Current understanding of inflation

Andrei Linde

Department of Physics, Stanford University, Stanford, CA 94305, USA

Available online 16 February 2005

Abstract

I will discuss the development of inflationary theory and its present status, including recent progress in describing de Sitter space, dark energy and inflationary universe in string theory.

© 2005 Published by Elsevier B.V.

1. Introduction

Textbooks on astrophysics usually describe inflation as an exponentially rapid expansion of the universe in a supercooled vacuum state (false vacuum), which was formed as a result of the high temperature cosmological phase transitions breaking symmetry between the weak, strong and electromagnetic interactions. Usually the textbooks do not mention that none of the theories of this type were found to correctly describe current observational data, and therefore 20 years ago the basic principles of inflationary cosmology have been changed dramatically. In this review, I will briefly describe the history and the present status of inflationary cosmology.

The first model of inflationary type was proposed by Starobinsky (1980). It was based on investigation of conformal anomaly in quantum gravity. This model did not suffer from the graceful exit problem, but it was rather complicated and did not aim on solving homogeneity, horizon and monopole problems.

A much simpler inflationary model with a very clear physical motivation was proposed by Guth (1981). His model, which is now called “old inflation,” was based on the theory of supercooling during the cosmological phase transitions (Kirzhnits and Linde, 1976). The universe was expanding exponentially while the scalar field was captured in the false vacuum state at the top of the effective potential. This scenario played a profound role in the development of inflationary cosmology because it clearly explained how inflation may solve the major cosmological problems. Its main difficulty was related to the false vacuum decay, which occurs due to bubble formation. If the bubbles are formed near each other, their collisions make the universe extremely inhomogeneous. If they are formed far away from each other, each of them...
represents a separate open universe with a vanishingly small $\Omega$. As Guth pointed out, both options are unacceptable (Guth, 1981).

This problem was resolved with the invention of the new inflationary theory (Linde, 1982; Albrecht and Steinhardt, 1982). In this theory, just as in old inflation, the stage of inflation may begin either in the false vacuum, or in an unstable state at the top of the effective potential. Then the inflaton field $\phi$ slowly rolls down to the minimum of its effective potential, and inflation still continued at that stage. This stage is of crucial importance: density perturbations produced during the slow-roll inflation are inversely proportional to $\phi$ (Mukhanov and Chibisov, 1981; Hawking, 1982; Starobinsky, 1982; Guth and Pi, 1982; Bardeen et al., 1983; Mukhanov, 1985; Mukhanov et al., 1992). Thus the key difference between the new inflationary scenario and the old one is that the useful part of inflation, which is responsible for the homogeneity of our universe in the new inflation scenario, does not occur in the false vacuum state, where $\phi = 0$.

The new inflation scenario was plagued by its own problems. Density perturbations generated in its original versions were unacceptably large. One could avoid this problem by making the coupling constant of the scalar field extremely small. But then the inflaton field could not be in a state with a mass $m$ and with the potential energy density $V(\phi) = \frac{m^2}{2} \phi^2$. Since this function has a minimum at $\phi = 0$, one may expect that the scalar field $\phi$ should oscillate near this minimum. This is indeed the case if the universe does not expand, in which case equation of motion for the scalar field coincides with equation for harmonic oscillator, $\ddot{\phi} = -m^2 \phi$.

However, because of the expansion of the universe with Hubble constant $H = \dot{a}/a$, an additional term $3H\dot{\phi}$ appears in the harmonic oscillator equation:

$$\ddot{\phi} + 3H\dot{\phi} = -m^2 \phi. $$

The term $3H\dot{\phi}$ can be interpreted as a friction term. The Einstein equation for a homogeneous universe containing scalar field $\phi$ looks as follows:

$$H^2 + \frac{k}{a^2} = \frac{1}{6} \left( \phi^2 + m^2 \phi^2 \right).$$

Here $k = -1$, 0, 1 for an open, flat or closed universe respectively. We work in units $M_p^{-2} = 8\pi G = 1$.

