any sense) (see § 5). Here we would like to describe some attempts to answer this question based on the idea of quantum creation of the inflationary universe.

The possibility of the quantum creation of the universe from 'nothing' or from 'some other universe' has been extensively discussed during the last ten years (see, for example, Tryon (1973), Fomin (1973, 1975), Brout *et al* (1978, 1979), Zeldovich (1981), Grishchuk and Zeldovich (1982), Sato *et al* (1982), Vilenkin (1982, 1983a, b), Hartle and Hawking (1983), Moss and Wright (1983), Linde (1984c, d) and Starobinsky (1984a, b, c)). The theory of such processes is far from being completely developed, and even the very concept of creation from 'nothing' or from 'some other universe' deserves a more detailed examination (Linde 1983a). Nevertheless some simple qualitative description of these processes can easily be suggested (Zeldovich 1981, Grishchuk and Zeldovich 1982, Linde 1984c, d).

It is usually believed that space is foam-like at small scales (Wheeler 1964, Hawking 1978), which means that at a scale $\Delta l \lesssim M_{\rm P}^{-1}$ quantum gravity effects lead to very large fluctuations of the metric and of all matter fields. Now let us assume that, as a result of such fluctuations, there appears a domain filled with a slowly changing field φ with the energy density $V(\varphi)$. If the size of this domain Δl is greater than the horizon radius $H^{-1} = (3M_{\rm P}^2/8\pi V(\varphi))^{1/2}$, the interior of this domain exponentially expands as a de Sitter space independent of any events outside this domain (Gibbons and Hawking 1977). Since the typical size of such a domain is $\Delta l \sim M_{\rm P}^{-1}$, quantum gravity effects can lead to a creation of an inflationary universe if $M_{\rm P}^{-1} \gtrsim H^{-1} = (3M_{\rm P}^2/8\pi V(\varphi))^{1/2}$, which yields

$$V(\varphi) \gtrsim M_{\rm P}^4 \tag{13.1}$$

whereas creation of an inflationary universe with $V(\varphi) \ll M_{\rm P}^4$ should be strongly suppressed. Note that in all early versions of the inflationary universe scenario (Guth 1981, Albrecht and Steinhardt 1982, Linde 1982a) inflation occurs only at $V(\varphi) \ll M_{\rm P}^4$, whereas in the scenario discussed in § 11 (Linde 1983d, e) inflation can occur even at $V(\varphi) \gtrsim M_{\rm P}^4$. This makes quantum creation of the inflationary universe possible (Linde 1984c, d).

These qualitative arguments are rather general and can be applied both to the process of creation of the universe from 'nothing' or from 'some other universe'. In what follows we will concentrate on the investigation of the first of these possibilities. Note that the condition $\Delta l \lesssim M_{\rm P}^{-1}$ implies, in particular, that the Friedmann universe created from 'nothing' should be closed.

A first attempt at quantitative investigation of the creation of a closed inflationary universe from 'nothing' was made by Vilenkin (1982, 1983a). However, in our opinion (Linde 1983a) his approach to this problem, being intuitively appealing, was not well motivated. As a result, he obtained the expression for the probability of the universe creation $P \sim \exp(3M_{\rm P}^4/8V(\varphi))$, from which it would follow that quantum gravity effects become stronger at small $V(\varphi)$ and that quantum fluctuations of metric become greater at greater length scales. Such a conclusion would be in contradiction with the well-known fact that quantum gravity effects at large scales (at small momenta) are negligibly small.

Recently a very interesting approach to the problem of quantum creation of the universe was suggested by Hartle and Hawking (1983). This approach is based on the computation of the ground-state wavefunction $\Psi_0(a, \varphi)$ of a closed universe with a scale factor a filled with a homogeneous field φ (DeWitt 1967, Wheeler 1968) which, according to Hartle and Hawking (1983), in the semiclassical approximation is given by

$$\Psi_0(a, \varphi) \sim \exp[-S_{\rm E}(a, \varphi)]. \tag{13.2}$$

Here $S_{\rm E}(a, \varphi)$ is the Euclidean action corresponding to the Euclidean solutions of the Lagrange equations for $a(\tau)$ and $\varphi(\tau)$ with the boundary conditions $a(0) = a$, $\varphi(0) = \varphi$. The main idea behind the derivation of equation (13.2) can be explained as follows. Let us consider the Green function of a particle, which moves from the point $(0, t')$

to the point $(\boldsymbol{x}, 0)$:

$$\langle \boldsymbol{x}, 0|0, t'\rangle = \sum_n \Psi_n(\boldsymbol{x})\Psi_n(0)\exp(\mathrm{i}E_n t')$$
$$= \int \mathrm{d}\boldsymbol{x}(t)\exp[\mathrm{i}S(\boldsymbol{x}(t))] \tag{13.3}$$

where $\Psi_n(\boldsymbol{x})$ is a complete set of energy eigenstates corresponding to the energies $E_n \geqslant 0$. To obtain $\Psi_0(\boldsymbol{x})$ one should make a rotation $t \to \mathrm{i}\tau$ and take the limit as $\tau' \to -\infty$. In the sum, only the term $n=0$ with $E_0 = 0$ survives and the integral transforms into $\int \mathrm{d}\boldsymbol{x}(\tau)\exp[-S_{\mathrm{E}}(\boldsymbol{x}(\tau))]$. A generalisation of this result for the case of interest in the semiclassical approximation would yield equation (13.2).

A possible Euclidean solution, corresponding to the quantum creation of an inflationary universe with a slowly changing scalar field φ, is the Euclidean section of a closed de Sitter space with $a(\tau) = H^{-1}(\varphi)\cos H\tau$. Here $H(\varphi) = [(8\pi/3M_{\mathrm{P}}^2)V(\varphi)]^{1/2}$ and $0 \leqslant H\tau \leqslant \pi/2$. The corresponding action is $S_{\mathrm{E}}(a, \varphi) = -3M_{\mathrm{P}}^4/16V(\varphi)$. According to Hartle and Hawking (1983), this means that the probability of quantum creation of a closed inflationary universe with $a(t) = H(\varphi)\cosh Ht$ is given by $P \sim |\Psi_0(a, \varphi)|^2 \sim \exp(3M_{\mathrm{P}}^4/8V(\varphi))$, which would coincide with the result of Vilenkin (1982, 1983a). However, the negative sign of the action suggests that something should be improved in this approach. The point is that the Lagrangian of the scale factor a in the mini-superspace approach used by Hartle and Hawking (1983) has a 'wrong' sign (Gibbons *et al* 1978). The part of the action which depends on a can be written as follows (Hartle and Hawking 1983):

$$S(a) = -\frac{1}{2}\int \mathrm{d}\eta\left[\left(\frac{\mathrm{d}a}{\mathrm{d}\eta}\right)^2 - a^2 + \frac{\Lambda}{3}a^4\right]\frac{3\pi M_{\mathrm{P}}^2}{2} \tag{13.4}$$

where η is the conformal time, $\eta = \int \mathrm{d}t/a(t)$ and Λ is the cosmological constant, $\Lambda = 8\pi V(\varphi)/M_{\mathrm{P}}^2$ for a slowly changing field φ. Equation (13.4) implies that the energy of 'excitations' of the scale factor a near $a=0$ is negative, $E_n \leqslant 0$. This is related to the fact that the total energy of a closed universe is zero, being a sum of the positive energy of matter and the negative energy of the scale factor a. In such a case, as far as the evolution of the field φ can be neglected, to obtain $\Psi_0(a, \varphi)$ by means of equation (13.2) one should rotate t not to $-\mathrm{i}\tau$ but to $+\mathrm{i}\tau$, which leads to the following result:

$$\Psi_0(a, \varphi) \sim \exp[S_{\mathrm{E}}(a, \varphi)] \sim \exp\left(-\frac{3\pi M_{\mathrm{P}}^2}{2\Lambda}\right) \tag{13.5}$$

(Linde 1984c, d).

