Particle physics and inflationary cosmology

It seems likely that the universe is an eternal, self-reproducing entity divided into many mini-universes, with low-energy physics and perhaps even dimensionality differing from one to the other.

Andrei Linde

With the invention of unified theories of strong, weak, electromagnetic and gravitational interactions, elementary-particle physics has entered a very interesting and unusual stage of its development. The end of the 1960s saw the introduction of the Glashow–Weinberg–Salam unification of the weak and electromagnetic interactions. In 1974 came the grand unified theories of the strong, weak and electromagnetic interactions. Two years later we had supergravity, giving us the first hope of unifying all fundamental interactions, including gravitation. The beginning of the 1980s witnessed a renewal of interest in the Kaluza–Klein theories and supergravity in higher-dimensional space–time. Nowadays superstring theory is the leading candidate for the role of "theory of everything."

The 1983 discovery of the W and Z weak vector bosons predicted by the Glashow–Weinberg–Salam theory has shown that the unified theories have made a very good start. Optimists are full of enthusiasm, predicting that at this rate we will soon reach our final goal. It would be highly desirable, however, to verify experimentally that we are still going in the right direction. Unfortunately, this appears to be an extremely difficult task.

The energy scale at which the unified nature of all four fundamental interactions is expected to become manifest is not very different from the Planck mass M_P, about 10^{-5} grams, where quantum gravity effects become important. (The Planck mass is that mass for which the Compton wavelength l_P, about 10^{-33} cm, equals the Schwarzschild radius.) Its rest energy $M_P c^2$, about 10^{19} GeV, corresponds to the kinetic energy of a small airplane. By contrast, the 80-km-circumference Superconducting Super Collider the Americans hope to build in the near future will accelerate particles up to 10^4 GeV. The largest accelerator ring one could build on Earth, with a circumference of 40 000 km, could not accelerate particles beyond about 10^8 GeV, a center-of-mass energy occasionally to be seen in cosmic-ray collisions. But this still leaves us 12 orders of magnitude short of the energy necessary for a direct test of the unified theories. Of course there are some indirect tests, such as the searches for proton decay and for supersymmetric partners of ordinary particles. But trying to get a correct theory of all fundamental interactions with only such low-energy experiments is like looking for the correct unified electroweak theory by studying nothing but radio waves.

The universe as accelerator

The only accelerator that could ever produce particles energetic enough for a direct testing of the unified theories of all fundamental interactions is our universe itself. The Big Bang scenario as it stood ten years ago,[1] which I will call the hot-universe theory, asserts that the universe was born at some moment $t=0$ about 15 billion years ago, in a state of infinitely high temperature T and infinitely large density ρ. Of course one cannot really speak of classical space–time for the earliest moments, when kT/c^2 was greater than the Planck mass, and ρ exceeded $M_P / l_P{}^3$, making quantum fluctuations of the metric predominant. (We will use a convenient unit system that sets k, c and h all equal to 1, so that l_P equals $1/M_P$ and the Planck density is $M_P{}^4$, roughly 10^{94} g/cm^3.) It is just at such times, when the average particle energy exceeded M_P, that the unity of all four fundamental interactions would have been manifest.

With the rapid expansion of the universe, the average energy of particles, given by the temperature, decreases rapidly, and the universe becomes cold. The temperature falls as the reciprocal of R, the scale factor, or "radius," of the universe. This means that particle interactions at extremely large energies can have occurred only at the very early stages of universal evolution. One might think it very difficult to extract useful and reliable information from the unique experiment carried out about 10^{10} years ago. Thus it came as a great surprise to those who study elementary particles that the investigation of physical processes at the very early stages of the universe can rule out most of the existing unified theories.

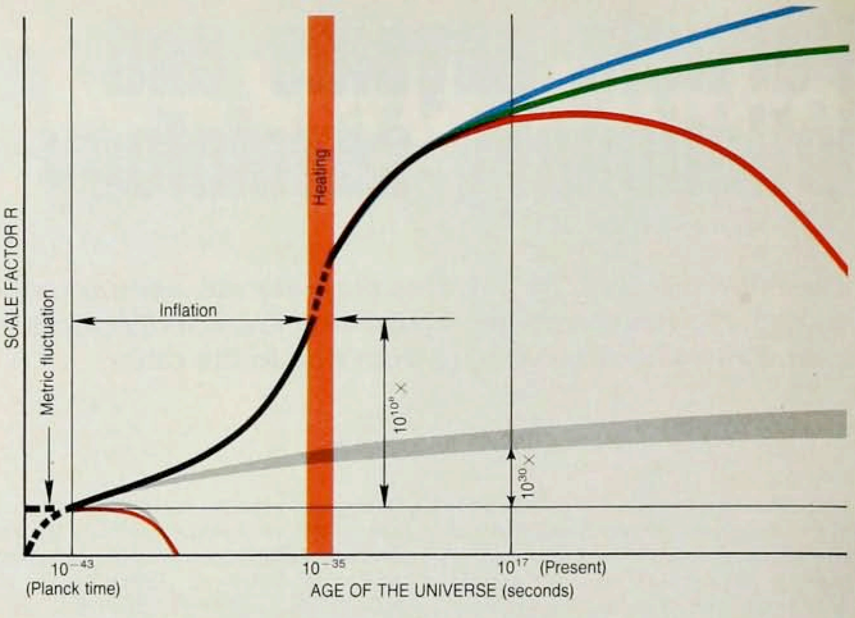
For example, all the grand unified theories predict the existence of superheavy stable particles carrying magnetic charge: magnetic monopoles. These objects have a typical mass 10^{16} times that of the proton. According to the standard hot-universe theory, monopoles should appear at the very early stages of the universe, and they should now be as abundant as protons. In that case the mean density of matter in the universe would be about 15 orders of magnitude higher than its present value of about 10^{-29} g/cm^3.

Originally there was some hope that this problem would disappear when more complicated theories were considered. It turned out, however, that the problems became even more complicated in the new theories. For example, according to the models based on one version of supergravity, the universe should contain not only monopoles but also heavy gravitinos (superpartners of the graviton, with spin $3/2$) and oscillating classical "Polonyi" fields. In most of these theories the predicted abundance of gravitinos (or their decay products) contradicts cosmological data by about 10 orders of magnitude, and the typical energy density stored in the Polonyi fields contradicts the data by 15 orders of magnitude. Most of the higher-dimensional Kaluza–Klein the-

Andrei Linde is professor of physics at the Lebedev Physical Institute in Moscow.

ories considered in the early 1980s predict the present vacuum energy density to be on the order of M_P^4, too large for the observational data by 125 orders of magnitude. The situation with the currently fashionable superstring theories (PHYSICS TODAY, July 1985, page 17) seemed, at first glance, to be somewhat better. To be fair, however, one must say that no consistent cosmological model based on superstrings has been suggested thus far.