If the scalar field $\phi$ initially was large, the Hubble parameter $H$ was large too, according to the second equation. This means that the friction term $3H\dot{\phi}$ was very large, and therefore the scalar field was moving very slowly, as a ball in a viscous liquid. Therefore, at this stage the energy density of the scalar field, unlike the density of ordinary matter, remained almost constant, and expansion of the universe continued with a much greater speed than in the old cosmological theory. Due to the rapid growth of the scale of the universe and a slow motion of the field $\phi$, soon after the beginning of this regime one has $\dot{\phi} \ll 3H\phi$, $H^2 \gg \frac{k}{a^2}$, $\dot{\phi}^2 \ll m^2\phi^2$, so the system of equations can be simplified...

2. Chaotic inflation

Consider the simplest model of a scalar field $\phi$ with a mass $m$ and with the potential energy density $V(\phi) = \frac{m^2}{2} \phi^2$. Since this function has a minimum at $\phi = 0$, one may expect that the scalar field $\phi$ should oscillate near this minimum. This is indeed the case if the universe does not expand, in which case equation of motion for the scalar field coincides with equation for harmonic oscillator, $\ddot{\phi} = -m^2 \phi$.

Moreover, old and new inflation represented a substantial but incomplete modification of the big bang theory. It was still assumed that the universe was in a state of thermal equilibrium from the very beginning, that it was relatively homogeneous and large enough to survive until the beginning of inflation, and that the stage of inflation was just an intermediate stage of the evolution of the universe. In the beginning of the 1980s these assumptions seemed natural and practically unavoidable. That is why it was so difficult to overcome a certain psychological barrier and abandon all of these assumptions. This was done with the invention of the chaotic inflation scenario (Linde, 1983; Linde, 1990).
The first equation shows that if the field $\phi$ changes slowly, the size of the universe in this regime grows approximately as $e^{Ht}$, where $H = \frac{m}{\sqrt{6}}$. This is the stage of inflation, which ends when the field $\phi$ becomes much smaller than $M_p = 1$.

Thus, inflation does not require supercooling and tunnelling from the false vacuum (Guth, 1981), or rolling from an artificially flat top of the effective potential (Linde, 1982; Albrecht and Steinhardt, 1982). It appears in the theories that can be as simple as a theory of a harmonic oscillator (Linde, 1983; Linde, 1990).

The first models of chaotic inflation were based on the theories with 1 polynomial potentials, such as $V(\phi) = -\frac{1}{2}m^2\phi^2 + \frac{1}{4}\phi^4$. But the main idea of this scenario is quite generic. One should consider any particular potential $V(\phi)$, polynomial or not, with or without spontaneous symmetry breaking, and study all possible initial conditions without assuming that the universe was in a state of thermal equilibrium, and that the field $\phi$ was in the minimum of its effective potential from the very beginning.

This idea strongly deviated from the standard lore of the hot big bang theory and was psychologically difficult to accept. It took several years until it finally became clear that the idea of chaotic initial conditions is most general, and it is much easier to construct a consistent cosmological theory without making unnecessary assumptions about thermal equilibrium and high temperature phase transitions in the early universe.

### 3. Shift symmetry and chaotic inflation in supergravity

In the simplest versions of chaotic inflation scenario with the potentials $V \sim \phi^n$, the process of inflation occurs at $\phi > 1$, in Planck units. However, for a long time it seemed very difficult to realize inflation in supergravity at $\phi > 1$. The effective potential of the complex scalar field $\Phi$ in supergravity is