We would like to point out that no general Euclidean prescription is known at present for the quantisation of interacting fields with both positive and negative energies, since the vacuum state of such a theory is unstable due to the gravitational interaction of species of both types (Linde 1984a). Fortunately, during the process of creation of the inflationary universe the evolution of the scalar field φ is unimportant ($\dot\varphi/\varphi \ll \dot a/a$, $\dot\varphi^2 \ll V(\varphi)$), its effective potential $V(\varphi)$ serves just as a cosmological constant in the action of the scale factor $S(a)$ (13.4), and equation (13.5) is valid. However, at present, when the universe is very large and the evolution of the scale factor a is extremely slow, the situation is reversed. The evolution of the scale factor is unimportant for quantisation of all other fields, and one can use the standard Euclidean formalism with rotation $t \to -\mathrm{i}\tau$ for the quantisation of all other fields with positive energy.

Another subtle point is that, due to the gauge invariance with respect to the choice of the time scale, there are no physical degrees of freedom (no particles) associated with the scale factor a, and therefore one should take into account the contribution of the corresponding ghosts to $\Psi_0(a, \varphi)$ (Hawking 1984c). However, just as in the theory of instantons, in the semiclassical approximation one can forget about ghosts and consider the gauge field a as a real physical field with negative energy.

Equation (13.5) has also recently been obtained by a different method (Starobinsky 1984b). The wavefunction which describes the creation of the universe from 'nothing' is given by the amplitude of tunnelling from the point $a=0$ to the point $a=H^{-1}(\varphi)$. The Lagrange equations for the scale factor a coincide exactly with the Lagrange equations in the theory with $\tilde{S}(a) = -(2/3\pi M_p^2)S(a)$ (Linde 1984a) (see also appendix 2). Therefore the semiclassical tunnelling amplitude coincides with the tunnelling amplitude in the theory $\tilde{S}(a)$ with the usual sign of energy of the scale factor a. This corresponds to the tunnelling from $a=0$ to $a=H^{-1}$ through the potential barrier

$$\tilde{V}(a) = \frac{1}{2}\left(a^2 - \frac{\Lambda}{3}a^4\right) \tag{13.6}$$

which leads again to equation (13.5).

Note that the amplitude of creation of a universe with $\Lambda \leqslant 0$ vanishes, since tunnelling from $a=0$ in the theory (13.6) with $\Lambda \leqslant 0$ is impossible.

Now let us apply our results to the theory of quantum creation of the inflationary universe. From equation (13.5) it follows that the wavefunction of the universe with a scale factor a filled with a slowly changing homogeneous field φ is given by

$$\Psi_0(a, \varphi) \sim \exp\left(-\frac{3M_p^4}{16V(\varphi)}\right). \tag{13.7}$$

This equation implies that quantum creation of an inflationary universe with $V(\varphi) \ll M_p^4$ is exponentially suppressed, just as we expected. However, there is no exponential suppression of creation of an inflationary universe with $V(\varphi) \gtrsim M_p^4$. In particular, in the theory $V(\varphi) = (\lambda/4)\varphi^4$ with $\lambda \lesssim 10^{-2}$ there is no suppression of creation of the universe with $\varphi \gtrsim \lambda^{-1/4}M_p \gtrsim 5M_p$, which expands more than exp (70) times during the exponential expansion stage (see § 11). This means that the process of quantum creation of the universe in a wide class of elementary particle theories with a large probability leads to the creation of the inflationary universe which, after inflation, acquires the size $l \gtrsim 10^{28}$ cm (Linde 1984c, d).

The results obtained above seem rather plausible. They are in agreement with the qualitative scenario of quantum creation of the universe suggested by Zeldovich (1981) and Grishchuk and Zeldovich (1982). One should bear in mind, however, that the derivation of equations (13.2) and (13.5) is based on some reasonable assumptions about the ground-state wavefunction of the universe, but this derivation is still far from being completely rigorous (Hartle and Hawking 1983, Linde 1984c, d). The physical interpretation of these results is even less definite but is strikingly interesting.

One could imagine that the metric $g_{\mu\nu}$ is some physical field which depends on the coordinates $(x_0, x_1, \ldots)$ and which, just as the scalar field φ, may or may not have a classical part. The scalar field φ acquires a classical part due to the instability of the state $\varphi=0$ in the theory (2.1) with $M^2 = -\mu^2 < 0$. Any quantum fluctuation of the field φ with a momentum $|\boldsymbol{k}| < M$ grows exponentially, $\delta\varphi(\boldsymbol{k}) \sim \exp[(\mu^2 - \boldsymbol{k}^2)^{1/2}t]$, which finally leads to the appearance of the classical field $\varphi = \varphi_0$, corresponding to the minimum of $V(\varphi)$ (Linde 1979) (see also the discussion of the generation of the 'almost classical' field $\varphi = (\langle\varphi^2\rangle)^{1/2}$ in § 8). A similar effect occurs with any fluctuation of the metric which gives rise to the inflationary universe. Any fluctuation of the metric at a scale $\Delta l \gtrsim H^{-1}(\varphi) = (3M_p^2/8\pi V(\varphi))^{1/2}$ containing the field $\varphi \gtrsim 5M_p$ grows as a closed de Sitter space ($a(t) \sim H^{-1} \cosh Ht$), and at large t the metric $g_{\mu\nu}$ acquires an exponentially large classical part (8.1). As we mentioned at the beginning of this section, such fluctuations can appear with a large probability if $\Delta l \lesssim M_p^{-1}$ or, equivalently, if $V(\varphi) \gtrsim M_p^4$, which is in agreement with equation (13.7).

At later stages of expansion the instability of the initial distribution of the field φ leads to reheating of the universe. Thus, the possibility of the existence of life, according to this scenario, appears as a combined effect of the gravitational instability (exponential expansion) and the instability with respect to the field φ. Even later the closed universe contracts again to a singularity, and the 'classical part' of $g_{\mu\nu}$ disappears again. From such a point of view the universe looks like a quasivirtual state, as a quantum fluctuation, which, due to the existence of the stage of inflation, has an exponentially large lifetime. Such fluctuations may appear again and again, which leads to a scenario similar to the eternally oscillating universe scenario suggested by Markov (1981, 1983, 1984) (see also Linde (1983a) and Hawking (1984d)). In a different interpretation one should not speak about an oscillating universe; all these oscillations represent the same universe which is split into infinitely many slices according to the many-world interpretation of quantum mechanics (Everett 1957, DeWitt 1967, Wheeler 1968, Hartle and Hawking 1983, Hawking 1984c). Actually, however, there may be no difference between these interpretations. Indeed, the concept of time is well-defined only at a sufficiently large a, when the semiclassical description of evolution of the universe is possible, and therefore one cannot say whether it is the next oscillation that starts after the singularity or whether it is just another slice of the same universe. This is a possible answer to the question of what was before the cosmological singularity.

One should note that the chain of oscillations can be broken if a universe is created in which, after the symmetry breaking, $V(\varphi)$ remains positive ($\Lambda > 0$). Such a universe remains exponentially expanding forever. However, this does not preclude a further process of quantum creation of universes, since the process of quantum creation of universes from the universe with $\Lambda > 0$ remains possible.

A more difficult problem arises if one considers the entropy behaviour in the oscillating universe scenario, namely, after the end of inflation the universe acquires a huge total entropy $S \gtrsim 10^{87}$ and therefore it will be not in the ground state $\Psi_0(a, \varphi)$ but rather in some highly excited state. The total energy E of all matter fields in the universe (and, correspondingly, the energy of the scale factor a) in this state is extremely large, and it increases towards the singularity ($E \sim a^{-1} \sim (t - t_c)^{-1/2}$ in the hot collapsing universe, where t_c is the time at which the universe becomes singular). Therefore one could argue that the behaviour of the wavefunction of the universe $\Psi(a, \varphi)$ after inflation is essentially semiclassical, and thus the quantisation of the scale factor of the universe does not save us from the final cosmological singularity unless the structure of the energy–momentum tensor at high densities drastically changes (e.g. $T_{\mu\nu} \to g_{\mu\nu}M_p^4$ at $\rho \to M_p^4$ (Markov 1982), see appendix 1). Since the collapsing universe is not in the ground state, the universe which appears after the collapse may also be created not in the ground state, and the total entropy of the universe will increase in each new cycle, which would lead to the well-known difficulties of the oscillating universe scenario (Zeldovich and Novikov 1975).