We see that it is not difficult to obtain strong cosmological constraints on the unified particle theories. On the contrary, there is a problem as to whether it is possible to reconcile elementary-particle theory and cosmology. To answer this question it is necessary to check whether the standard hot-universe theory is as good as it seemed at first to be, and whether it is possible to modify it so as to remove some of the difficulties we have mentioned.



Expansion of the universe for three possible inflationary scenarios—open (blue), flat (green) and closed (red). The scale factor R, which, at least in the closed scenario, can be thought of as the radius of the universe, is plotted as a function of time. The inflationary universe starts either from a singular Big Bang or from a large quantum fluctuation of a pre-existing space–time metric. After the Planck time, inflation sets in, expanding the universe by a factor of 10^{10^8} in 10^{-35} seconds. By contrast, the noninflationary hot-universe model (gray) yields an expansion factor of only 10^{30} to the present day, 10^{17} seconds after the beginning. The inflationary epoch is followed, at 10^{-35} seconds, by a brief interlude of heating, after which the further evolution of the universe is adequately described by the standard noninflationary model. Figure 1

Problems of the standard theory

There are many difficult problems associated with the hot-universe theory, but for a long time most of them seemed more metaphysical than physical, and thus of no immediate concern for scientists. The most important of these problems are:

▶ The singularity problem. The state of infinite density in which the universe was born at time 0 is called a singularity. One may wonder what there was *before* the Big Bang singularity: Where did the universe come from? If space–time itself does not exist for times less than 0, how could everything appear from nothing?

▶ The flatness problem. At school we are taught that the world is described by Euclidean geometry, and that two parallel lines never cross. When we come to the university, our professors tell us that, according to general relativity, the geometry of the universe is, in general, different from the Euclidean geometry of flat space. The universe may be open, in which case parallel lines diverge from one another, or it may be closed, in the way that the surface of a sphere is closed, so that parallel lines cross one another like the meridian lines on a globe.

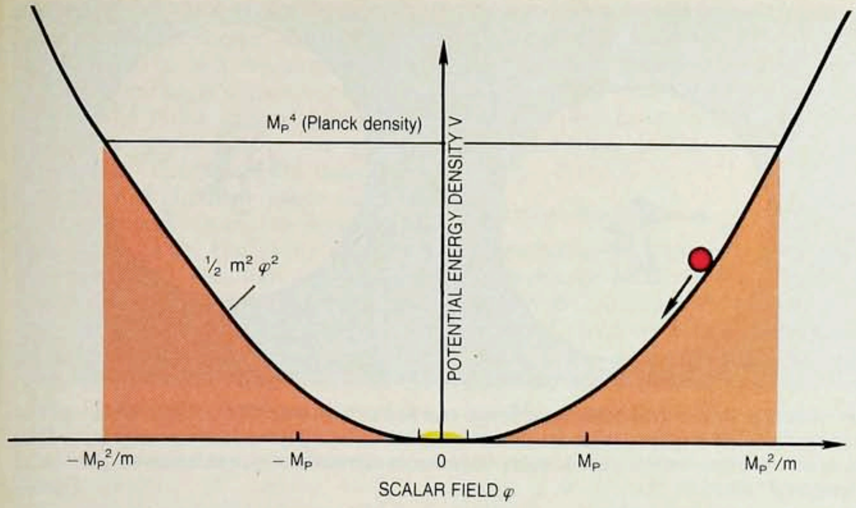
The only natural length parameter in general relativity is the Planck length l_P. Therefore one would expect our space to be very curved, with a typical radius of curvature on the order of 10^{-33} cm. We see, however, that our universe is just about flat on a scale of 10^{28} cm, the radius of the observable part of the universe. This 10^{10}-light-year distance to the Big Bang horizon is 60 orders of magnitude larger than l_P. Why is our universe so flat, its geometry almost exactly Euclidean?

▶ The horizon problem. This problem can be formulated in several ways. The simplest way to understand it is to consider an infinite (open or flat) universe. Why, the problem asks, did all the causally disconnected regions of the infinite universe start their expansion simultaneously (at $t = 0$)? Who gave the command?

▶ The problems of homogeneity and galaxy formation. Astronomical observations show that our universe on the very large scale is extremely homogeneous. On the scale of 10^{10} light years the distribution of matter departs from perfect homogeneity by less than a part in a thousand. Why is the universe so homogeneous? On much smaller scales, on the other hand, the universe is not homogeneous at all. It contains stars, galaxies, clusters of galaxies, voids and other structures larger than 10^8 light years. What is the origin of these important inhomogeneities?

▶ The uniqueness problem. The essence of this problem was once formulated by Albert Einstein: "What I am really interested in is whether God could have created the world differently." A few years ago it would have seemed rather meaningless to ask why our space–time is four-dimensional, why α, the fine-structure constant, is close to $^1/_{137}$, why the vacuum energy density (or equivalently, the cosmological constant) is so small, and so on. However, in the Kaluza–Klein and superstring theories, space–time originally has considerably more than four dimensions, but the extra dimensions have been "compactified," shrunk into thin tubes. Therefore one may wonder why compactification stopped with four effective space–time dimensions, not two or five, or some other number. Moreover, in superstring theories there are thousands of possible compactification schemes to four dimensions, each giving different low-energy particle physics (in particular, a different value of α) and a different vacuum energy density. A similar situation may occur even without recourse to compactification. In supersymmetric grand unified theories, for example, there are dozens of stable (or practically stable) vacuum states, each corresponding to a different type of symmetry breaking between fundamental interactions, and thus to a different phenomenology of low-energy particle physics. Thus it becomes rather difficult to understand why compactification and symmetry breaking have brought about just the world we know, and no other.

All these problems (and others we have not listed) are very difficult; we had no idea how to solve them within the standard hot-universe theory. That is why it is very encouraging that



Classical evolution of φ, the homogeneous scalar field in the chaotic inflation scenario, with a quantum of mass m, much smaller than the Planck mass M_P. For $|\varphi|$ larger than $\frac{1}{5}M_P$ (but smaller than $M_P{}^2/m$), the field (represented by the red ball) rolls very slowly down toward the potential energy minimum at $\varphi = 0$. This is the inflationary epoch, during which the universe expands quasiexponentially. When $|\varphi|$ becomes smaller than $\frac{1}{5}M_P$, it oscillates rapidly around the minimum, and its energy is transformed into heat. Figure 2

most of these problems, one after another, have in the last few years been either completely resolved or considerably relaxed in the context of one comparatively simple scenario of the evolution of the universe—the inflationary scenario.