$$V = e^K \left[ K_{\phi \bar{\phi}} |D_{\phi} W|^2 - 3|W|^2 \right].$$

Here $W(\Phi)$ is the superpotential, $\Phi$ denotes the scalar component of the superfield $\Phi$; $D_{\Phi} W = \frac{\partial W}{\partial \Phi} + \frac{i}{2} \epsilon_{\Phi} W$. The kinetic term of the scalar field is given by $K_{\phi \bar{\phi}} \partial_{\phi} \partial_{\bar{\phi}} \Phi$. The standard textbook choice of the Kähler potential corresponding to the canonically normalized fields $\Phi$ and $\phi$ is $K = \phi \bar{\phi}$, so that $K_{\phi \bar{\phi}} = 1$. This immediately reveals a problem: At $\phi > 1$ the potential is extremely steep. It blows up as $e^{\phi^2}$, which makes it difficult to realize chaotic inflation in supergravity at $\phi \equiv \sqrt{2}/|\phi| > 1$. Moreover, the problem persists even at small $\phi$. If, e.g., one considers the simplest case when the superpotential does not depend on the inflaton field $\phi$, then Eq. (4) implies that at $\phi \ll 1$ the effective mass of the inflaton field is $m^2 = 3H^2$. This violates the condition $m^2 \ll H^2$ required for successful slow-roll inflation (so-called $\eta$-problem).

Some progress in SUGRA inflation during the last decade was achieved in the context of the models of the hybrid inflation type (Linde, 1991; Linde, 1994), where chaotic inflation may occur at $\phi \ll 1$. Among the best models are the F-term inflation, where different contributions to the effective mass term $m^2$ cancel (Copeland et al., 1994; Dvali et al., 1994; Linde and Riotto, 1997), and D-term inflation (Binetruy and Dvali, 1996; Halyo, 1996), where the dangerous term $e^K$ does not affect the potential in the inflaton direction. A recent version of this scenario, P-term inflation, which unifies F-term and D-term models, was proposed in (Kallosh and Linde, 2003).

A natural realization of the simplest version of the chaotic inflation model in supergravity was proposed only few years ago. The idea is to take the Kähler potential $K = \frac{1}{2}(\phi + \bar{\phi})^2 + X \bar{X}$ of the fields $\Phi$ and $X$, with the superpotential $m\Phi X$ (Kawasaki et al., 2000). This Kähler potential can be obtained from the standard potential $K = \phi \bar{\phi} + X \bar{X}$ by adding terms $\phi^2/2 + \bar{\phi}^2/2$, which do not give any contribution to the kinetic term of the scalar fields $K_{\phi \bar{\phi}} \partial_{\phi} \partial_{\bar{\phi}} \Phi$. In other words, the new Kähler potential leads to canonical kinetic terms for the fields $\Phi$ and $X$, so it is as simple and legitimate as the standard textbook Kähler potential. However, the new Kähler potential has shift symmetry: it does not depend on the imaginary part
of the field $\Phi$. The shift symmetry is broken only by the superpotential.

As a result, the dangerous term $e^K$ continues growing exponentially in the direction $(\Phi + \tilde{\Phi})$, but it remains constant in the direction $(\Phi - \tilde{\Phi})$.

Decomposing the complex field $\Phi$ into two real scalar fields, $\Phi = \frac{1}{\sqrt{2}} (\eta + i\phi)$, one can find the resulting potential $V(\phi, \eta, X)$ for $\eta, |X| \ll 1$: $V = e^\frac{\phi^2}{2} [1 + \eta^2] + m^2|X|^2$. This potential has a deep valley, with a minimum at $\eta = X = 0$. The fields $\eta$ and $X$ rapidly fall down towards $\eta = X = 0$, after which the potential for the field $\phi$ becomes $V = e^\frac{\phi^2}{2}$. This provides a very simple realization of the chaotic inflation scenario in supergravity (Kawasaki et al., 2000).

It is amazing that for almost 20 years nothing but inertia was keeping us from using the version of the supergravity which was free from the famous $\eta$ problem. As we will see shortly, the situation with inflation in string theory is very similar, and may have a similar resolution.

4. Towards inflation in string theory

4.1. de Sitter vacua in string theory

For a long time, it seemed rather difficult to obtain inflation in M/string theory. The main problem here was the stability of compactification of internal dimensions. For example, ignoring non-perturbative effects to be discussed below, a typical effective potential of the effective 4d theory obtained by compactification in string theory of type IIB can be represented in the following form:

$$V(\sigma, \rho, \phi) \sim e^{2\sigma}\sqrt{6}\bar{V}(\phi).$$

Here $\sigma$ and $\rho$ are canonically normalized fields representing the dilaton field and the volume of the compactified space; $\phi$ stays for all other fields.