A possible answer is that in the very vicinity of the singular point one could not measure either the total energy E (since the indefiniteness of the total energy $\sim (t-t_c)^{-1} \gg E \sim (t-t_c)^{-1/2}$) or the total entropy, and the concept of time is also ill-defined, since there is no clock which could work at $t \to t_c$. Therefore, there is no reason to believe that the total entropy of the universe must grow with each new cycle; the growth of the entropy may be a quasivirtual effect, which occurs only inside each quasivirtual universe in the semiclassical stage of its evolution. Another possible solution of the entropy problem is discussed in appendix 1.

If one takes a risk to discuss a manifold without any classical part of $g_{\mu\nu}$, one can take another step and consider spaces with different possible signatures of $g_{\mu\nu}$. From such a point of view, a newly born universe can have the Euclidean signature $(++++)$, or it can have the signature $(--++)$, or it can consist of different domains with different signatures of $g_{\mu\nu}$ (Sakharov 1984). One can also imagine that the manifold $(x_0, x_1, \ldots)$ has a number of dimensions much greater than four, but some of the classical components of $g_{\mu\nu}$ are very small (compactification) or they have no classical part at all, whereas the other components of $g_{\mu\nu}$ may have a different signature in different parts of the manifold $(x_0, x_1, \ldots)$ (Sakharov 1984). It is worth noting that life of our type in the Euclidean universe would be impossible due to the absence of particle-like states. This is a possible reason why we live in the Minkowski slice of the universe, which is four-dimensional for the reasons discussed in the previous section.

Thus we see that the investigation of quantum cosmology reveals many interesting possibilities which have not been explored so far. At present we do not know which of these possibilities will be realised in a future theory, but in any case it seems very likely that something similar to the inflationary universe scenario should be present in a complete cosmological theory as a bridge between the creation of the universe and its later stages of evolution as described by the hot universe theory.

14. Conclusions

During the three years of its existence the inflationary universe scenario has been considerably modified. According to the first version of this scenario, inflation occurs *before* the symmetry breaking phase transition from some strongly supercooled vacuum state (Guth 1981). Then it was realised that this scenario leads to some unacceptable cosmological consequences. In 1982 the new inflationary universe scenario was suggested, which was free of the main difficulties of the 'old' scenario (Linde 1982a, b, c, d, Albrecht and Steinhardt 1982). According to the new scenario, inflation occurs *during* the symmetry breaking phase transition. In 1983 it was discovered that the idea of inflation can be realised in a much better way in the chaotic inflation scenario (Linde 1983c, d), which was *not* based on the theory of high-temperature phase transitions in the early universe, and this scenario was implemented in the context of $N=1$ supergravity coupled to matter (Linde 1983g, Goncharov and Linde 1984a, b, d). Very recently a new version of the chaotic inflation scenario was suggested, which was based on the idea of quantum creation of the universe (Hawking 1984c, d, Linde 1984c, d). We do not know what will be the next modification of the inflationary universe scenario. We do not know what impact new theories of the type of extended supergravities, Kaluza–Klein theories, etc, will have on this scenario (see, for example, Shafi and Wetterich (1983)). However, whereas the details of the inflationary universe scenario will certainly be modified with the development of elementary particle theories, some of the basic features of this scenario discussed in the present paper will presumably remain intact.

One may wonder whether the inflationary universe scenario can be experimentally tested. The answer to this question is two-fold. First of all, as we have noted in the introduction, the only laboratory in which particles with energies $E \sim 10^{15}$ GeV have ever interacted with each other is our universe at the very early stages of its evolution. The main results of this cosmological experiment are the flatness, homogeneity, isotropy and baryon asymmetry of the universe, the absence of primordial monopoles and domain walls and the existence of galaxies. All of these experimental data can be explained in the context of the inflationary universe scenario, and no other possibility of explaining simultaneously all these data (except the new version of the Starobinsky model (Starobinsky 1983b, Kofman *et al* 1984)) is known to us at present.

However, the inflationary universe scenario not only explains a large set of existing experimental data, it also predicts some new ones. According to the inflationary universe scenario, the density of matter ρ in the observable part of the universe should now be almost exactly equal to the critical density $\rho_c \sim 10^{-29}$ g cm^{-3}: $\Omega = \rho/\rho_c = 1 + O(10^{-4})$ (Guth 1983). (The term $O(10^{-4})$ appears due to the density perturbations in the observable part of the universe.) Though it is practically impossible to verify this prediction by a direct measurement of ρ, one can verify it indirectly by measuring the masses of neutrinos. Indeed, the neutrino density in the universe can be theoretically evaluated by means of the hot universe theory (Zeldovich and Novikov 1975) and it can be shown that the neutrino energy density in the universe would exceed the critical density ρ_c if $m_{\nu_e} > 25$ eV. Thus the measurement of the neutrino mass (Lubimov *et al* 1980, 1983) can be considered as an experimental test for the inflationary universe scenario. This test is not crucial, since there exist some ways to considerably reduce the theoretical estimate of the present neutrino density in the universe (Doroshkevich and Khlopov 1984a, b). However, it would not be very easy to reconcile the inflationary universe scenario with the existence of a stable neutrino with $m_\nu \gtrsim 10^3$ eV (M Yu Khlopov 1984 private communication).

Another possible test is related to the large-scale fluctuations of the microwave background radiation temperature $\Delta T/T$, which should not be much smaller than 0.3 $\delta\rho/\rho \sim 3 \times 10^{-5}$ according to the theory of galaxy formation from adiabatic perturbations with the scale-independent spectrum $\delta\rho/\rho \sim 10^{-4}$ and $\Omega = \rho/\rho_c \approx 1$ (Rubakov *et al* 1982, Shandarin *et al* 1983, Starobinsky 1983b, Mukhanov and Chibisov 1984). Meanwhile, from the present observational data it follows that the large-scale fluctuations $\Delta T/T$ are rather small, $\Delta T/T \lesssim (4\text{–}6) \times 10^{-5}$ (Ceccarelli *et al* 1983, Fixsen *et al* 1983). These results are still not in disagreement with the inflationary universe scenario, but in a few years time with an increase of accuracy in the measurement of $\Delta T/T$ we will have a chance of verifying one of the important predictions of the inflationary universe scenario: the almost scale-independent spectrum of density perturbations, $\delta\rho/\rho \sim 10^{-4}$. This test is also not absolutely crucial. One can have both $\delta\rho/\rho \sim 10^{-4}$ and $\Delta T/T \lesssim 10^{-5}$ in the theory with heavy unstable neutrinos (Doroshkevich and Khlopov 1984a, b, Turner *et al* 1984). Another possible way of reducing the large-scale fluctuations $\Delta T/T$ is to suggest a version of the inflationary universe scenario in which $\delta\rho/\rho \sim 10^{-4}$ at the galaxy scale, but $\delta\rho/\rho \ll 10^{-4}$ at larger length scales. This is actually possible in some theories, though it does not seem to be very natural. Finally, one can imagine, for example, that the density perturbations after inflation are very small, $\delta\rho/\rho \ll 10^{-4}$, and the perturbations of density with a slightly different spectrum $\delta\rho(k)/\rho$

are generated by cosmic strings (Zeldovich 1980, Vilenkin and Shafi 1983). However, this would diminish the aesthetic attractiveness of the inflationary universe scenario.