The inflationary paradigm

At present there are many different versions of the inflationary-universe scenario. The main feature of all these versions, sometimes called the inflationary paradigm, is the existence of some stage of evolution at which the universe expands exponentially (or quasiexponentially) while it is in a vacuumlike state containing some (almost) homogeneous classical fields, but no (or almost no) particles. Such an expansion is what we refer to as inflation. After inflation, the initial vacuumlike state decays into particles. They interact with one another, and after the establishment of thermodynamic equilibrium the universe becomes hot. From then on it can be described by the usual hot-universe theory. (See figure 1.)

Historically, the possible existence of an inflationary stage was first envisaged by Erast Gliner (Institute for Physics and Technology, Leningrad) about 20 years ago.[2] In the 1970s I came to the realization that homogeneous classical scalar fields φ, which are present in all unified theories of elementary particles, can play the role of an unstable vacuum state, and that their decay can heat up the universe.[3] A very interesting model was proposed in 1979 by Alexey Starobinsky (Landau Institute, Moscow), who pointed out that the regime of exponential expansion and subsequent reheating of the universe occurs in the theory of gravity when one takes account of quantum corrections.[4]

The crucial step in the development of the inflationary paradigm was taken by Alan Guth at MIT in 1980. He suggested exploiting the stage of exponential expansion in some supercooled vacuumlike state to solve simultaneously the primordial monopole problem, the flatness problem and the horizon problem.[5] His idea was very clear and attractive, but, as Guth himself pointed out, the universe after inflation in this scenario becomes too inhomogeneous.

In October 1981 I suggested an improved version of the inflation idea, which has come to be called the new inflationary scenario.[6] This resolved some of the difficulties of Guth's original formulation and made possible the solution of some of the other cosmological problems listed here. Some months later the same idea was independently proposed by Andreas Albrecht and Paul Steinhardt at the University of Pennsylvania,[7] and soon the new inflationary scenario became quite popular. However, it turns out to be very difficult to realize this scenario in the context of realistic theories of elementary particles.[8] Moreover, the orthodox version of this scenario was just a modest variation on the standard hot Big Bang theory. It was still assumed that there was an initial singularity at $t = 0$, that after the Planck time t_P (or l_P/c, about 10^{-43} seconds) the universe became hot, and that inflation was just a brief interlude in the evolution of the universe.

In 1983 I proposed another inflationary scenario, which I call the chaotic inflation scenario.[9] In my opinion, this scenario is much simpler and more natural than other versions of inflation. Therefore I shall concentrate here on the chaotic inflation scenario and on some of its recent extensions, which have given rise to the idea of an eternally existing chaotic inflationary universe.[10,11] A more technical discussion of various versions of inflation can be found in a number of review articles[8,11,12] and in some books soon to be published.[13,14]

Chaotic inflation

One of the main features of the unified particle theories is that in addition to the spinor fields ψ describing electrons, neutrinos and quarks, and the vector fields A_μ mediating their interactions, these theories also contain some scalar fields φ. Scalar fields have long been used for the phenomenological description of pions. The role of scalar fields in the new theories is more fundamental and more complicated. If, for example, the potential energy $V(\varphi)$ of a scalar field φ has one minimum at some particular value φ_1, then the whole universe gradually becomes filled with the constant classical field φ_1. Such a field is almost unobservable; it looks the same to a moving observer as it does to an observer at rest, and it mimics the appearance of empty space.

However, this field can change the masses of particles. By giving different masses to vector particles mediating different interactions, such scalar fields, referred to as Higgs fields (after Peter Higgs of the University of Edinburgh), are responsible for the symmetry breaking between the weak, strong and electromagnetic interactions in grand unified theories. Another manifestation of these fields, which will be especially important for us, is connected with their potential energy density $V(\varphi)$, which in some cases may lead to the exponential (or quasiexponential, if the field φ changes slowly) expansion of the universe.

As a simplest example, let us consider the theory of a massive scalar field φ, whose quantum is a particle of nonzero mass m, much smaller than the Planck mass. Assuming that the field is minimally coupled to gravity, its potential energy density $V(\varphi)$ is simply $\frac{1}{2}m^2\varphi^2$, as illustrated in figure 2. Suppose the field φ is initially nearly homogeneous in some domain of space–time that looks locally like an expanding universe with a growing scale factor $R(t)$ and a Hubble constant $H(t)$, defined as $\dot{R}/R$. Its evolution can then be described by the usual Klein–Gordon equation modified by cosmic expansion,

$$\ddot{\varphi} + 3H\dot{\varphi} = -m^2\varphi \qquad (1)$$

and by the Einstein equation

$$H^2 + \frac{\kappa}{R^2} = \frac{4\pi}{3M_P^2}(\dot{\varphi}^2 + m^2\varphi^2) \qquad (2)$$

Here κ is $+1$, -1 or 0, for a closed, open or flat universe, respectively. For the closed case ($\kappa = +1$), one can regard $R(t)$ as the literal radius of this universe (see PHYSICS TODAY, May, page 17). The term $3H\dot{\varphi}$ in equation 1 plays the role of a friction term in the equation of motion of the field φ. Equation 2 tells us that if φ is initially large enough, the friction term will also be large. It can be shown[9,15] that if the initial field φ_0 is greater than $\frac{1}{5}M_P$, friction makes the variation of the field φ very slow, so that one can neglect $\ddot{\varphi}$ in equation 1 and $\dot{\varphi}^2$ in equation 2. In that case any solution rapidly approaches the regime

$$\varphi(t) = \varphi_0 - \frac{mM_P}{2\sqrt{3\pi}}\, t \qquad (3)$$

$$R(t) = R_0 \exp\left(\frac{2\pi}{M_P^2}\left[\varphi_0^2 - \varphi^2(t)\right]\right) \qquad (4)$$