If $\sigma$ and $\rho$ were constant, then the potential $\bar{V}(\phi)$ could drive inflation. However, this does not happen because of the steep exponent $e^{2\sigma}\sqrt{6}$, which rapidly pushes the dilaton field $\sigma$ to $-\infty$, and the volume modulus $\rho$ to $+\infty$. As a result, the radius of compactification becomes infinite; instead of inflating, 4d space decompactifies and becomes 10d.

Thus, in order to describe inflation one should first learn how to stabilize the dilaton and the volume modulus. The dilaton stabilization was achieved in Giddings et al. (2002). The most difficult problem was to stabilize the volume. The solution of this problem was found in Kachru et al. (2003a) (KKLT construction). It consists of two steps.

First of all, due to a combination of effects related to warped geometry of the compactified space and nonperturbative effects, it is possible to obtain a supersymmetric AdS minimum of the effective potential for $\rho$. This fixed the volume modulus, but in a state with a negative vacuum energy. Then one adds an anti-D3 brane with the positive energy $\sim \rho^{-2}$. This addition uplifts the minimum of the potential to the state with a positive vacuum energy.

Instead of adding an anti-D3 brane, one can also add a D7 brane with fluxes. This results in the appearance of a D-term which has a similar dependence on $\rho$, but leads to spontaneous supersymmetry breaking (Burgess et al., 2003). In either case, one ends up with a metastable dS state which can decay by tunnelling and formation of bubbles of 10d space with vanishing vacuum energy density. The decay rate is extremely small (Kachru et al., 2003a), so for all practical purposes, one obtains an exponentially expanding de Sitter space with the stabilized volume of the internal space.

This scenario at present represents the only existing model of dark energy derived in the context of string theory. It corresponds to the (meta-stable) de Sitter state, with the equation of state $w = -1$. Even though it might be possible to obtain the models of dark energy with $w \neq 1$, no models of such type have been proposed so far.

4.2. Inflation in string theory and shift symmetry

During the last few years there were many suggestions how to obtain hybrid inflation in string theory by considering motion of branes in the compactified space (see Dvali and Tye, 1999;
Can we avoid fine-tuning altogether? One of the possible ideas is to find theories with some kind of shift symmetry. Another possibility is to construct something like D-term inflation, where the flatness of the potential is not spoiled by the term $e^K$. Both of these ideas were explored in a recent paper (Hsu et al., 2003) based on the model of D3/D7 inflation in string theory (Kallosh, 2001; Herdeiro et al., 2001; Dasgupta et al., 2002). In this model the Kahler potential is given by $K = -3 \log(\rho + \bar{\rho} - k(\phi, \bar{\phi}))$, where the function $k(\phi, \bar{\phi})$ for the inflaton field $\phi$, at small $\phi$, was taken in the simplest form $k(\phi, \bar{\phi}) = \phi \bar{\phi}$. If one makes the simplest assumption that the superpotential does not depend on $\phi$, then the $\phi$ dependence of the potential (4) comes from the term $e^K = (\rho + \bar{\rho} - \phi \bar{\phi})^{-3}$. Expanding this term near the stabilization point $\rho = 0$, one finds that the inflaton field has a mass $m_\phi^2 = 2H^2$. Just like the similar relation $m_\phi^2 = H^2$ in the simplest models of supergravity, this is not what we want for inflation.

One way to solve this problem is to consider $\phi$-dependent superpotentials. By doing so, one may fine-tune $m_\phi^2$ to be $O(10^{-2})H^2$ in a vicinity of the point where inflation occurs (Kachru et al., 2003b). Whereas fine-tuning is certainly undesirable, in the context of string cosmology it may not be a serious drawback. Indeed, if there exist many realizations of string theory (Bousso and Polchinski, 2000; Susskind, 2003; Douglas, 2003), then one might argue that all realizations not leading to inflation can be discarded, because they do not describe the universe in which we could live. Meanwhile, those non-generic realizations, which lead to eternal inflation, describe inflationary universes with an indefinitely large and ever-growing volume of inflationary domains. This makes the issue of fine-tuning less problematic.\(^2\)