It is very difficult to make any predictions concerning the future development of the inflationary universe scenario. Two years ago, the difficulties of this scenario seemed absolutely insurmountable. A year ago, it was not quite clear whether the new version of this scenario would be sufficiently good or not. At present, it seems that there are no difficulties which would prevent a consistent realisation of the inflationary universe scenario, but it is not excluded, of course, that some better scenario will be suggested later. However, whatever the fate of the inflationary universe scenario, its very existence may have irreversible consequences for the development of the theory of evolution of the universe. During the last three years it has been understood that it is possible to solve simultaneously about ten different cosmological problems in the context of one comparatively simple scenario. It was shown that the theory of the very early stages of evolution of the universe may differ considerably from the usual hot universe theory. It became clear that the universe at very large scales may consist of many locally isotropic and homogeneous mini-universes, and inside each of these mini-universes the properties of elementary particles, the vacuum energy and even the dimensionality of space–time may be very different.

There still remain many problems to be solved, related both to cosmology and to elementary particle physics, some of which we will discuss in the appendices to this review. However, in our opinion, the new results which have already been obtained during the investigation of the inflationary universe scenario clearly indicate how many interesting and important effects can be revealed by the investigation of the cosmological consequences of elementary particle theory.

Acknowledgments

I would like to express my deep gratitude to G V Chibisov, V P Frolov, A S Goncharov, S W Hawking, R E Kallosh, M Yu Khlopov, D A Kirzhnits, M A Markov, V F Mukhanov, I L Rozental, A A Starobinsky and Ya B Zeldovich for many valuable discussions.

Appendix 1. The oscillating inflationary universe and gravitational confinement

As we have emphasised in § 5, one of the most important and intriguing cosmological problems is the problem of a general cosmological singularity. The most difficult aspect of this problem is not the existence of the singularity itself, but the question of what was *before* the singularity and what will be *after* the collapse of the universe (which will eventually occur if the universe is closed with cosmological constant $\Lambda = 0$). This problem lies somewhere at the boundary between physics and metaphysics. For example, from the point of view of an Oriental philosopher the singularity problem would not be much of a problem at all. He would recall that, according to the theory of reincarnation the 'soul complex' (or the 'principle of consciousness') survives after bodily death, and then reincarnates again in a new body, which is sufficiently suitable for it (Evans-Wentz 1980). Usually the reincarnated person does not remember his previous incarnation, at least consciously. Similarly, one could consider the universe as a 'body' for something like a world-soul, which after the collapse reincarnates inside some new universe which is sufficiently suitable for it. It may be rather difficult to reconcile such a philosophy with the usual scientific approach, but one must confess that there is a striking similarity between this philosophy and the anthropic principle, combined with the idea of an eternally oscillating (or splitting) universe considered in § 13.

One of the virtues of the inflationary universe scenario is that it provides a simple physical solution to many problems which, for a long time, have seemed almost metaphysical. Since the necessary ingredient of the inflationary universe scenario is the de Sitter stage, and de Sitter space is known to be non-singular (Hawking and Ellis 1973), it is very tempting to speculate about a possible connection between this scenario and the singularity problem. Unfortunately, it does not seem possible to solve the singularity problem in the context of the inflationary universe scenario without taking into account quantum gravity effects (Linde 1983a). The most interesting approach to the singularity problem which is known to us at present is related to what can be called 'quantum cosmology' (see § 13). However, a complete quantum description of evolution of the universe is rather complicated, especially in the absence of a consistent quantum theory of gravity. Therefore, in the meantime one may try to suggest a phenomenological description of the effects which might appear when the energy density of the universe becomes of the order of $\rho_p \sim M_p^4$ and quantum gravity effects become important. Here we would like to discuss a possibility related to the eternally oscillating universe scenario first suggested by Markov (1981, 1983, 1984).

Markov assumed that the universe is closed, and that at the moment of maximal contraction the energy–momentum tensor of matter becomes $T_{\mu\nu} \sim g_{\mu\nu}\rho_p$, so that the universe enters a de Sitter stage. Closed de Sitter space is non-singular, $a(t) = H^{-1} \cosh Ht$, where $H \sim [(8\pi/M_p^2)\rho_p]^{1/2} \sim M_p$, and after the moment of maximal contraction $t = 0$ the universe becomes expanding again. After a period of exponential expansion the de Sitter universe transforms into the hot Friedmann universe as in the inflationary universe scenario. Then the closed Friedmann universe contracts again, and again transforms into the de Sitter one, etc.

The possibility of an eternally oscillating non-singular inflationary universe scenario seems very attractive. The main problem of this scenario is that the hot Friedmann universe cannot evolve into de Sitter space without violation of the second law of thermodynamics. Indeed, the pressure p in the hot universe is given by

$$p = -F = Ts - \rho \tag{A1.1}$$

where ρ is the energy density, s is the entropy density and F is the free energy density. Therefore, in order to have $p = -\rho$ ($T_{\mu\nu} \sim g_{\mu\nu}\rho$) one should have vanishing entropy s at the end of the universe contraction. This means that, at the end of the universe contraction, the entropy of the universe should be decreasing, which would contradict the second law of thermodynamics (a similar possibility had been suggested earlier by Sakharov in his version of the oscillating universe scenario (Sakharov 1967, 1970, 1979, 1980)).

In the hot universe theory one could argue that at the temperature T exceeding $M_p \sim 10^{19}$ GeV the time remaining before the singularity is of the order of $\frac{1}{50}M_p/T^2 \ll T^{-1}$. At that time the standard thermodynamic quantities become ill-defined at the classical level, since during the time $\Delta t \ll T^{-1}$ one cannot register a particle with energy $\sim T$. In particular, one cannot measure the entropy of the universe at $T > M_p$, and it is even less clear what is the entropy of the universe in the singular state (see § 13).

However, these arguments do not apply to the non-singular oscillating universe scenario, in which we have a lot of time to take measurements.

A possible solution of this problem is connected with a dynamical violation of the second law of thermodynamics in a non-linear quantum field theory, in which there exists an upper bound on the energy density, $\rho \leqslant \rho_{\mathrm{P}}$ (Markov 1982). Here we would like to suggest one more possibility, which may help to realise the non-singular oscillating inflationary universe scenario (Linde 1983a). It is known that at $R \gtrsim G^{-1}$ perturbative results in quantum gravity become unreliable (here, R is the curvature tensor and all indices are omitted). The only other non-Abelian theory which is well-known to us is quantum chromodynamics. Perturbative results in QCD are unreliable at small momenta, which is an indication of the colour confinement. The only objects which can exist as free particles (in- and out-states) are colourless, i.e. do not interact directly with the Yang–Mills field. Quarks are assumed to interact with each other with a force which (asymptotically) does not depend on the distance between them. Therefore, any isolated quark would have an infinite energy, which just means that quarks cannot exist as free particles. Phenomenologically the existence of the distance-independent forces between two quarks can be described as the interaction between two quark currents j_1 and j_2 of the type of $j_1(k_1-k)j_2(k_2+k)/k^4$ in momentum space (though such an interpretation of the quark interactions in QCD is not quite precise). A similar term of the type $R^2(k_1-k)R^2(k_2+k)/k^4$ has been obtained by Green *et al* (1982) and Grisaru and Siegel (1982) as the one-loop correction to the effective action in O(8) supergravity. We will not speculate here about a possible phenomenological significance of this result, but just *assume* that at sufficiently large curvature $R \gtrsim G^{-1}$ in quantum gravity there occurs a phase transition to the gravitational confinement phase. We are unable to prove or disprove this conjecture at present, since it would require a much deeper understanding of quantum gravity than we have at present. Some estimates indicate that the regime of gravitational confinement may occur even at $R \ll G^{-1}$, but the confinement length in these cases is much greater than the size of the observable part of the universe. We hope to discuss this question elsewhere, and here we shall only try to understand which consequences our assumption may lead to.

First of all, let us try to understand what the hypothesis of gravitational confinement could mean. As we have already mentioned, any isolated coloured particle in QCD would have an infinite energy, and therefore only colourless particles can exist as free particles in QCD. However, *all particles are coloured with respect to the gravitational interaction.* Therefore, in the gravitational confinement phase (if such a phase can exist) *any* particle would have infinite energy. In other words, the gravitational confinement phase is a vacuum state in curved space, in which *no* free particles and, more generally, no inhomogeneities of the curvature tensor (which are also gravitationally 'coloured') can exist. Therefore, after the phase transition to the gravitational confinement phase the universe should transform into the maximally symmetric and homogeneous space with $\rho \sim \rho_{\mathrm{P}}$ and without particles, i.e. into de Sitter space.