From these equations it follows that for times t earlier than $\varphi_0/(mM_P)$ the universe is expanding quasiexponentially:

$$R(t) = R_0 \exp(H(\varphi)t) \qquad (5)$$

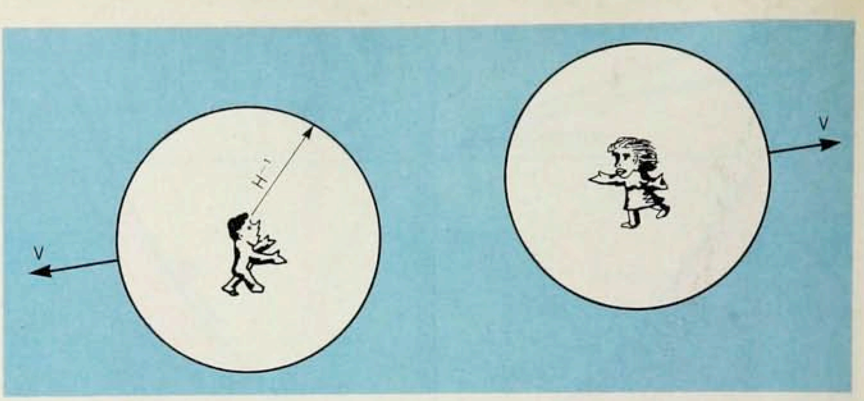
where the Hubble "constant" is given by

$$H(\varphi) \approx \sqrt{\frac{4\pi}{3}}\,\frac{m\varphi}{M_P} \qquad (6)$$

The exponent Ht is much larger than unity if φ exceeds the Planck mass in these early moments. This is just the inflationary regime we wanted to obtain.[9]

When the field φ becomes smaller than about $\frac{1}{5}M_P$, the friction term in equation 1 becomes small, and φ oscillates rapidly near its equilibrium value of zero. If this field interacts with other matter fields, these oscillations lead to abundant particle production and the heating of the universe.

The existence of an inflationary regime of this type is a general feature of a wide class of theories, including all theories in which $V(\varphi)$ varies as φ^n, higher-derivative theories of gravity, some versions of supergravity, and Kaluza–Klein theories. The role of the scalar field φ in these theories can be played by the curvature scalar, the logarithm of the radius of compactification or the like.[8,14] The remaining question to be discussed here is whether the initial conditions necessary for the realization of the inflationary regime in this scenario are sufficiently natural.[15]



Each observer in the inflationary universe can see only events that occur at distances closer than H^{-1}, the reciprocal of his present Hubble constant (times the speed of light, if one uses ordinary units). In that sense, every observer is *surrounded* by a black hole. (After a drawing by A. Linde.) Figure 3

Initial conditions

The main idea of the chaotic inflation scenario is to consider a chaotic initial distribution of the scalar field φ and to investigate its evolution without making the *ad hoc* assumption that the initial value of the field corresponds to the minimum of its potential energy $V(\varphi)$. In this sense chaotic inflation is much more general than the new inflationary universe, in which it is *assumed* that the universe *ab initio* was in thermodynamic equilibrium, and that φ was at the minimum of its temperature-dependent potential $V(\varphi,T)$.

The chaotic initial distribution of the scalar field φ is not, however, entirely unspecified. Indeed it is usually believed that for energy densities exceeding M_P^4 (or at length scales smaller than the Planck length) quantum fluctuations of the metric are so large that it becomes meaningless to speak about classical space–time; rather, one is dealing with something like a fluctuating space–time foam. Therefore when one speaks about initial conditions in some domain of a classical space–time, one automatically implies that this domain has a size greater than the 10^{-33}-cm Planck length and a density less than M_P^4. Thus $V(\varphi)$ and the squares of the spatial and temporal gradients of φ are also less than M_P^4. It follows that any initial distribution of the field φ inside any domain of a classical space–time is smooth enough that one can treat the field as essentially homogeneous within the domain.

As an example, consider first a closed universe that just appears out of the space–time foam, or from a singularity. A typical initial size of such a universe should be on the order of the Planck length M_P^{-1}, and its energy density initially should be on the order of M_P^4. Because the gradient of φ is less than M_P^2, it follows that the initial value of the field in such a universe cannot differ from some average value φ_0 by much more than M_P. One might expect that φ_0 should correspond to the minimum of $V(\varphi)$. However, fluctuations of the energy density at the moment of creation are on the order of M_P^4. Therefore the field initially "does not know" where the minimum of $V(\varphi)$ is, and later (equation 3 tells us) it rolls down to this minimum very slowly. The only constraint on the initial value of φ_0 is the requirement that $V(\varphi_0)$ cannot exceed M_P^4, implying that the typical initial value φ_0 is of order M_P^2/m, much larger than any variation of φ.

Thus a typical initial distribution of the field φ inside a closed universe is essentially homogeneous, with a potential energy density of order M_P^4. If in such a universe the initial value of $|\nabla\varphi|^2$ is a quite reasonable factor of two or three smaller than $V(\varphi)$, it can be shown that the spatial gradient of φ very quickly becomes exponentially small, and the closed universe becomes an inflationary universe described by equations 1–6, grown from a single bubble in the space–time foam.[15]

Thus we see that the initial conditions necessary for the realization of the chaotic inflation scenario in a compact (closed) universe are quite natural. In an infinite universe the situation at first glance might seem more complicated. If the initial value of φ is much larger than M_P, then, just as in the closed-universe case, φ remains rather homogeneous on a scale much larger than the Planck length, which is roughly the value of H^{-1} at that early time t. This would imply some acausal correlation between the initial values of φ in causally disconnected parts of a domain of initial size much larger than the Planck length or ct, which seems rather improbable.

The resolution of this complication is

very simple.[15] There is no reason to expect that the energy density ρ can simultaneously become smaller than $M_P{}^4$ in all causally disconnected parts of a domain much larger than $M_P{}^{-1}$, since this would imply an acausal correlation. Therefore, the typical initial size of a classical space–time domain emerging from the space–time foam must be on the order of $M_P{}^{-1}$. Such domains look like separate islands of classical space–time rising out of the foam. Outside each domain the gradient of φ is not restricted to be smaller than $M_P{}^2$, and there is no acausal correlation between the values of the field in different domains. Only later can these domains become connected.

In such a case the whole universe looks like a cluster of mini-universes, some of them inflationary. Just as in the closed-universe case, a typical initial value φ inside these mini-universes would be $M_P{}^2/m$. From equation 4 we see that by the time φ approaches zero the universe has expanded by an inflating factor

$$\exp(2\pi M_P{}^2/m^2)$$

For m on the order of $10^{-4}\,M_P$ (which, as we shall see, is necessary for eventual galaxy formation) this implies that in our simplest model the inflationary domains of the universe typically expand to $10^{10^{8}}$ times their original size!