\(^2\) One should note that decreasing of $m_\phi^2$ is not the only way to get inflation; one may reach the same goal by considering theories with non-minimal kinetic terms (see e.g., Armendariz-Picon et al., 1999; Dimopoulos and Thomas, 2003; Silverstein and Tong, 2004).
theory, which remains valid even after the KKLT volume stabilization. The answer to this question was found only very recently, and it appears to be model-dependent. It was shown in Hsu and Kallosh (2004) that in a certain class of models, including D3/D7 models (Kallosh, 2001; Herdeiro et al., 2001; Dasgupta et al., 2002; Hsu et al., 2003; Angelantonj et al., 2004; Koyama et al., 2004), the shift symmetry of the effective 4d theory is not an assumption but an unambiguous consequence of the underlying mathematical structure of the theory. This may allow us to obtain a natural realization of inflation in string theory.

5. Eternal inflation and stringy landscape

Even though we are still at the very first stages of implementing inflation in string theory, it is very tempting to speculate about possible generic features and consequences of such a construction.

First of all, KKLT construction shows that the vacuum energy after the volume stabilization is a function of many different parameters in the theory. One may wonder how many different choices do we actually have. There were several attempts to investigate this issue, counting different flux vacua (Bousso and Polchinski, 2000; Douglas, 2003). The possible numbers, depending on specific assumptions, may vary in the range from $10^{20}$ to $10^{1000}$. Some of these vacuum states with positive vacuum energy can be stabilized using the KKLT approach. Each of such states will correspond to a metastable vacuum state.

Where is our place in this enormously large landscape of all possible vacuum states? In order to answer this question one may try to use anthropic considerations. However, anthropic principle leads to unambiguous predictions only if for some of the possible vacuum states the probability of emergence of life of our type vanishes. It is much more difficult to make predictions in the situations where life may appear albeit with a very small probability. Indeed, in an eternally expanding infinitely large universe life may eventually appear even if the probability of such event is very small. In order to make probabilistic predictions based on the anthropic principle one would need to have an unambiguous measure of probability; different points of view on this issue can be found (e.g., Linde, 1990; Linde et al., 1994; Garcia-Bellido et al., 1994; Garriga and Vilenkin, 2001).

I do believe that anthropic principle is a simple and powerful superselection rule which should have its place in the toolbox of every physicist. However, one should use it with caution, and it may not be restrictive enough to explain all properties of the world. After all, its main idea is that we can live only in those parts of the universe where we can survive. Some people even call anthropic principle the principle of mediocrity. This does not sound very inspir-}

ing. It would be nice to find an explanation of some beautiful features of this word, such as the existence of various symmetries, which may or may not be necessary for our survival.

One of the physical mechanisms that may help us to explain some of the features of our world is related to the moduli trapping near the points of enhanced symmetry (Kofman et al., 2004). In supersymmetric theories, and in string theory describing brane interactions, one often has potentials with flat directions, which sometimes intersect. Excitations of the fields along the flat directions (moduli fields) have small or even vanishing mass. At the intersections, the number of the massless particles increases, which corresponds to enhancement of symmetry.

An interesting effect found in Kofman et al. (2004) is that the motion of one of the fields along its flat direction leads to particle production of the second field when the first one reaches the intersection. These particles have masses proportional to the degree of deviation of the first field from the intersection. As a result, these particles produce a force which returns the first field to the intersection and traps both of the fields there. This effect provides a dynamical mechanism which brings the fields to the position with large number of symme-

tries. Since symmetry is often associated with beauty, one can summarize the nature of this effect in a simple and intuitive way: beauty is attractive.

In this scenario, inflation leads to creation of a chaotic eternally expanding universe consisting of exponentially large number of domains with different properties (Linde et al., 1994; Garcia-Bellido et al., 1994). Some of these domains can be
selected by the anthropic principle as the parts of the universe where life as we know it can exist. Then some dynamical mechanisms, such as the moduli trapping mechanism described above, might help us not only to survive but to live well, by making our world more symmetric and beautiful.

References

Hsu, J.P., Kallosh, R. 2004. JHEP 0404, 042.