In order to investigate the possible cosmological consequences of our hypothesis let us consider a hot closed universe contracting to a would-be singularity. The temperature of the universe increases, and at $T \gtrsim 10^{15}$ GeV the phase transitions with symmetry restoration in grand unified theories occur. When the density of matter becomes greater than $\rho_{\mathrm{P}} = G^{-2}$, the curvature tensor becomes of the order of G^{-1} and the phase transition to the gravitational confinement phase occurs. In this state no real particle excitations can exist (though virtual particles, in general, can exist) and the entropy of the universe in this state vanishes. This does not contradict the second law of thermodynamics, since any thermodynamical description of the gravitational confinement phase is impossible because of the absence of particles and of any possibility of the disorder of matter. Since the universe is closed, after the transition to the gravitational confinement phase it evolves into the closed de Sitter universe, which contracts to the minimal value of the scale factor $a \sim H^{-1}$ and then expands as $H^{-1} \cosh Ht$. Such a universe is stable with respect to the deconfining phase transition at the contraction stage but is unstable at the expansion stage. Indeed, any bubble of the non-confining phase (with $\rho \lesssim \rho_p$), which may appear inside the confining phase, at the contraction stage automatically transforms into the confining phase due to the increase in density during the contraction. On the other hand, at the exponential expansion stage sufficiently large bubbles with $\rho \lesssim \rho_p$ grow, the density inside them decreases and all the universe eventually transforms into the hot Friedmann universe filled with the ordinary (non-confining) phase with $\rho < \rho_p$. If this phase transition occurs slowly enough, the inflationary universe scenario can be realised at this stage of expansion of the universe. The inflationary universe scenario can also be realised later (see §§ 10–12). After inflation the closed Friedmann universe expands to some large radius, contracts and again transforms into the de Sitter universe in the confinement phase.

We have suggested, therefore, a possible realisation of the Markov non-singular oscillating inflationary universe scenario. Since the duration of the de Sitter stage fluctuates from cycle to cycle, there will be large fluctuations in the duration of each cycle, but the average duration will be time-independent. This is one of the most important differences between the scenario discussed above and the usual oscillating universe model, in which the duration of each cycle grows infinitely with the number of cycles due to the growth of the total entropy of the universe in each cycle (see, for example, Zeldovich and Novikov 1975). In our case the universe does not remember anything about its previous 'incarnations'. All entropy and all inhomogeneities of the contracting universe disappear in the purgatory of the de Sitter stage and are then generated anew in each new cycle.

Our discussion of the gravitational confinement hypothesis and of its cosmological consequences is, of course, far from being rigorous. It is also quite possible that this hypothesis is not necessary for the realisation of the oscillating inflationary universe scenario (see, for example, Markov (1981, 1983, 1984), Hawking (1984d) and § 13). However, the problem we are discussing now is extremely complicated and, just as in the theory of quark confinement, in the absence of a complete theory it might be useful to have several different models which may describe different aspects of the same physical reality.

Appendix 2. The cosmological constant problem

As was pointed out in § 5, one of the most difficult problems related both to cosmology and to elementary particle physics is the cosmological constant problem. There have been many attempts made to understand why the vacuum energy at present is almost precisely zero (Dolgov 1983, Zee 1983, Rajeev 1983, Rubakov and Shaposhnikov 1983, Cremmer *et al* 1983, Hawking 1984a, b), but in our opinion the problem is still not solved.

In this appendix we would like to discuss two different possibilities to solve the cosmological constant problem. The first possibility is connected with a doubling of all matter fields (Linde 1984a). Let us assume that the total Lagrangian of matter can be written as follows:

$$\mathscr{L} = -\frac{R}{16\pi G} + L(\varphi) - L(\tilde{\varphi}) \qquad \text{(A2.1)}$$

where $-R/16\pi G$ is the Einstein Lagrangian, $L(\varphi)$ is the Lagrangian of matter fields φ, $-L(\tilde{\varphi})$ is the Lagrangian of the same functional form as $L(\varphi)$ but is a function of the other fields $\tilde{\varphi}$ and it enters into $\mathscr{L}$ (A2.1) with a *negative sign.*

At first sight, the theory (A2.1) is unstable with respect to the generation of infinitely large fields due to the negative sign of $L(\tilde{\varphi})$. However, all Lagrange equations for the field $\tilde{\varphi}$ are *the same* as for the field φ and there is *no* instability in the classical theory $-L(\tilde{\varphi})$. The only difference between these theories is that the field φ in the theory $L(\varphi)$ will be in the minimum φ_0 of the effective potential $V(\varphi)$, whereas the field $\tilde{\varphi}$ will be in the *maximum* (at $\tilde{\varphi} = \tilde{\varphi}_0$) of the effective potential $V(\tilde{\varphi})$. Since the Lagrangians $L(\varphi)$ and $L(\tilde{\varphi})$ are of the same functional form, $\varphi_0 = \tilde{\varphi}_0$ and $V(\varphi_0) = -V(\tilde{\varphi}_0)$. Therefore, the total vacuum energy density of the theory (A2.1) will be zero *independently of the value of* $V(\varphi_0)$:

$$\rho_{\text{vac}} = V(\varphi_0) + V(\tilde{\varphi}_0) = 0. \qquad \text{(A2.2)}$$

This is exactly what is needed for a solution of the cosmological constant problem.

One may wonder whether the new particles $\tilde{\varphi}$, which we call 'down' particles to distinguish them from the usual 'up' particles φ, are harmless. To answer this question let us note that 'up' and 'down' particles interact with each other only gravitationally. Therefore, only very large inhomogeneously distributed amounts of 'down' matter (like planets) could affect us in some way. Moreover, in the inflationary universe scenario one could see no 'down' particles at all. Indeed, let us consider a domain of the field $\varphi \gtrsim 5M_{\rm p}$ in which the field $\tilde{\varphi}$ was initially much smaller than φ. Such a domain will exponentially expand (see § 11). If $\tilde{\varphi} \ll \varphi$, then the process of $\tilde{\varphi}$-particle production finishes *before* the end of inflation, and then inflation pushes all 'down' particles away from the observable part of the universe.

There still exist many problems associated with our suggestion. First of all, there is the vacuum instability due to gravitational interaction between 'up' and 'down' particles. Indeed, a pair of 'down' and a pair of 'up' particles can be produced from the vacuum without violation of energy conservation. However, this process does not produce any energy density and does not lead to the creation of a preferred reference frame; this is just a vacuum reconstruction process. A preliminary investigation of this question indicates that such a boiling of the vacuum state may not lead to any unacceptable observational consequences.

Another problem may be even more difficult. It is possible to double the matter fields φ, but unless gravity is doubled as well, quantum gravity corrections to $V(\varphi)$ and $V(\tilde{\varphi})$ can violate equation (A2.2), though it might be possible to make quantum gravity corrections to $V(\varphi)$ and $V(\tilde{\varphi})$ small if these potentials are sufficiently flat near their minima.

We realise that the possibility of having 'up' and 'down' worlds may sound absolutely crazy, but at least at the level at which quantum gravity effects can be neglected we see no reasons why such a scenario could not work. Quantum gravity effects may lead to considerable complications of this scenario, but the general idea of the vacuum energy cancellation between the 'up' and 'down' vacua seems so simple and natural that it would be a pity to abandon such a possibility without a detailed investigation.