That much inflation is more than enough to solve many of the problems of the standard hot-universe theory. After a $10^{10^{8}}$-fold expansion, the geometry of space inside the expanding domain does not differ discernibly from the Euclidean geometry of flat space, just as the surface of a vastly expanded balloon would look very much like the surface of a flat plane. Formally, this follows from the fact that the curvature term κ/R^2 in equation 2 has become negligibly small compared with H^2 after inflation. So much for the flatness problem.

The horizon problem is also solved. The observable part of the present-day universe, with a radius of order 10^{28} cm or ten billion light years, is only a small part of the domain inflated from the smallest possible beginnings—10^{-33} cm. Therefore we no longer have a problem about different parts of such a domain starting their expansion simultaneously. Furthermore, after expansion by a factor of $10^{10^{8}}$, all initial inhomogeneities, monopoles and domain boundaries have been swept beyond the horizon. This solves the primordial monopole problem and the homogeneity problem.

The scenario I have described is a general scheme of chaotic inflation as it can be understood at the level of classical physics. Usually one assumes that quantum effects are not essential for the description of something as large as our universe. Surprisingly, however, quantum effects play a very important role in the inflationary scenario.

Quantum fluctuations

According to quantum field theory, empty space is not entirely empty. It is filled with quantum fluctuations of all types of physical fields. These fluctuations can be regarded as waves of physical fields with all possible wavelengths, moving in all possible directions. If the values of these fields, averaged over some macroscopically large time, vanish, then the space filled with these fields seems to us empty and can be called the vacuum.

In the exponentially expanding universe the vacuum structure is much more complicated.[16,17] The wavelengths of all vacuum fluctuations of the scalar field φ grow exponentially with the expanding universe. When the wavelength of any particular fluctuation becomes greater than H^{-1}, this fluctuation stops propagating, and its amplitude freezes at some nonzero value $\delta\varphi(x)$ because of the large friction term $3H\dot{\varphi}$ in the equation of motion of the field φ. (The reciprocal of the Hubble constant is ordinarily an approximation to the "age" of the universe, but during inflation it remains almost constant. In our unit system, H^{-1} also denotes the length c/H.) The amplitude of this fluctuation then remains almost unchanged for a very long time, whereas its wavelength grows exponentially. Therefore the appearance of such a frozen fluctuation is equivalent to the appearance of a classical field $\delta\varphi(x)$ that does not vanish after averaging over macroscopic intervals of space and time.

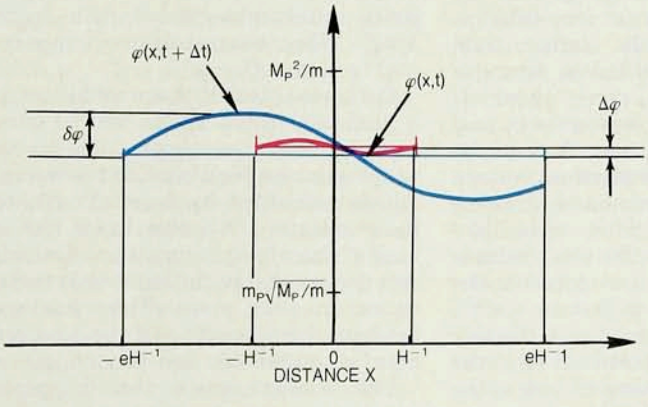
Because the vacuum contains fluctuations of all wavelengths, inflation leads to the creation of more and more new perturbations of the classical field with wavelengths greater than H^{-1}. The average amplitude of such perturbations generated during a time interval H^{-1} (in which the universe expands by a factor of e) is given by[16]

$$|\delta\varphi(x)| \approx \frac{H}{2\pi} = \frac{m\varphi}{\sqrt{3\pi}M_P} \qquad (7)$$

Perturbations of the field lead to perturbations of density that are just right for subsequent galaxy formation if m, the mass of the quantum of φ, is about $10^{-4}\,M_P$. The discovery of this mechanism, which gives a possible explanation of the origin of galaxies, was regarded in 1982 as one of the major successes of the inflationary scenario. Now it appears that the long-wave perturbations of φ are responsible not only for such local structure of the observable part of the universe, but also for the global structure of the universe.[10,11,14]

Eternal inflation

A very unusual feature of the inflationary universe is that processes separated by distances l greater than H^{-1} proceed independently of one another. This is so because during exponential expansion any two objects separated by more than H^{-1} are moving away from each other with a velocity v exceeding the speed of light. (This does not contradict special relativity because v is not the speed of any signal; it is just the rate at which the general expansion of the universe separates two distant points.) As a result, any observer in the inflationary universe can see



Evolution of φ in an inflationary domain of initial radius H^{-1}. The scalar field $\varphi(x,t)$ is nearly homogeneous because quantum perturbations have wavelengths larger than the domain. In a time interval H^{-1}, the domain has an e-fold expansion. If φ is sufficiently large, its overall decrease $\Delta\varphi$ during this interval is much smaller than the fluctuating perturbation amplitudes $\delta\varphi$. Therefore the domain volume is divided into about e^3 mini-universes of radius H^{-1}, in about half of which φ will grow. The red and blue curves show φ before and after the H^{-1} time interval.
Figure 4

only those processes occurring nearer than H^{-1}, as illustrated in figure 3.

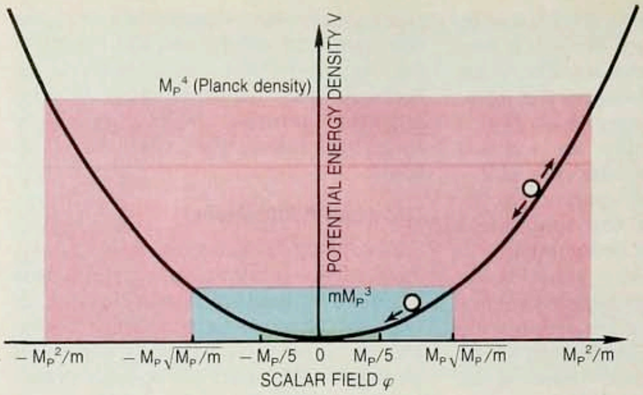
An important consequence of this general result is that the process of inflation in any spatial domain of radius H^{-1} occurs independently of any events outside it. Any two inflationary domains displaced by more than H^{-1} cannot collide or eat one another, or do each other any damage. Their expansion is due not to the annexation of the territory of their neighbors, but rather to the peaceful (and very rapid) growth of their own volume, as allowed by general relativity. In this sense any inflationary domain of initial size exceeding $2H^{-1}$ can be considered as a separate mini-universe, expanding independently of what occurs in the rest of the cosmos.