Another possible way of solving the cosmological constant problem is related to quantum cosmology. The vacuum energy density may depend on the topology of the compactified part of space (Sakharov 1984) and on some classical fields of the type of the antisymmetric tensor field $A_{\mu\nu\lambda}$ (Ogievetsky and Sokatchev 1980, Duff and van Nieuwenhuizen 1980, Aurelia *et al* 1980). This field, just as the scalar field φ, appears simultaneously with quantum creation of the universe (Hawking 1984a, b). However, contrary to the scalar field φ, the field strength $F_{\mu\nu\lambda\sigma}$ of the field $A_{\mu\nu\lambda}$, which gives the contribution $V(F)$ to the vacuum energy $V(\varphi, F) = V(\varphi) + V(F)$, remains constant during the subsequent classical evolution of the universe (Ogievetsky and Sokatchev 1980, Duff and van Nieuwenhuizen 1980, Aurelia *et al* 1980). As follows from equation (13.7), the universe is created most probably in a state with $V(\varphi, F) \gtrsim M_{\rm p}^4$ (Linde 1984c, d, Starobinsky 1984b). However, this does not impose any constraints on the value of $V(F) = V(\varphi, F) - V(\varphi)$ since the value of $V(\varphi)$ at the initial stages of inflation can be arbitrarily large (see § 11). Therefore, *after* symmetry breaking any value of vacuum energy density $V(\varphi_0, F) = V(\varphi_0) + V(F)$ may appear with approximately the same probability. At $|V(\varphi_0, F)| \gg 10^{-29}$ g cm^{-3} life of our type would be impossible. The value $|V(\varphi_0, F)| \lesssim 10^{-29}$ g cm^{-3} *a priori* does not seem very probable. The eternally oscillating universe scenario is not of much help here, since the closed universe with $V(\varphi_0, F) > 0$ can expand forever, which would break the chain of oscillations (see § 13). What may occur, however, is a multiple quantum production of 'new' universes from the 'old' ones. Such a process looks like an infinite chain reaction, which is possible due to the gravitational instability discussed in § 13 (see also a paper by Englert and Nicolai (1983) in which similar ideas were suggested). During this process infinitely many universes can be produced, in some of which $|V(\varphi_0, F)| \lesssim 10^{-29}$ g cm^{-3} and life of our type may exist. This is a possible solution of the cosmological constant problem based on the implementation of the anthropic principle in quantum cosmology.

References

Abbott L F and Burges C J C 1983 *Phys. Lett.* **131B** 49–52

Abbott L F, Fahri E and Wise M B 1982 *Phys. Lett.* **117B** 29–34

Abers G S and Lee B W 1973 *Phys. Rep.* **9** 1–141

Albrecht A, Dimopoulos S, Fischler W, Kolb E W, Raby S and Steinhardt P J 1983 *Nucl. Phys.* B **229** 528–40

Albrecht A and Steinhardt P J 1982 *Phys. Rev. Lett.* **48** 1220–3

Albrecht A, Steinhardt P J, Turner M S and Wilczek F 1982 *Phys. Rev. Lett.* **48** 1437–40

Altshuller B L 1972 *Proc. 3rd Soviet Gravitational Conf., Erevan* p 6

Aurelia A, Nicolai H and Townsend P K 1980 *Nucl. Phys.* B **176** 509–22

Barbieri R, Ferrara S and Savoy C A 1982 *Phys. Lett.* **119B** 343–7

Bardeen J, Steinhardt P J and Turner M S 1983 *Phys. Rev.* D **28** 679–93

Barrow J 1983 *Fundamentals of Cosmic Phys.* **7** 1–117

Billoire A and Tamvakis K 1982 *Nucl. Phys.* B **200** 329–44

Birrell N D and Davies P C W 1982 *Quantum Fields in Curved Space* (Cambridge: Cambridge University Press)

Boucher W 1983 *Preprint* Department of Applied Mathematics and Theoretical Physics, University of Cambridge

Boucher W and Gibbons G W 1983 *The Very Early Universe* ed G W Gibbons, S W Hawking and S Siklos (Cambridge: Cambridge University Press) pp 273–8

Brandenberger R and Kahn R 1984 *Phys. Rev.* D to be published

Brout R, Englert F and Gunzig E 1978 *Ann. Phys., NY* **115** 78–106

Brout R, Englert F and Spindel P 1979 *Phys. Rev. Lett.* **43** 417–20

Buccella F, Deredinger J P, Ferrara S and Savoy C A 1982 *Phys. Lett.* **115B** 375–9
Bunch T S and Davies P C W 1978 *Proc. R. Soc.* A **360** 117–34
Carr B J and Rees M J 1979 *Nature* **278** 605–12
Ceccarelli C, Melchiorri F, Pietraneira L, Dall' Oglio G and Melchiorri B O 1983 *Astrophys. J. Lett.* **269** L27–30
Coleman S 1977 *Phys. Rev.* D **15** 2929–36
Coleman S and De Luccia F 1980 *Phys. Rev.* D **21** 3305–15
Coleman S and Weinberg E 1973 *Phys. Rev.* D **7** 1888–910
Collins C B and Hawking S W 1973 *Astrophys. J.* **180** 317–34
Cremmer E, Ferrara S, Kounnas C and Nanopoulos D V 1983 *Phys. Lett.* **133B** 61
Cremmer E and Sherk J 1976 *Nucl. Phys.* B **108** 409–16
Cremmer E, Julia B, Scherk J, Ferrara S, Girardello L and van Nieuwenhuizen P 1979 *Nucl. Phys.* B **147** 105–31
Daniel M 1981 *Phys. Lett.* **98B** 391–4
Deser S and Zumino B 1976 *Phys. Lett.* **62B** 335–8
DeWitt B S 1967 *Phys. Rev.* **160** 1113–48
Dicke R H 1961 *Nature* **192** 440–1
Dimopoulos S and Georgi H 1981 *Nucl. Phys.* B **193** 150–62
Dimopoulos S and Raby S 1983 *Nucl. Phys.* B **221** 479–512
Dimopoulos S and Susskind L 1978 *Phys. Rev.* D **18** 4500–9
Dolan L and Jackiw R 1974 *Phys. Rev.* D **9** 3320–41
Dolgov A D 1983 *The Very Early Universe* ed G W Gibbons, S W Hawking and S Siklos (Cambridge: Cambridge University Press) pp 449–58
Dolgov A D and Linde A D 1982 *Phys. Lett.* **116B** 329–34
Dolgov A D and Zeldovich Ya B 1981 *Rev. Mod. Phys.* **53** 1–41
Doroshkevich A G and Khlopov M Yu 1984a *Yad. Fiz.* **39** 808
—— 1984b *Mon. Not. R. Astron. Soc.* to be published
Dowker J S and Critchley R 1976 *Phys. Rev.* D **13** 3224–32
Dragon N V 1982 *Phys. Lett.* **113B** 288–92
Dreitlein J 1974 *Phys. Rev. Lett.* **33** 1243–4
Duff M J and van Nieuwenhuizen P 1980 *Phys. Lett.* **94B** 179–82
Ehrenfest P 1917 *Proc. Amsterdam Acad.* **20** 200
Ellis J, Gaillard M K and Nanopoulos D V 1979 *Phys. Lett.* **80B** 360
Ellis J, Linde A D and Nanopoulos D V 1982 *Phys. Lett.* **118B** 59–63
Ellis J and Nanopoulos D V 1982 *Phys. Lett.* **116B** 133–6
Ellis J, Nanopoulos D V, Olive K A and Tamvakis K 1983a *Nucl. Phys.* B **221** 524–48
—— 1983b *Phys. Lett.* **120B** 331–4
Englert F and Brout R 1964 *Phys. Rev. Lett.* **13** 321
Englert F and Nicolai H 1983 *CERN Preprint* TH 3711
Evans-Wentz W Y 1980 *The Tibetan Book of the Dead* ed W Y Evans-Wentz (Oxford: Oxford University Press) pp 1–81
Everett H 1957 *Rev. Mod. Phys.* **29** 454
Fayet P and Ferrara S 1977 *Phys. Rep.* **32** 249
Fixsen D J, Cheng E S and Wilkinson D T 1983 *Phys. Rev. Lett.* **50** 620–2
Fomin P I 1973 *Preprint* Kiev ITF-73-137 P
—— 1975 *Dokl. Akad. Nauk Ukr. SSR* **9A** 831–5
Fradkin E S 1980 *Proc. Seminar 'Quark 80'* (Moscow: I Ya I Press) pp 80–98
Frampton P H and Kephart T W 1982 *Phys. Rev. Lett.* **48** 1237–41
Freedman D Z, van Nieuwenhuizen P and Ferrara S 1976 *Phys. Rev.* D **13** 3214–8
Freund P G O 1982 *Nucl. Phys.* B **209** 146
Freund P G O and Rubin M A 1980 *Phys. Lett.* **97B** 233–5
Gell-Mann M, Ramond P and Slansky R 1978 *Rev. Mod. Phys.* **50** 721–44
Gelmini G B, Kounnas C and Nanopoulos D V 1984 *Primordial inflation with flat potentials. CERN Preprint* TH 3777
Gelmini G B, Nanopoulos D V and Olive K A 1983 *Phys. Lett.* **131B** 53–6
Georgi H and Glashow S 1974 *Phys. Rev. Lett.* **32** 438–40
Gibbons G W and Hawking S W 1977 *Phys. Rev.* D **15** 2738–51
Gibbons G W, Hawking S W and Perry M 1978 *Nucl. Phys.* B **138** 141
Gildener E 1976 *Phys. Rev.* D **14** 1667–72
Glashow S L 1961 *Nucl. Phys.* **22** 579

Gliner E B 1965 *Sov. Phys.–JETP* **22** 378–82
—— 1970 *Dokl. Akad. Nauk SSSR* **192** 771–4
Gliner E B and Dyminikova I G 1975 *Pis. Astron. Zh.* **1** No 5 7–9
Gol'fand Ya A and Likhtman E P 1971 *JETP Lett.* **13** 323
Goncharov A S and Linde A D 1984a *Zh. Eksp. Teor. Fiz.* **86** 1594–9
—— 1984b *Phys. Lett.* **139B** 27–30
—— 1984c *Elementary Particles and Atomic Nuclei* submitted
—— 1984d *A simple realisation of the inflationary universe scenario in SU(1, 1) supergravity. Preprint* Lebedev Physical Institute
Green M B, Schwarz J H and Brink L 1982 *Nucl. Phys.* B **198** 474–92
Grisaru M T and Siegel W 1982 *Phys. Lett.* **110B** 49–54
Grishchuk L P and Zeldovich Ya B 1982 *Quantum Structure of Space–Time* ed M Duff and C Isham (Cambridge: Cambridge University Press) pp 409–22
Gross D J and Wilczek F 1973 *Phys. Rev. Lett.* **30** 1343–6
Guralnik G S, Hagen C R and Kibble T W B 1964 *Phys. Rev. Lett.* **13** 585
Gurevich L E 1975 *Astrophys. Space Sci.* **38** 67–78
Gurovich V Ts and Starobinsky A A 1979 *Sov. Phys.–JETP* **50** 844
Guth A H 1981 *Phys. Rev.* D **23** 347–56
—— 1983 *The Very Early Universe* ed G W Gibbons, S W Hawking and S Siklos (Cambridge: Cambridge University Press) pp 171–204
Guth A H and Pi S-Y 1982 *Phys. Rev. Lett.* **49** 1110–3
Guth A H and Weinberg E 1983 *Nucl. Phys.* B **212** 321–64
Hartle J B and Hawking S W 1983 *Phys. Rev.* D **28** 2960–75
Hawking S W 1978 *Nucl. Phys.* B **144** 349–62
—— 1982 *Phys. Lett.* **115B** 295–7
—— 1984a *Phil. Trans. R. Soc.* A to be published
—— 1984b *Phys. Lett.* **134B** 4034
—— 1984c *Quantum Cosmology. Les Houches Lectures* (Amsterdam: North-Holland) to be published
—— 1984d *Nucl. Phys.* to be published
Hawking S W and Ellis G F R 1973 *The Large Scale Structure of the Universe* (Cambridge: Cambridge University Press)
Hawking S W, Moss I G and Stewart J M 1982 *Phys. Rev.* D **26** 2681–93
Hawking S W and Moss I G 1982 *Phys. Lett.* **110B** 35–8
—— 1983 *Nucl. Phys.* B **224** 180–92
Higgs P W 1964a *Phys. Lett.* **12** 132
—— 1964b *Phys. Rev. Lett.* **13** 508
—— 1966 *Phys. Rev.* **145** 1156
't Hooft G 1974 *Nucl. Phys.* B **79** 279–84
Ibanez L E 1982 *Phys. Lett.* **118B** 73
Ignatiev A Yu, Krasnikov N V, Kuzmin V A and Tavkhelidze A N 1978 *Phys. Lett.* **76B** 436–8
Kaluza Th 1921 *Sitzungsber. Preuss. Akad. Wiss. Berlin, Math: Phys.* **K1** 966
Khlopov M Yu and Linde A D 1984 *Phys. Lett.* **138B** 265–8
Kibble T W B 1967 *Phys. Rev.* **155** 1554
—— 1976 *J. Phys. A: Math. Gen.* **9** 1387–98
—— 1980 *Phys. Rep.* **67** 183–99
Kirzhnits D A 1972 *JETP Lett.* **15** 529–31
Kirzhnits D A and Linde A D 1972 *Phys. Lett.* **42B** 471–4
—— 1974 *Zh. Eksp. Teor. Fiz.* **67** 1263–75 (Engl. transl. 1975 *Sov. Phys.–JETP* **40** 628–34)
—— 1976 *Ann. Phys., NY* **101** 195–238
—— 1982 *Science and Mankind* (Moscow: Znanie) pp 165–77
Klein O 1926 *Z. Phys.* **37** 895
Kofman L A, Linde A D and Starobinsky A A 1984 *Phys. Lett.* submitted
Kolb E W and Wolfram S 1980 *Astrophys. J.* **239** 428
Kuzmin V A 1970 *JETP Lett.* **12** 335
Landau L D and Lifshitz E M 1973 *Field Theory* (Moscow: Nauka)
—— 1976 *Statistical Physics* (Moscow: Nauka)
Langacker P 1981 *Phys. Rep.* **68** 185
Lazarides G, Shafi Q and Walsh T F 1982 *Nucl. Phys.* B **195** 157–72