To investigate the behavior of such a mini-universe, taking account of quantum fluctuations, let us consider an inflationary domain of initial size roughly H^{-1} containing a sufficiently homogeneous field whose initial value φ greatly exceeds M_P. Equation 3 tells us that during a typical time interval $\Delta t = H^{-1}$ the field inside this domain will be reduced by

$$\Delta\varphi = \frac{M_P{}^2}{4\pi\varphi} \qquad (8)$$

If φ is much less than $\frac{1}{2}M_P\sqrt{M_P/m}$ (about 50 M_P), this decrease $\Delta\varphi$ is much larger than the amplitude of the quantum fluctuations $\delta\varphi$ generated during the same time. But for larger φ (up to the classical limit of $10^4\,M_P$), $\delta\varphi(x)$ will exceed $\Delta\varphi$. Because the typical wavelength of the fluctuation field $\delta\varphi(x)$ generated during this time is H^{-1}, the whole domain volume after Δt will effectively have become divided into e^3 separate domains (mini-universes) of diameter H^{-1}. In almost half of these domains the field φ *grows* by $|\delta\varphi(x)| - \Delta\varphi$, which is not very different from $|\delta\varphi(x)|$ or $H/2\pi$, rather than decreasing, as illustrated in figure 4. During the next time interval $\Delta t = H^{-1}$ the field grows again in half of these mini-universes. It can be shown that the total physical volume occupied by a permanently growing field φ increases with time like $\exp(3 - \ln 2)Ht$, and the total volume occupied by a field that does not decrease grows almost as fast as $\frac{1}{2}e^{3Ht}$.

Because the value of the Hubble constant $H(\varphi)$ is proportional to φ, the main part of the physical volume of the universe is the result of the expansion of domains with nearly the maximal possible field value, $M_P{}^2/m$, for which $V(\varphi)$ is close to $M_P{}^4$. There are also exponentially many domains with smaller values of φ. Those domains in which φ eventually becomes smaller than about 50 M_P give rise to the mini-universes of *our* type. In such domains, φ eventually rolls down to the minimum of $V(\varphi)$, and these mini-universes are subsequently describable by the usual Big Bang theory. However, the main part of the physical volume of the entire universe remains forever in the inflationary phase.[10]



Quantum fluctuations affect the evolution of the scalar field φ. At high field values (red) φ moves like a Brownian particle on the potential hillside $V(\varphi)$. An uphill jump is improbable, but the reward is rapid growth of the domain. Thus most of the volume of the universe comes from the expansion of domains where φ has jumped all the way up to the Planck energy density. At intermediate values (blue), the upward diffusion of φ is inefficient, and its motion is described by equation 3 in the text. Inflation takes place while φ is in the red and blue regions. When φ is small (green), it oscillates near the potential minimum. Figure 5

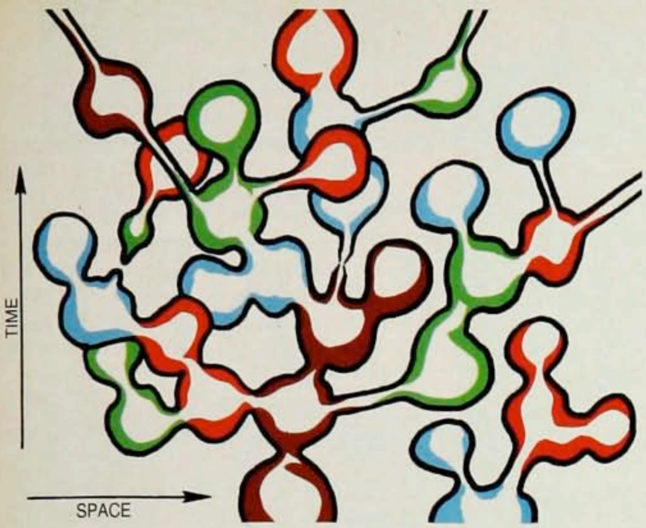
Similar results are also valid for the old[5,18] and the new inflationary scenarios,[19] in which the main part of the volume of the universe can always remain in a state corresponding to a local extremum of $V(\varphi)$ at $\varphi = 0$. In our chaotic-inflation case,[10] the results are even more surprising: Not only can the universe stay permanently on the top of a hill; it can also climb perpetually up the wall toward the largest possible values of its potential energy density, as shown in figure 5. How can it be that the universe unceasingly produces inflationary mini-universes in energetically unfavorable states with large $V(\varphi)$? What energy source supports such a process?

The answer is that the probability for a successful climb up the wall is very small indeed, but those domains in which φ jumps high enough are immediately rewarded by huge growth of their volumes. All this looks like a kind of Darwinian approach to cosmology, though one might argue that there exists an even more direct analogy between this scenario and the development of social life and economics.

The energy source that supports inflation is the gravitational energy associated with $R(t)$, the scale factor ("radius") of the universe. This gravitational energy is negative, so that the total energy of a closed universe, being the sum of the positive energy of matter and the negative gravitational energy, is zero. Just this unbounded reservoir of gravitational energy makes possible the exponentially rapid growth of the total energy of matter during inflation. The negative gravitational energy seeks any opportunity to become more negative, that is, to make inflation a nonstop process. One could say that it is the Gravitational Energy Fund that sponsors permanent production and growth of inflationary domains with large energy density.

Thus in our scenario the universe, in which there was initially at least one domain of a size on the order of H^{-1} filled with a sufficiently large and homogeneous field φ, unceasingly reproduces itself and becomes immortal. One mini-universe produces many others, and this process goes on without end, even if some of the mini-universes eventually collapse. (Everyone must die, but humanity, if it is clever enough, may exist eternally.) But this means that it is vanishingly improbable that our mini-universe would have been the first in the sequence of mini-universes illustrated in figure 6. Moreover, it no longer seems necessary to assume that there actually was some *first* mini-universe appearing from nothing or from an initial singularity at some moment $t = 0$ before which there was no space–time at all.

From general topological theorems about singularities in cosmology it does not actually follow that our universe was created *as a whole* at some moment before which the universe did not exist. The usual supposition that the whole universe appears from the unique Big Bang singularity at $t = 0$ is based on the implicit assumption that the universe *as a whole* is sufficiently homogeneous. Indeed, the observable part of our universe is very homogeneous.



Global structure of a chaotic, self-reproducing inflationary universe. Locally (out to the 10^{10}-light-year horizon) the universe looks quite homogeneous, but its global structure is complex. Mini-universes at the Planck energy density are "mutants" that may forget completely the "genetic code" (color) of their parent universe. They may even have a different space–time dimensionality. The typical thickness of a tube connecting two mini-universes after inflation is exponentially large, but if it corresponds to a compactified inflationary universe it can be as thin as the Planck length (10^{-33} cm). If the tube then evaporates by Hawking radiation, the parent and offspring mini-universes have lost their umbilical space–time connection. (After a drawing by A. Linde.) Figure 6

Observed density fluctuations are less than a part in a thousand and there has been no reason to expect that the universe is inhomogeneous on a larger scale beyond the 10^{10}-light-year horizon. In a homogeneous universe one can use the density $\rho(t)$ as a measure of time. In that case it can be shown that the universe appears *as a whole* from a singularity at $t = 0$. The initial density $\rho(0)$ is infinite, and it becomes possible to describe the whole universe in terms of classical space–time after the Planck time M_P^{-1} (about 10^{-43} seconds), when the energy density everywhere simultaneously becomes smaller than the Planck density M_P^4.

With the invention of the inflationary scenario the situation changes drastically. At present only inflation can explain why the observable part of the universe is so homogeneous, but from inflation it also follows that on a much larger scale the universe is extremely *in*homogeneous. In some parts of the universe the energy density ρ is now on the order of M_P^4, 125 orders of magnitude higher than the 10^{-29} or 10^{-30} g/cm^3 we can (or hope to) see. In such a scenario there is no reason to assume that the universe was initially homogeneous and that all its causally disconnected parts started their expansions simultaneously.

If the universe is infinitely large (like the Friedmann open or flat universe), then it cannot have had a single beginning; a simultaneous creation of infinitely many causally disconnected regions is totally improbable. Therefore the universe cannot be infinite *ab initio*, or it must exist eternally as a huge self-reproducing entity. Some of its parts appear at different times from singularities, or may die in a singular state. New parts are constantly being created from the space–time foam when $V(\varphi)$ exceeds the Planck density, or they may revert to the foamlike state again as a result of large fluctuations in φ. But the evolution of the universe as a whole has no end, and it may have had no beginning.

Mutation

Darwin having been mentioned, the reader may ask: Does the universe simply reproduce itself, or are mutations also possible? To answer this question, we recall that in realistic theories of elementary particles there exist many different types of scalar fields φ_i. The potential energy $V(\varphi_i)$ often has many different local minima, in which the universe may live for an extremely long time, much greater than the 10^{10} years of our observable domain. For example, in the supersymmetric SU(5) theory, $V(\varphi_i)$ has several different minima of almost equal depth. Because the laws governing the interactions of elementary particles at the low energies at which we do experiments depend on the values of the classical fields φ_i, each of these minima corresponds to a different low-energy physics. In one of them the SU(5) symmetry between all types of interactions remains unbroken—that is, the scalar fields φ_i remain equal to zero. In other minima various symmetry breaking patterns are realized, and in only one of these minima is the broken symmetry of the weak, strong and electromagnetic interactions that which we in fact observe.

During inflation there are large-scale fluctuations of all the fields φ_i. As a result, the inflationary universe becomes divided into an exponentially large number of inflationary mini-universes, with the scalar fields taking all possible values. As inflation ends in some mini-universes, these scalar fields roll down to all possible minima of $V(\varphi_i)$. The universe becomes divided into many different exponentially large domains, realizing all possible types of symmetry breaking between the fundamental interactions. In some of these mini-universes the low-energy physics is quite different from our own. We cannot now see them because the size of our own domain is much greater than the size of its 10^{10}-light-year observable portion. We could not live in these domains because our kind of life requires our kind of low-energy physics. Therefore I would not recommend crossing the wall to a neighboring domain if it should appear over the horizon in the distant future.

It is very important that in the inflationary universe there is lots of room for all possible types of symmetry breaking and for all possible types of life. There is, therefore, no longer any need to require that in the true theory the minimum of $V(\varphi_i)$ corresponding to our type of symmetry breaking be the only one or the deepest one. This new cosmopolitan viewpoint may greatly simplify the task of building realistic models of the elementary particles.

The change of the values of the scalar fields φ_i—that is to say, the change of the vacuum state—is the simplest kind of mutation that may occur during inflation. Much more interesting possibilities appear if one considers chaotic inflation in the higher-dimensional Kaluza–Klein theories. In those domains in which the energy density of the field φ grows to the Planck density, quantum fluctuations of the metric at a length scale of M_P^{-1} become of order unity. In such domains an inflationary d-dimensional universe can squeeze locally into a tube of smaller dimensionality $d - n$ (or vice versa), as illustrated in figure 7. If this tube is also inflationary (in $d - n$ dimensions) and the initial length of the tube is greater than M_P^{-1} (which is quite probable near the Planck density), then its further expansion proceeds independently of its prehistory and of the fate of its mother universe. In an eternally existing universe such processes should occur even if their probability is very small. In fact the probability of such processes is small only if $V(\varphi)$ is far below M_P^4.

Thus the inflationary universe becomes divided into different mini-universes in which all possible types of compactification produce all sorts of dimensionalities.[10,11] By this argument we find ourselves inside a four-dimensional domain with our kind of low-energy physics not because other kinds of mini-universes are impossible or improbable, but simply because our kind of life cannot exist in other domains.

This may have important implications for the building of realistic Kaluza–Klein and superstring theories.[14] For example, it is extremely complicated, if not impossible, to construct a theory in which only one type of compactification can occur, leading precisely to a four-dimensional inflationary universe with the low-energy particle physics of our experience. But from the point of view discussed here, there is no need to require that the results compactification and inflation have wrought in our realm be the only possible results, or the best. It is enough to find a theory in which such a compactification is possible. This problem is still difficult, but it is much easier than the one we have been trying to solve.