Linde A D 1974 *JETP Lett.* **19** 183–4
—— 1976 *JETP Lett.* **23** 64–7
—— 1977 *Phys. Lett.* **70B** 306–8
—— 1979 *Rep. Prog. Phys.* **42** 389–437
—— 1980a *Phys. Lett.* **92B** 119–22
—— 1980b *Preprint* Bielefeld University BI-TP 80/20
—— 1981a *Phys. Lett.* **99B** 391–5
—— 1981b *Phys. Lett.* **100B** 37–40
—— 1982a *Phys. Lett.* **108B** 389–93
—— 1982b *Phys. Lett.* **114B** 431–5
—— 1982c *Phys. Lett.* **116B** 340–2
—— 1982d *Phys. Lett.* **116B** 335–9
—— 1983a *The Very Early Universe* ed G W Gibbons, S W Hawking and S Siklos (Cambridge: Cambridge University Press) pp 205–49
—— 1983b *Phys. Lett.* **123B** 185–7
—— 1983c *JETP Lett.* **37** 724–7
—— 1983d *JETP Lett.* **38** 176–8
—— 1983e *Phys. Lett.* **129B** 177–80
—— 1983f *Phys. Lett.* **131B** 330–3
—— 1983g *Phys. Lett.* **132B** 317–20
—— 1983h *Nucl. Phys.* B **216** 421–45
—— 1984a *Proc. Shelter Island Conf. II* (Cambridge, Mass.: MIT Press) to be published
—— 1984b *Usp. Fiz. Nauk* **144** to be published
—— 1984c *Lett Nuovo Cim.* **39** 401–5
—— 1984d *Zh. Eksp. Teor. Fiz.* **87** to be published
Lubimov V A *et al* 1980 *Phys. Lett.* **94B** 266
—— 1983 *Conf. on High Energy Physics, Brighton* Talk
Markov M A 1981 *Preprint* Institute of Nuclear Research P-0227
—— 1982 *JETP Lett.* **36** 265–7
—— 1983 *The Very Early Universe* ed G W Gibbons, S W Hawking and S Siklos (Cambridge: Cambridge University Press) pp 353–72
—— 1984 to be published
Misner C W 1969a *Phys. Rev. Lett.* **22** 1071
—— 1969b *Phys. Rev.* **186** 1319, 1328
Moss I G and Wright W A 1983 *Preprint* University of Newcastle
Mottola E and Lapedes A 1983 *Phys. Rev.* D **27** 2285
Mukhanov V F and Chibisov G V 1981 *JETP Lett.* **33** 532–5
—— 1982 *Sov. Phys.–JETP* **56** 258–65
—— 1984 *Pis. Astron. Zh.* to be published
Nanopoulos D V, Olive K A and Srednicki M 1983a *Phys. Lett.* **127B** 30–4
Nanopoulos D V, Olive K A, Srednicki M and Tamvakis K 1983b *Phys. Lett.* **123B** 41
Nanopoulos D V, Olive K A and Tamvakis K 1982 *Phys. Lett.* **115B** 15
Nanopoulos D V and Tamvakis K 1982 *Phys. Lett.* **110B** 449–52
Nath P, Arnowitt R and Chamseddine A H 1982 *Phys. Rev. Lett.* **49** 970
van Nieuwenhuizen P 1981 *Phys. Rep.* **68** 191–397
Ogievetsky V I and Sokatchev E 1980 *Yad. Fiz.* **32** 1142–51
Okun' L B 1982 *Leptons and Quarks* (Amsterdam: North-Holland)
Olive K and Seckel D 1983 *Preprint. Fermilab-Conf.* 83/105-AST
Ovrut B A and Steinhardt P J 1983 *Phys. Lett.* **133B** 161
Parke S 1983 *Phys. Lett.* **121B** 313–5
Parke S and Pi S-Y 1981 *Phys. Lett.* **107B** 54–8
Pi S-Y 1984 *Inflation without tears. Preprint* Boston University
Politzer H D 1973 *Phys. Rev. Lett.* **30** 1346–9
Polyakov A M 1974 *JETP Lett.* **20** 194
Preskill J P 1979 *Phys. Rev. Lett.* **43** 1365–8
Rajeev S G 1983 *Phys. Lett.* **125B** 144–6
Rees M J 1972 *Phys. Rev. Lett.* **28** 1071
Rindler W 1956 *Mon. Not. R. Astron. Soc.* **116** 663

Rozental I L 1980 *Usp. Fiz. Nauk* **131** 239–56
—— 1984 *Elementary Particles and Structure of the Universe* (Moscow: Nauka)
Rubakov V A, Sazhin M V and Veryaskin A V 1982 *Phys. Lett.* **115B** 189–92
Rubakov V A and Shaposhnikov M E 1983 *Phys. Lett.* **125B** 136–8
Sakharov A D 1965 *Sov. Phys.–JETP* **22** 241–9
—— 1967 *JETP Lett.* **5** 24–7
—— 1970 *Preprint* Institute of Applied Mathematics
—— 1979 *Sov. Phys.–JETP* **49** 594–9
—— 1980 *Sov. Phys.–JETP* **52** 349–51
—— 1984 *Zh. Eksp. Teor. Fiz.* **87** to be published
Sakai N 1982 *Z. Phys.* C **11** 153
Salam A 1968 *Elementary Particle Theory* ed N Svartholm (Stockholm: Almquist and Wiksell)
Sato K, Kodama H, Sasaki M and Maeda K 1982 *Phys. Lett.* **108B** 103–7
Shandarin S F, Doroshkevich A G and Zeldovich Ya B 1983 *Usp. Fiz. Nauk* **139** 83
Shafi Q and Vilenkin A 1984 *Phys. Rev. Lett.* **52** 691–4
Shafi Q and Wetterich C 1983 *Phys. Lett.* **129B** 387–90
Sher M 1982 *Phys. Rev.* D **24** 1699–701
Sikivie P 1982 *Phys. Rev. Lett.* **48** 1156–8
Slansky R 1981 *Phys. Rep.* **79** 1–128
Srednicki M 1982a *Nucl. Phys.* B **202** 327
—— 1982b *Nucl. Phys.* B **206** 132
Starobinsky A A 1979 *JETP Lett.* **30** 682–5
—— 1980 *Phys. Lett.* **91B** 99–102
—— 1982 *Phys. Lett.* **117B** 175–8
—— 1983a *JETP Lett.* **37** 66–9
—— 1983b *Pis. Astrofiz. Zh.* **9** 579–83
—— 1983c *Quantum Gravity* (New York: Plenum) pp 103–29
—— 1984a *Fundamental Interactions* (Moscow: MGPI Press)
—— 1984b to be published
—— 1984c *Zh. Eksp. Teor. Fiz.* to be published
Steigman G and Turner M S 1983 *Phys. Lett.* **128B** 295
Taylor J C 1976 *Gauge Theories of Weak Interactions* (Cambridge: Cambridge University Press)
Toussaint D, Treiman S B, Wilczek F and Zee A 1979 *Phys. Rev.* D **19** 1036–45
Tryon E P 1973 *Nature* **246** 396–7
Turner M S, Steigman G and Krauss L M 1984 *Preprint* Fermilab-Publ 84/32-A
Veltman M 1974 *Preprint* Rockefeller University
—— 1975 *Phys. Rev. Lett.* **34** 77
Vilenkin A 1981 *Phys. Rev. Lett.* **46** 1169–72
—— 1982 *Phys. Lett.* **117B** 25–32
—— 1983a *Phys. Rev.* D **27** 2848–55
—— 1983b *Nucl. Phys.* B **226** 527–46
Vilenkin A and Ford L H 1982 *Phys. Rev.* D **26** 1231–41
Vilenkin A and Shafi Q 1983 *Phys. Rev. Lett.* **51** 1716–9
Volkov D V and Akulov V P 1972 *JETP Lett.* **16** 438–41
Voloshin M B, Kobzarev I Yu and Okun' L B 1974 *Yad. Fiz.* **20** 1229–34
Wald R 1983 *Phys. Rev.* D **28** 2118–20
Weinberg S 1967 *Phys. Rev. Lett.* **19** 1264–6
—— 1972 *Gravitation and Cosmology* (New York: Wiley)
—— 1974 *Phys. Rev.* D **9** 3357–78
—— 1976 *Phys. Rev. Lett.* **36** 294–6
—— 1979 *Phys. Rev. Lett.* **42** 850–3
—— 1982a *Phys. Rev. Lett.* **48** 1303–6
—— 1982b *Phys. Rev. Lett.* **48** 1776–9
Wess J and Zumino B 1974 *Nucl. Phys.* B **70** 39–49
Wheeler J A 1964 *Relativity, Groups and Topology* ed B S DeWitt and C M DeWitt (New York: Gordon and Breach)
—— 1968 *Batelle Rencontres* ed C M DeWitt and J A Wheeler (New York: Benjamin)
Witten E 1981 *Nucl. Phys.* B **186** 412–28

Yoshimura M 1978 *Phys. Rev. Lett.* **41** 281–4
Zee A 1980 *Preprint* University of Pennsylvania UPR-0151T
—— 1983 *Preprint* University of Washington
Zeldovich Ya B 1972 *Mon. Not. R. Astron. Soc.* **160** 1p
—— 1980 *Mon. Not. R. Astron. Soc.* **192** 663–7
—— 1981 *Pis. Astron. Zh.* **7** 579–82
Zeldovich Ya B and Khlopov M Yu 1978 *Phys. Lett.* **79B** 239–41
Zeldovich Ya B, Kobzarev M Yu and Okun' L B 1974 *Phys. Lett.* **50B** 340–3
Zeldovich Ya B and Novikov I D 1975 *Structure and Evolution of the Universe* (Moscow: Nauka) (Engl. transl. 1983 *Relativistic Astrophysics* vol II (Chicago: University of Chicago Press))