Cosmology and particle physics

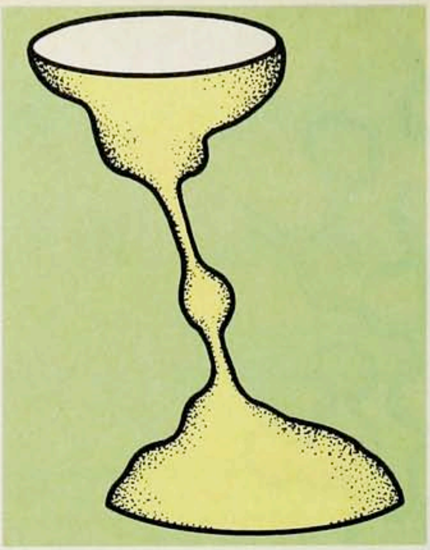
One can compare cosmologists and particle physicists to two groups of people tunneling through a huge mountain from opposite sides. This analogy is not quite exact. If the two groups fail to meet they will simply make two tunnels instead of one. But if the particle physicists do not meet the cosmologists on their way through the mountain of the unknown, they will not have any complete theory at all.

The inflationary-universe scenario is a spectacular manifestation of the interplay between elementary-particle theory and cosmology. Of course inflation is not a magic word that will automatically solve all our problems and open all doors. In some theories of elementary particles it is difficult to realize inflation, whereas many other theories do not lead to a good cosmology even with the help of the inflation. At the moment, however, it does seem necessary to have something like inflation to obtain a consistent cosmology at peace with particle physics.

The inflationary scenario is still developing rapidly. We are witnessing a gradual change in the overall viewpoint of cosmology. Just a few years ago there was no doubt that the universe was born in a single Big Bang singularity. It seemed absolutely clear that space–time was four-dimensional from the start, and that it remains four-dimensional everywhere. It was believed that if the universe is closed its total size is on the order of its observable part, 10^{28} cm, and that in about 10^{11} years the whole universe will collapse and disappear in a Big Crunch. If, on the other hand, the universe is flat or open, then it is infinite, and it was commonly believed that its properties everywhere are approximately the same. Such a universe would exist without end, but after the decay of its protons, predicted by the unified elementary-particle theories, there would be no baryonic matter to support life of our type. The only possible choice seemed to be between the "hot ending" in the Big Crunch and the "cold ending" in infinite empty space.

Now it seems more likely that the universe is an eternally existing, self-producing entity, that it is divided into many mini-universes much larger than our observable portion, and that the laws of low-energy physics and even the dimensionality of space–time may be different in each of these mini-universes. This modification of our picture of the universe and of our place in it is one of the most important consequences of the inflationary scenario.

We have even gained a new understanding of why it was necessary to write a scenario when the performance is already over. The answer is that the performance is still going on, and it will continue eternally. In different parts of the universe different observers see its endless variations. We cannot see the whole play in all its greatness, but we can try to imagine its most essential parts, and perhaps ultimately understand its meaning.



Local compactification or decompactification of the inflationary universe alters the apparent dimensionality of space–time in a particular domain. The radius of a compactified tube can fluctuate with fluctuations of the scalar field φ. (After a drawing by A. Linde.) Figure 7

References

1. S. Weinberg, *The First Three Minutes*, Basic, New York (1977).
2. E. B. Gliner, Zh. Eksp. Teor. Fiz. **49**, 542 (1965) [Sov. Phys. JETP **22**, 378 (1965)]; Dokl. Akad. Nauk SSSR, **192**, 771 (1970) [Sov. Phys. Dokl. **15**, 559 (1970)].
3. A. D. Linde, Pis'ma Zh. Eksp. Teor. Fiz. **19**, 320 (1974) [JETP Lett. **19**, 183 (1974)]; Rep. Progr. Phys. **42**, 389 (1979).
4. A. A. Starobinsky, Pis'ma Zh. Eksp. Teor. Fiz. **30**, 719 (1979) [JETP Lett. **30**, 682 (1979)].
5. A. H. Guth, Phys. Rev. D **23**, 347 (1981).
6. A. D. Linde, Phys. Lett. B **108**, 389 (1982).
7. A. Albrecht, P. J. Steinhardt, Phys. Rev. Lett. **48**, 1220 (1982). A. H. Guth, P. J. Steinhardt, Sci. Am., May 1984, p. 90.
8. A. D. Linde, Rep. Prog. Phys. **47**, 925 (1984).
9. A. D. Linde, Phys. Lett. **129B**, 177 (1983).
10. A. D. Linde, Mod. Phys. Lett. A **1**, 81 (1986); Phys. Lett. B **175**, 395 (1986). A. S. Goncharov, A. D. Linde, V. F. Mukhanov, J. Mod. Phys. A **2**, 561 (1987).
11. A. D. Linde, in *300 Years of Gravitation*, S. W. Hawking, W. Israel, eds., Cambridge U. P., Cambridge, England (1987).
12. R. H. Brandenberger, Rev. Mod. Phys. **57**, 1 (1985). M. S. Turner, in *Proc. Cargese School on Fundamental Physics and Cosmology*, J. Audouze, J. Tran Thanh Van, eds., Editions Frontières, Gif-sur-Yvette, France (1986). S. K. Blau, A. H. Guth, in *300 Years of Gravitation*, S. W. Hawking, W. Israel, eds., Cambridge U. P., Cambridge, England (1987).
13. R. H. Brandenberger, *Inflation and Cosmic Strings*, World Scientific, Singapore (1987).
14. A. D. Linde, *Particle Physics and Inflationary Cosmology*, Gordon and Breach, New York, to be published.
15. A. D. Linde, Phys. Lett. B **162**, 281 (1985).
16. A. Vilenkin, L. Ford, Phys. Rev. D **26**, 1231 (1982). A. D. Linde, Phys. Lett. B **116**, 335 (1982).
17. V. F. Mukhanov, G. V. Chibisov, Pis'ma Zh. Eksp. Teor. Fiz. **10**, 549 (1981) [JETP Lett. **33**, 532 (1981)]. S. W. Hawking, Phys. Lett. B **115**, 295 (1982). A. A. Starobinsky, Phys. Lett. B **117**, 175 (1982). A. H. Guth, S.-Y. Pi, Phys. Rev. Lett. **49**, 1110 (1982). J. Bardeen, P. J. Steinhardt, M. S. Turner, Phys. Rev. D **28**, 679 (1983).
18. J. R. Gott, Nature **295**, 304 (1982). K. Sato, H. Kodama, M. Sasaki, K. Maeda, Phys. Lett. B **108**, 35 (1982).
19. P. J. Steinhardt, in *The Very Early Universe*, G. W. Gibbons, S. W. Hawking, S. Siklos, eds., Cambridge U. P., Cambridge, England (1982), p. 251. A. D. Linde, *Nonsingular Regenerating Inflationary Universe*, Cambridge University preprint (1982). A. Vilenkin, Phys. Rev. D **27**, 2848 (1983). □